

# Dossier zur Nutzenbewertung gemäß § 35a SGB V

*Luspatercept (Reblozyl®)*

Celgene GmbH

## **Modul 4 B**

*Erwachsene Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind*

Medizinischer Nutzen und  
medizinischer Zusatznutzen,  
Patientengruppen mit therapeutisch  
bedeutsamem Zusatznutzen

# Inhaltsverzeichnis

	Seite
<b>Tabellenverzeichnis</b> .....	<b>4</b>
<b>Abbildungsverzeichnis</b> .....	<b>11</b>
<b>Abkürzungsverzeichnis</b> .....	<b>23</b>
<b>4 Modul 4 – allgemeine Informationen</b> .....	<b>26</b>
4.1 Zusammenfassung der Inhalte von Modul 4.....	27
4.2 Methodik .....	52
4.2.1 Fragestellung .....	52
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung .....	53
4.2.3 Informationsbeschaffung .....	55
4.2.3.1 Studien des pharmazeutischen Unternehmers .....	55
4.2.3.2 Bibliografische Literaturrecherche .....	55
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken .....	56
4.2.3.4 Suche auf der Internetseite des G-BA.....	58
4.2.3.5 Selektion relevanter Studien .....	59
4.2.4 Bewertung der Aussagekraft der Nachweise.....	60
4.2.5 Informationssynthese und -analyse .....	63
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien .....	63
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	63
4.2.5.2.1 Patientencharakteristika .....	64
4.2.5.2.2 Patientenrelevante Endpunkte .....	66
4.2.5.3 Meta-Analysen.....	91
4.2.5.4 Sensitivitätsanalysen .....	92
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren .....	93
4.2.5.6 Indirekte Vergleiche .....	100
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen .....	102
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	102
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel .....	102
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	102
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	104
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken .....	106
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA .....	108
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	109
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	110
4.3.1.2.1 Studiendesign und Studienpopulationen .....	110
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	128
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	129
4.3.1.3.1 Endpunkte – RCT.....	130

4.3.1.3.1.1	Mortalität – Gesamtüberleben .....	133
4.3.1.3.1.2	Morbidität – Transfusionsfreiheit .....	138
4.3.1.3.1.3	Morbidität – Reduktion der Transfusionslast .....	153
4.3.1.3.1.4	Morbidität – Reduktion der Eisenüberladung .....	163
4.3.1.3.1.5	Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens .....	169
4.3.1.3.1.6	Morbidität – Hospitalisierungen .....	193
4.3.1.3.1.7	Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens .....	208
4.3.1.3.1.8	Lebensqualität gemessen anhand des QoL-E-Fragebogens .....	228
4.3.1.3.1.9	Sicherheit .....	247
4.3.1.3.2	Subgruppenanalysen – RCT .....	285
4.3.1.4	Liste der eingeschlossenen Studien - RCT .....	314
4.3.2	Weitere Unterlagen .....	315
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien .....	315
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche .....	315
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche .....	315
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen .....	316
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT .....	316
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT .....	319
4.3.2.1.4	Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT .....	319
4.3.2.2	Nicht randomisierte vergleichende Studien .....	319
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien .....	319
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien .....	320
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien .....	321
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien .....	321
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien .....	322
4.3.2.2.4	Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien .....	323
4.3.2.3	Weitere Untersuchungen .....	323
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen .....	323
4.3.2.3.2	Charakteristika der weiteren Untersuchungen .....	324
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen .....	324
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen .....	324
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen .....	325
4.3.2.3.4	Liste der eingeschlossenen Studien – weitere Untersuchungen .....	325
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens .....	326
4.4.1	Beurteilung der Aussagekraft der Nachweise .....	326
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß .....	328
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht .....	346
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte .....	346
4.5.1	Begründung für die Vorlage indirekter Vergleiche .....	346

4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	346
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen .....	347
4.5.4	Verwendung von Surrogatendpunkten.....	347
4.6	Referenzliste.....	349
<b>Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche .....</b>		<b>358</b>
<b>Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken.....</b>		<b>362</b>
<b>Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....</b>		<b>364</b>
<b>Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken).....</b>		<b>365</b>
<b>Anhang 4-E : Methodik der eingeschlossenen Studien – RCT .....</b>		<b>373</b>
<b>Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten .....</b>		<b>390</b>
<b>Anhang 4-G : Kaplan-Meier-Kurven der Zeit bis zum ersten UE auf Ebene der SOC und PT.....</b>		<b>417</b>
<b>Anhang 4-H : Ergänzende Subgruppenanalysen für statistisch nicht signifikante Interaktionstests.....</b>		<b>474</b>

**Tabellenverzeichnis**

	<b>Seite</b>
Tabelle 4-1: Zusammenfassung der Therapieeffekte von Luspatercept .....	37
Tabelle 4-2: Kriterien für den Ein- und Ausschluss von Studien in die Nutzenbewertung .....	54
Tabelle 4-3: Für die Nutzenbewertung herangezogene patientenrelevante Endpunkte (Hauptanalyse) .....	68
Tabelle 4-4: Übersicht der Dimensionen und Summenskalen des QoL-E Fragebogens .....	82
Tabelle 4-5: Übersicht der über den EORTC QLQ-C30-Fragebogen erhobenen Symptome der Anämie auf Ebene der UE.....	86
Tabelle 4-6: Analysepopulationen auf Studienebene aus MEDALIST .....	89
Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel .....	103
Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	104
Tabelle 4-9: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel .....	107
Tabelle 4-10: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel.....	108
Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	109
Tabelle 4-12: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel .....	111
Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	113
Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	113
Tabelle 4-15: Behandlungsdauer der Studienpopulation in MEDALIST .....	125
Tabelle 4-16: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	128
Tabelle 4-17: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel .....	129
Tabelle 4-18: Operationalisierung des Endpunkts Mortalität – Gesamtüberleben .....	133
Tabelle 4-19: Bewertung des Verzerrungspotenzials für den Endpunkt Mortalität – Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel .....	133
Tabelle 4-20: Beobachtungszeiten für den Endpunkt Mortalität – Gesamtüberleben aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	134
Tabelle 4-21: Ergebnisse für den Endpunkt Mortalität – Gesamtüberleben aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	135
Tabelle 4-22: Operationalisierung des Endpunkts Morbidität – Transfusionsfreiheit.....	138

Tabelle 4-23: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Transfusionsfreiheit in MEDALIST .....	139
Tabelle 4-24: Ergebnisse für den Endpunkt Morbidität – Transfusionsfreiheit aus MEDALIST – Anteil der Patienten mit einer Transfusionsfreiheit; Woche 1 – 24.....	140
Tabelle 4-25: Zusatzanalyse 1: Ergebnisse für den Endpunkt Morbidität – Transfusionsfreiheit aus MEDALIST – Zeit bis zum Erreichen einer Transfusionsfreiheit; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	142
Tabelle 4-26: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Transfusionsfreiheit aus MEDALIST – Längste Dauer eines transfusionsfreien Zeitraums; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	145
Tabelle 4-27: Zusatzanalyse 4: Ergebnisse für den Endpunkt Morbidität – Transfusionsfreiheit aus MEDALIST – Mittlere Änderung des Hb-Werts .....	149
Tabelle 4-28: Operationalisierung des Endpunkts Morbidität – Reduktion der Transfusionslast.....	153
Tabelle 4-29: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Reduktion der Transfusionslast in MEDALIST.....	154
Tabelle 4-30: Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST – Anteil der Patienten mit einem HI-E; Woche 1 – 24.....	155
Tabelle 4-31: Zusatzanalyse 1: Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST – Zeit bis zum Erreichen eines HI-E; Woche eins bis zum Datenschnitt vom 01.07.2019.....	156
Tabelle 4-32: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST – Mittlere Änderung der Transfusionslast.....	159
Tabelle 4-33: Zusatzanalyse 3: Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST – Mittlere Änderung der Transfusionshäufigkeit .....	160
Tabelle 4-34: Operationalisierung des Endpunkts Morbidität – Reduktion der Eisenüberladung .....	163
Tabelle 4-35: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Reduktion der Eisenüberladung in MEDALIST .....	164
Tabelle 4-36: Ergebnisse zum Endpunkt Morbidität – Reduktion der Eisenüberladung aus MEDALIST – Mittlere Änderung des sFerritin-Spiegels .....	165
Tabelle 4-37: Ergebnisse zum Endpunkt Morbidität – Reduktion der Eisenüberladung aus MEDALIST – Mittlere Änderung der täglichen Dosis Eisenchelator .....	166
Tabelle 4-38: Operationalisierung des Endpunkts Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens .....	169
Tabelle 4-39: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens in MEDALIST .....	171
Tabelle 4-40: Rücklaufquoten des EORTC QLQ-C30-Fragebogens aus MEDALIST.....	172
Tabelle 4-41: Ergebnisse zum Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Anteil der Patienten mit einer Verbesserung der Symptomatik um mindestens zehn Punkte; Woche 1 – 24.....	173

Tabelle 4-42: Zusatzanalyse 1: Ergebnisse zum Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte; Woche 1 – 48 .....	174
Tabelle 4-43: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Werte der Symptomatik zu den Erhebungszeitpunkten; Woche 1 – 48.....	185
Tabelle 4-44: Zusatzanalyse 3: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik im Studienverlauf (MMRM-Analyse); Woche 1 – 24.....	188
Tabelle 4-45: Zusatzanalyse 3: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik im Studienverlauf (MMRM-Analyse); Woche 1 – 48.....	190
Tabelle 4-46: Operationalisierung des Endpunkts Morbidität – Hospitalisierungen.....	193
Tabelle 4-47: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Hospitalisierungen in MEDALIST .....	194
Tabelle 4-48: Ergebnisse für den Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Anteil der Patienten mit mindestens einer Hospitalisierung; Woche 1 – 24 .	195
Tabelle 4-49: Ergebnisse für den Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Gründe für Hospitalisierungen; Woche 1 – 24 .....	196
Tabelle 4-50: Sensitivitätsanalyse: Ergebnisse zum Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Anteil der Patienten mit mindestens einer Hospitalisierung aufgrund von UE und/oder Transfusionen; Woche 1 – 24 .....	197
Tabelle 4-51: Ergebnisse für den Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Hospitalisierungen aufgrund von UE auf Ebene der SOC und PT .....	198
Tabelle 4-52: Zusatzanalyse 1: Ergebnisse zum Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Zeit bis zur ersten Hospitalisierung; Woche eins bis zum Datenschnitt vom 01.07.2019.....	204
Tabelle 4-53: Zusatzanalyse 1: Ergebnisse zum Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Gründe für Hospitalisierungen; Woche eins bis zum Datenschnitt vom 01.07.2019.....	206
Tabelle 4-54: Operationalisierung des Endpunkts Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens.....	208
Tabelle 4-55: Bewertung des Verzerrungspotenzials für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens in MEDALIST .....	210
Tabelle 4-56: Ergebnisse zum Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Anteil der Patienten mit einer Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte; Woche 1 – 24 .....	211
Tabelle 4-57: Zusatzanalyse 1: Ergebnisse zum Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Zeit bis zur Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte; Woche 1 – 48.....	212

Tabelle 4-58: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Werte der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten; Woche 1 – 48.....	221
Tabelle 4-59: Zusatzanalyse 3: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 24.....	224
Tabelle 4-60: Zusatzanalyse 3: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 48.....	225
Tabelle 4-61: Operationalisierung des Endpunkts Lebensqualität gemessen anhand des QoL-E-Fragebogens .....	228
Tabelle 4-62: Bewertung des Verzerrungspotenzials für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens in MEDALIST .....	229
Tabelle 4-63: Rücklaufquoten des QoL-E-Fragebogens aus MEDALIST .....	230
Tabelle 4-64: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 24.....	232
Tabelle 4-65: Zusatzanalyse 1: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 48.....	234
Tabelle 4-66: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Werte der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten; Woche 1 – 48.....	241
Tabelle 4-67: Operationalisierung des Endpunkts Sicherheit .....	247
Tabelle 4-68: Bewertung des Verzerrungspotenzials für den Endpunkt Sicherheit in MEDALIST.....	249
Tabelle 4-69: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Gesamtraten UE; Woche 1 – 24.....	250
Tabelle 4-70: Ergebnisse zum Endpunkt Sicherheit aus MEDALIST – Gesamtraten UE über die Zeit bis zum Datenschnitt vom 01.07.2019.....	251
Tabelle 4-71: Sensitivitätsanalyse: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Gesamtraten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen; Woche 1 – 24 .....	252
Tabelle 4-72: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Inzidenzen von UE auf Ebene der SOC und PT (jegliche UE ab Inzidenzen von mindestens 10 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche 1 – 24.....	253
Tabelle 4-73: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Inzidenzen von SUE auf Ebene der SOC und PT (SUE ab Inzidenzen von mindestens 5 % bzw.	

mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche 1 – 24.....	257
Tabelle 4-74: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Inzidenzen von schweren UE (CTCAE $\geq$ 3) auf Ebene der SOC und PT (Schwere UE (CTCAE $\geq$ 3) ab Inzidenzen von mindestens 5 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche 1 – 24 .....	258
Tabelle 4-75: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Inzidenzen von Therapieabbrüchen aufgrund von UE auf Ebene der SOC und PT; Woche 1 – 24.....	259
Tabelle 4-76: Zusatzanalyse 1: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten UE; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	261
Tabelle 4-77: Zusatzanalyse 1: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen; Woche eins bis zum Datenschnitt vom 01.07.2019.....	264
Tabelle 4-78: Zusatzanalyse 2: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten UE auf Ebene der SOC und PT (jegliche UE ab Inzidenzen von mindestens 10 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche eins bis zum Datenschnitt vom 01.07.2019.....	267
Tabelle 4-79: Zusatzanalyse 2: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten SUE auf Ebene der SOC und PT (SUE ab Inzidenzen von mindestens 5 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche eins bis zum Datenschnitt vom 01.07.2019.....	272
Tabelle 4-80: Zusatzanalyse 2: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) auf Ebene der SOC und PT (schwere UE ab Inzidenzen von mindestens 5 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche eins bis zum Datenschnitt vom 01.07.2019.....	273
Tabelle 4-81: Zusatzanalyse 2: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum Therapieabbruch aufgrund von UE auf Ebene der SOC und PT; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	275
Tabelle 4-82: Zusatzanalyse 3: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – UE von besonderem Interesse; Woche 1 – 24.....	277
Tabelle 4-83: Zusatzanalyse 3: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten UE von besonderem Interesse; Woche eins bis zum Datenschnitt vom 01.07.2019.....	279
Tabelle 4-84: Zusatzanalyse 4: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zur Progression in eine AML .....	281
Tabelle 4-85: Zusatzanalyse 5: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Progression in ein Hochrisiko-MDS .....	283
Tabelle 4-86: Matrix der durchgeführten Subgruppenanalysen.....	287
Tabelle 4-87: Ergebnisse des Interaktionstests der Subgruppenanalysen je Endpunkt für MEDALIST.....	289
Tabelle 4-88: Ergebnisse für die Subgruppenanalysen zum Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST; Woche 1 – 24 .....	297

Tabelle 4-89: Ergebnisse für die Subgruppenanalysen zum Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Fatigue; Woche 1 – 24 .....	298
Tabelle 4-90: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – körperliches Wohlbefinden; Woche 1 – 24.....	299
Tabelle 4-91: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – funktionales Wohlbefinden; Woche 1 – 24.....	300
Tabelle 4-92: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – soziales/familiäres Wohlbefinden; Woche 1 – 24.....	301
Tabelle 4-93: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – allgemeines Wohlbefinden; Woche 1 – 24.....	302
Tabelle 4-94: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Behandlungseffekt; Woche 1 – 24.....	303
Tabelle 4-95: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Gesamtsummenscore; Woche 1 – 24.....	304
Tabelle 4-96: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE aus MEDALIST; Woche 1 – 24.....	305
Tabelle 4-97: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – SUE aus MEDALIST; Woche 1 – 24.....	307
Tabelle 4-98: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – Therapieabbrüche aufgrund von UE aus MEDALIST; Woche 1 – 24 .....	308
Tabelle 4-99: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE: SOC Gastrointestinale Erkrankungen aus MEDALIST; Woche 1 – 24.....	309
Tabelle 4-100: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE: SOC Erkrankungen des Nervensystems aus MEDALIST; Woche 1 – 24.....	310
Tabelle 4-101: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE: PT Asthenie aus MEDALIST; Woche 1 – 24 .....	312
Tabelle 4-102: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE: PT Schwindelgefühl aus MEDALIST; Woche 1 – 24 .....	313
Tabelle 4-103: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche .....	316
Tabelle 4-104: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	316
Tabelle 4-105: Operationalisierung von <Endpunkt xxx>.....	317
Tabelle 4-106: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche .....	317

Tabelle 4-107: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	318
Tabelle 4-108: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien .....	320
Tabelle 4-109: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien .....	321
Tabelle 4-110: Operationalisierung von <Endpunkt xxx>.....	321
Tabelle 4-111: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien .....	322
Tabelle 4-112: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen ..	324
Tabelle 4-113: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	324
Tabelle 4-114: Zusammenfassung der Ergebnisse zum medizinischen Zusatznutzen von Luspatercept .....	332
Tabelle 4-115: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	346
Tabelle 4-116 (Anhang): Studiendesign und -methodik für Studie MEDALIST .....	374
Tabelle 4-117 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MEDALIST .....	391

**Abbildungsverzeichnis**

	<b>Seite</b>
Abbildung 4-1: Plausibilisierung möglicher Effektmodifikationen in MEDALIST .....	99
Abbildung 4-2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel .....	105
Abbildung 4-3: Übersicht des Studiendesigns der MEDALIST-Studie [1].....	123
Abbildung 4-4: Kaplan-Meier-Kurve des Gesamtüberlebens aus MEDALIST; Analysezeitraum bis zum Datenschnitt vom 01.07.2019 .....	136
Abbildung 4-5: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur Transfusionsfreiheit über $\geq 8$ Wochen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	143
Abbildung 4-6: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur Transfusionsfreiheit über $\geq 12$ Wochen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	143
Abbildung 4-7: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur Transfusionsfreiheit über $\geq 16$ Wochen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	144
Abbildung 4-8: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur Transfusionsfreiheit über $\geq 24$ Wochen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	144
Abbildung 4-9: Zusatzanalyse 2: Kaplan-Meier-Kurve der längsten Dauer eines transfusionsfreien Zeitraums in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	146
Abbildung 4-10: Zusatzanalyse 3: Swimmer-Plot der Transfusionszeitpunkte und transfusionsfreien Zeiträume über mindestens acht Wochen aus MEDALIST; Studienzeitraum bis zum Datenschnitt vom 01.07.2019 .....	148
Abbildung 4-11: Zusatzanalyse 4: Hb-Werte im Studienverlauf in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	150
Abbildung 4-12: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum Erreichen eines HI-E (insgesamt) in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	157
Abbildung 4-13: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum Erreichen eines HI-E I für Patienten mit hoher Transfusionslast zur Baseline ( $\geq 4$ EK-Einheiten pro acht Wochen) in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	157
Abbildung 4-14: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum Erreichen eines HI-E II für Patienten mit niedriger Transfusionslast zur Baseline ( $< 4$ EK-Einheiten pro acht Wochen) in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	158
Abbildung 4-15: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Fatigue; Woche 1 – 48.....	175

Abbildung 4-16: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Übelkeit und Erbrechen; Woche 1 – 48 .....	176
Abbildung 4-17: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Schmerz; Woche 1 – 48.....	176
Abbildung 4-18: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Dyspnoe; Woche 1 – 48.....	177
Abbildung 4-19: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Schlaflosigkeit; Woche 1 – 48 .....	177
Abbildung 4-20: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Appetitminderung; Woche 1 – 48 .....	178
Abbildung 4-21: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Obstipation; Woche 1 – 48.....	178
Abbildung 4-22: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Diarrhö; Woche 1 – 48.....	179
Abbildung 4-23: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Finanzielle Schwierigkeiten; Woche 1 – 48 .....	179
Abbildung 4-24: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Fatigue; Woche 1 – 48.....	180
Abbildung 4-25: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Übelkeit und Erbrechen; Woche 1 – 48 .....	181
Abbildung 4-26: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Schmerz; Woche 1 – 48.....	181
Abbildung 4-27: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Dyspnoe; Woche 1 – 48.....	182
Abbildung 4-28: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Schlaflosigkeit; Woche 1 – 48.....	182

Abbildung 4-29: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Appetitminderung; Woche 1 – 48.....	183
Abbildung 4-30: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Obstipation; Woche 1 – 48.....	183
Abbildung 4-31: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Durchfall; Woche 1 – 48.....	184
Abbildung 4-32: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Finanzielle Schwierigkeiten; Woche 1 – 48 .....	184
Abbildung 4-33: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur ersten Hospitalisierung in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	205
Abbildung 4-34: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur ersten Hospitalisierung aufgrund von UE und/oder Transfusionen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	205
Abbildung 4-35: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Körperliche Funktion; Woche 1 – 48.....	213
Abbildung 4-36: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Rollenfunktion; Woche 1 – 48 .....	213
Abbildung 4-37: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Emotionale Funktion; Woche 1 – 48 .....	214
Abbildung 4-38: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Kognitive Funktion; Woche 1 – 48.....	214
Abbildung 4-39: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Soziale Funktion; Woche 1 – 48 .....	215
Abbildung 4-40: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Gesundheitlicher Status insgesamt; Woche 1 – 48 .....	215
Abbildung 4-41: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – körperliche Funktion; Woche 1 – 48.....	216

Abbildung 4-42: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – Rollenfunktion; Woche 1 – 48.....	217
Abbildung 4-43: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – emotionale Funktion; Woche 1 – 48 .....	218
Abbildung 4-44: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – kognitive Funktion; Woche 1 – 48.....	219
Abbildung 4-45: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – soziale Funktion; Woche 1 – 48.....	220
Abbildung 4-46: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – gesundheitlicher Status insgesamt; Woche 1 – 48.....	221
Abbildung 4-47: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – körperliches Wohlbefinden; Woche 1 – 48.....	236
Abbildung 4-48: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – funktionales Wohlbefinden; Woche 1 – 48.....	237
Abbildung 4-49: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – soziales/familiäres Wohlbefinden; Woche 1 – 48.....	237
Abbildung 4-50: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – Fatigue; Woche 1 – 48.....	238
Abbildung 4-51: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – MDS-bezogene Störungen; Woche 1 – 48 .....	239
Abbildung 4-52: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – sexuelles Wohlbefinden; Woche 1 – 48.....	239

Abbildung 4-53: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – allgemeines Wohlbefinden; Woche 1 – 48.....	240
Abbildung 4-54: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – Gesamtsummenskala; Woche 1 – 48 .....	240
Abbildung 4-55: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – Behandlungseffekt; Woche 1 – 48 .....	241
Abbildung 4-56: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten UE in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	262
Abbildung 4-57: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten SUE in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	262
Abbildung 4-58: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq 3$ ) in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	263
Abbildung 4-59: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	263
Abbildung 4-60: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	265
Abbildung 4-61: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten SUE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	265
Abbildung 4-62: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq 3$ ) ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	266
Abbildung 4-63: Zusatzanalyse 4: Kaplan-Meier-Kurve für die Zeit bis zur Progression in eine AML aus MEDALIST; Analysezeitraum bis zum Datenschnitt vom 01.07.2019.....	282
Abbildung 4-64: Patientenfluss in MEDALIST zum Datenschnitt vom 08.05.2018 .....	389
Abbildung 4-65 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	417
Abbildung 4-66 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Asthenie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	418
Abbildung 4-67 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Fatigue aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	418
Abbildung 4-68: (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Ödem peripher aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	419

Abbildung 4-69 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Fieber aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	419
Abbildung 4-70 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Gastrointestinale Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	420
Abbildung 4-71 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Bauchschmerz aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	420
Abbildung 4-72 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Schmerzen im Oberbauch aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	421
Abbildung 4-73 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Obstipation aus MEDALIST; Woche Woche eins bis zum Datenschnitt vom 01.07.2019...	421
Abbildung 4-74 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Diarrhö aus MEDALIST; Woche Woche eins bis zum Datenschnitt vom 01.07.2019 .....	422
Abbildung 4-75 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Übelkeit aus MEDALIST; Woche Woche eins bis zum Datenschnitt vom 01.07.2019 .....	422
Abbildung 4-76 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Erbrechen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	423
Abbildung 4-77 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Hepatobiliäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	423
Abbildung 4-78 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Infektionen und parasitäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	424
Abbildung 4-79 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Bronchitis aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	424
Abbildung 4-80 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Grippe aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	425
Abbildung 4-81 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Infektion der oberen Atemwege aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	425
Abbildung 4-82 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Harnwegsinfektion aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019...	426
Abbildung 4-83 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT virale Infektion der oberen Atemwege aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	426
Abbildung 4-84 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen des Nervensystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	427
Abbildung 4-85 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Schwindelgefühl aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	427

Abbildung 4-86 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Kopfschmerzen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	428
Abbildung 4-87 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Synkope aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	428
Abbildung 4-88 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	429
Abbildung 4-89 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Rückenschmerzen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 ....	429
Abbildung 4-90 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Myalgie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	430
Abbildung 4-91 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Arthralgie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	430
Abbildung 4-92 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	431
Abbildung 4-93 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Husten aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	431
Abbildung 4-94 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Dyspnoe aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	432
Abbildung 4-95 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Epistaxis aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	432
Abbildung 4-96 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Stoffwechsel- und Ernährungsstörungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	433
Abbildung 4-97 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT verminderter Appetit aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019	433
Abbildung 4-98 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Hyperglykämie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	434
Abbildung 4-99 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Hyperurikämie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	434
Abbildung 4-100 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Herzerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019....	435
Abbildung 4-101 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Palpitationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	435
Abbildung 4-102 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Untersuchungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	436
Abbildung 4-103 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Aspartataminotransferase erhöht aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	436
Abbildung 4-104 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Gefäßerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019..	437

Abbildung 4-105 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Hypertonie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	437
Abbildung 4-106 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen der Haut und des Unterhautgewebes aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	438
Abbildung 4-107 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen der Niere und Harnwege aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	438
Abbildung 4-108 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Psychiatrische Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	439
Abbildung 4-109 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Verwirrheitszustand aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019	439
Abbildung 4-110 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Schlaflosigkeit aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	440
Abbildung 4-111 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	440
Abbildung 4-112 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Sturz aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	441
Abbildung 4-113 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen des Blutes und des Lymphsystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	441
Abbildung 4-114 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Anämie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	442
Abbildung 4-115 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Augenerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.	442
Abbildung 4-116 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen des Ohrs und des Labyrinths aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	443
Abbildung 4-117 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Vertigo aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	443
Abbildung 4-118 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Gutartige, bösartige und nicht spezifizierte Neubildungen (inkl. Zysten und Polypen) aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	444
Abbildung 4-119 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen der Geschlechtsorgane und der Brustdrüse aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	444
Abbildung 4-120 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Infektionen und parasitäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	445

Abbildung 4-121 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	445
Abbildung 4-122 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	446
Abbildung 4-123 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Gastrointestinale Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	446
Abbildung 4-124 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Herzerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019....	447
Abbildung 4-125 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	447
Abbildung 4-126 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen) aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	448
Abbildung 4-127 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Erkrankungen des Blutes und des Lymphsystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	448
Abbildung 4-128 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) des PT Anämie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	449
Abbildung 4-129 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) des PT Neutropenie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	450
Abbildung 4-130 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Gastrointestinale Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	450
Abbildung 4-131 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Infektionen und parasitäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	451
Abbildung 4-132 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Stoffwechsel- und Ernährungsstörungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	451
Abbildung 4-133 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019....	452
Abbildung 4-134 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Erkrankungen des Nervensystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	453

Abbildung 4-135 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	453
Abbildung 4-136 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) des PT Sturz aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	454
Abbildung 4-137 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Herzerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	455
Abbildung 4-138 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE $\geq$ 3) der SOC Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	455
Abbildung 4-139 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	456
Abbildung 4-140 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Erkrankungen des Nervensystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	456
Abbildung 4-141 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Hirnblutung aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	457
Abbildung 4-142 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Kopfschmerzen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	457
Abbildung 4-143 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Erinnerungsvermögen eingeschränkt aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	458
Abbildung 4-144 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Morbus Parkinson aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	458
Abbildung 4-145 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen) aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	459
Abbildung 4-146 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Kolonadenom aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	459
Abbildung 4-147 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Myelodysplastische Syndrome aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	460
Abbildung 4-148 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Refraktäre Anämie mit Blastenüberschuss aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	460

Abbildung 4-149 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Transformation zu einer AML aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	461
Abbildung 4-150 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Erkrankungen der Nieren und Harnwege aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	461
Abbildung 4-151 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Nierenversagen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	462
Abbildung 4-152 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT chronische Nierenerkrankung aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	462
Abbildung 4-153 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Infektionen und parasitäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	463
Abbildung 4-154 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Sepsis aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	463
Abbildung 4-155 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019....	464
Abbildung 4-156 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Asthenie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	464
Abbildung 4-157 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Fatigue aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019 .....	465
Abbildung 4-158 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Generelle Verschlechterung des physischen Gesundheitsstatus aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	465
Abbildung 4-159 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Gefäßerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	466
Abbildung 4-160 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Hämorrhagischer Schock aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	466
Abbildung 4-161 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	467
Abbildung 4-162 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Lungenfibrose aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	467

Abbildung 4-163 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Respiratorische Insuffizienz aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	468
Abbildung 4-164 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	468
Abbildung 4-165 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Arthralgie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	469
Abbildung 4-166 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Myalgie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	469
Abbildung 4-167 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Untersuchungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	470
Abbildung 4-168 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Alaninaminotransferase erhöht aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	470
Abbildung 4-169 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Aspartataminotransferase erhöht aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	471
Abbildung 4-170 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Erkrankungen des Blutes und des Lymphsystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	471
Abbildung 4-171 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Thrombozytose aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	472
Abbildung 4-172 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	472
Abbildung 4-173 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Verkehrsunfall aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019.....	473

**Abkürzungsverzeichnis**

<b>Abkürzung</b>	<b>Bedeutung</b>
AMIS	Arzneimittel-Informationssystem
AML	Akute myeloische Leukämie
ANC	absolute Neutrophilenzahl ( <i>absolute neutrophil count</i> )
ANCOVA	<i>Analysis of covariance</i>
ANOVA	<i>Analysis of variance</i>
ARR	absolute Risikodifferenz/absolute Risikoreduktion
ATC	<i>Anatomical Therapeutic Chemical;</i>
BÄK	Bundesärztekammer
BMI	Body-Mass-Index
BSC	Best-Supportive-Care
CMH	Cochran-Mantel-Haenszel
CONSORT	Consolidated Standards of Reporting Trials
CTCAE	<i>Common Terminology Criteria for Adverse Events</i>
DIMDI	Deutsches Institut für Medizinische Dokumentation
ECOG	<i>Eastern Cooperative Oncology Group</i>
EG	Europäische Gemeinschaft
EK	Erythrozytenkonzentrat
EMA	European Medicines Agency
EORTC QLQ-C30	<i>European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module</i>
EOT	<i>End of treatment</i>
EPO	Erythropoetin
ESA	Erythropoese-stimulierende Agenzien ( <i>erythropoiesis-stimulating agents</i> )
EU-CTR	<i>EU Clinical Trials Register</i>
EUMDS	<i>European LeukaemiaNet MDS</i>
FACT(-G)	<i>Functional Assessment of Cancer Therapy (-General)</i>
G-BA	Gemeinsamer Bundesausschuss
G-CSF	Granulozyten-Kolonie-stimulierender Faktor ( <i>granulocyte-colony-stimulating factor</i> )
GM-CSF	Granulozyten-Makrophagen-Kolonie-stimulierender Faktor ( <i>granulocyte-macrophage-colony-stimulating factor</i> )

<b>Abkürzung</b>	<b>Bedeutung</b>
Hb	Hämoglobin
HI-E	Erythroides Ansprechen ( <i>hematologic improvement-erythroid</i> )
HMA	hypomethylierende Agenzien
HR	Hazard Ratio
HRQoL	gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> )
HTB	<i>High Transfusion Burden</i>
ICTRP	<i>International Clinical Trials Registry Platform</i>
IMiD	Immunmodulierende Arzneimittel
IPSS-R	<i>International Prognostic Scoring System-revised</i>
IQWiG	Institut für Wirtschaftlichkeit im Gesundheitswesen
IRT	<i>Interactive Response Technology</i>
ITT	Intention-to-Treat
IWG	<i>International Working Group</i>
KI	Konfidenzintervall
KM	Knochenmark
LLT	Begriff niedrigster Ebene ( <i>lowest level term</i> )
LS	Methode der kleinsten Quadrate ( <i>least square</i> )
LTB	<i>Low Transfusion Burden</i>
MACE	<i>Major Adverse Cardiac Events</i>
Max	Maximum
MCID	minimale klinische Relevanzschwelle ( <i>minimally clinically important difference</i> )
MDS	Myelodysplastische Syndrome
MedDRA	<i>Medical Dictionary for Regulatory Activities</i>
Min	Minimum
MTC	Mixed Treatment Comparison
MW	Mittelwert
MWD	Mittelwertdifferenz
NCI	<i>National Cancer Institute</i>
N.E.	nicht erreicht oder nicht berechnet
OR	Odds Ratio
OS	Gesamtüberleben ( <i>overall survival</i> )

<b>Abkürzung</b>	<b>Bedeutung</b>
PT	Bevorzugter Begriff ( <i>preferred term</i> )
RARS	Refraktäre Anämie mit Ringsideroblasten ( <i>refractory anemia with ring sideroblasts</i> )
RCMD	Refraktäre Zytopenie mit multilineären Dysplasien ( <i>refractory cytopenia with multi-lineage dysplasia</i> )
RCT	Randomisiert kontrollierte Studie ( <i>randomized controlled trial</i> )
RR	Relatives Risiko
s. c.	subkutan
SD	Standardabweichung ( <i>standard deviation</i> )
SF3B1	<i>Splicing factor 3B subunit 1</i>
SGB	Sozialgesetzbuch
SMQ	<i>Standardised MedDRA Query</i>
SOC	Systemorganklasse ( <i>system organ class</i> )
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes UE
TOI	Behandlungseffekt ( <i>Treatment Outcome Index</i> )
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des Gemeinsamen Bundesausschusses
WHO	World Health Organization
zVT	zweckmäßige Vergleichstherapie

#### 4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

#### 4.1 Zusammenfassung der Inhalte von Modul 4

*Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.*

##### **Fragestellung**

In der vorliegenden Nutzenbewertung soll das Ausmaß des Zusatznutzens von Luspatercept (Reblozyl®) anhand von patientenrelevanten Endpunkten untersucht werden. Zielpopulation sind erwachsene Patienten mit transfusionsabhängiger Anämie aufgrund von myelodysplastischen Syndromen (MDS) mit Ringsideroblasten, mit sehr niedrigem, niedrigem oder intermediärem Risiko, die auf eine Erythropoetin (EPO)-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind.

Da es sich bei Luspatercept um ein Arzneimittel zur Behandlung eines seltenen Leidens gemäß § 35a Absatz 1 Satz 11 SGB V handelt, gilt der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Das Ausmaß des Zusatznutzens ist nachzuweisen.

##### **Datenquellen**

Zur Beantwortung der Fragestellung wurde die doppelblinde, randomisiert kontrollierte Phase III-Studie MEDALIST herangezogen.

In MEDALIST wurden erwachsene Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie, d. h. eine Therapie mit Erythropoese-stimulierenden Agenzien (*erythropoiesis-stimulating agents*, ESA), nicht zufriedenstellend angesprochen haben oder für eine solche Therapie nicht geeignet sind, hinsichtlich patientenrelevanter Endpunkte der Kategorien Mortalität, Morbidität, Lebensqualität und Sicherheit untersucht. Die Diagnose und Klassifikation der Erkrankung erfolgte anhand der etablierten und anerkannten Kriterien der WHO (*World Health Organization*) 2016 und dem *International Prognostic Scoring System-revised* (IPSS-R).

Durch die Wahl der Einschlusskriterien von MEDALIST wurde sichergestellt, dass die im Anwendungsgebiet von Luspatercept spezifizierten Patientencharakteristika erfüllt wurden.

Es wurden 229 Patienten in die Studie eingeschlossen und im Verhältnis 2:1 auf die beiden Behandlungsarme Luspatercept + Best-Supportive-Care (BSC) (n = 153) – im Folgenden als „Luspatercept“ bezeichnet – und Placebo + BSC (n = 76) – im Folgenden als „Placebo“ bezeichnet – randomisiert. Die Placebo-Gabe erfolgte analog zur Luspatercept-Gabe, um die Verblindung aufrecht zu erhalten. Unter BSC wurde insbesondere die bedarfsgerechte Behandlung mit Transfusionen mit Erythrozytenkonzentraten (EK), ggf. in Kombination mit Eisenchelatoren, verstanden. Dies entspricht dem derzeitigen Behandlungsstandard bei Niedrigrisiko-MDS-Patienten mit transfusionsabhängiger Anämie gemäß Anwendungsgebiet von Luspatercept.

In MEDALIST erfolgte die Auswertung der patientenrelevanten Endpunkte zu Woche 24 und zu Woche 48. Patienten wurden in der primären Behandlungsphase mit Luspatercept oder Placebo bis Woche 24 behandelt. Ab Woche 25 wurden nur solche Patienten entsprechend ihrer ursprünglichen Zuordnung weiter behandelt, die nach Ermessen des Prüfarztes in den ersten 24 Wochen einen klinischen Nutzen der Behandlung und keine Krankheitsprogression aufwiesen. Aufgrund der Imbalance in den Anteilen der Patienten, die von Woche 25 bis Woche 48 behandelt wurden (Luspatercept vs. Placebo: 51 % vs. 15,8 %) und der daraus resultierenden Unsicherheit hinsichtlich der Interpretierbarkeit, werden die Aussagen zum medizinischen Zusatznutzen auf den Ergebnissen der ersten 24 Wochen abgeleitet. Davon abweichend wurde für das Gesamtüberleben der gesamte Studienzeitraum herangezogen. Die Ergebnisse bis Woche 48 werden der Vollständigkeit halber als Zusatzanalysen ergänzend berichtet.

Für die herangezogene Phase-III-Studie MEDALIST lagen zum Zeitpunkt der Dossiereinreichung der präspezifizierte Datenschnitt vom 08.05.2018 und die im Rahmen der Zulassung des zu bewertenden Arzneimittels erfolgten zusätzlichen Datenschnitte vom 07.01.2019 und 01.07.2019 vor. Die zusätzlich durchgeführten Datenschnitte im Jahr 2019 erfolgten zur Unterstützung der Beantwortung einer Rückfrage der *European Medicines Agency* (EMA). Zum Zeitpunkt des präspezifizierten Datenschnitts (08.05.2018) hatten alle Patienten ihre Behandlung bis Woche 48 abgeschlossen oder die Behandlung frühzeitig abgebrochen. Demnach hatten alle Patienten die primäre Behandlungsphase (Woche 1 – 24) abgeschlossen. Gemäß Studienprotokoll durften Patienten entsprechend ihrer Zuordnung weiter behandelt werden, sofern sie weiterhin einen klinischen Nutzen der Behandlung und keine Krankheitsprogression aufwiesen (Überprüfung alle 24 Wochen). Zum Zeitpunkt des Datenschnitts vom 07.01.2019 hatten alle Patienten des Placebo-Arms die Behandlung abgebrochen, im Luspatercept-Arm waren es 68,6 % der Patienten (105/153). Die Bewertung des Zusatznutzens von Luspatercept beruht daher im vorliegenden Anwendungsgebiet primär auf den Ergebnissen des präspezifizierten Datenschnitts vom 08.05.2018 (Ergebnisse der vergleichenden, doppelblinden, primären Behandlungsphase bis Woche 24). Die Ergebnisse der zusätzlichen, von der EMA geforderten Datenschnitte finden lediglich in den ergänzend berichteten Ereigniszeitanalysen, sowie in den Auswertungen zum Gesamtüberleben Berücksichtigung, um ein vollumfängliches Bild der Wirksamkeit und Sicherheit von Luspatercept darzustellen.

### **Ein-/Ausschlusskriterien für Studien**

Zur Identifizierung relevanter Studien wurden folgenden Selektionskriterien geprüft (siehe auch Tabelle 4-2):

1. **Patientenpopulation:** Erwachsene Patienten mit transfusionsabhängiger Anämie aufgrund von MDS mit Ringsideroblasten, mit sehr niedrigem, niedrigem oder intermediärem Risiko gemäß IPSS-R oder vergleichbarer Risikoklassifizierung, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind

2. **Intervention:** Luspatercept (1,0 mg/kg bis maximal 1,75 mg/kg Körpergewicht subkutan (s. c.) alle drei Wochen)
3. **Endpunkte:** Mindestens ein patientenrelevanter Endpunkt (Mortalität, Morbidität, Lebensqualität, Sicherheit) wurde berichtet
4. **Studiendesign:** randomisiert kontrollierte Studie (*randomized controlled trial*, RCT) oder Meta-Analyse aus RCT
5. **Studiendauer:** mindestens 24 Wochen
6. **Ergebnisdarstellung:** Berichterstattungen mit ausreichenden Informationen, um Methodik und Ergebnisse der Studien adäquat bewerten und einordnen zu können

Studien, die die genannten Kriterien nicht erfüllen, wurden ausgeschlossen. Studien, die die Kriterien nur teilweise erfüllten, wurden dann eingeschlossen, wenn mindestens 80 % der Studienpopulation die Kriterien erfüllte oder eine den Kriterien entsprechende Auswertung von Teilpopulation(en) möglich war.

## **Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen**

### ***Methoden zur Bewertung der Aussagekraft der Nachweise***

Die Aussagekraft der Nachweise wurde unter Berücksichtigung des Verzerrungspotentials auf Studien- und Endpunktebene bewertet und in den Bewertungsbögen der Dossievorlage dokumentiert (siehe auch Abschnitt 4.2.4).

In einem ersten Schritt wurde auf Studienebene das Verzerrungspotential anhand endpunktübergreifender Aspekte bewertet. Basierend auf den Einzelaspekten wurde das Verzerrungspotential der Studie als „niedrig“ oder „hoch“ eingestuft.

Bei einem hohen Verzerrungspotential auf Studienebene ist davon auszugehen, dass auch die Endpunkte hoch verzerrt sind, sodass auch hier, begründete Fälle ausgenommen, eine Einstufung des Verzerrungspotentials als „hoch“ vorgenommen wurde.

Eine Einteilung als niedrig wurde vorgenommen, wenn eine Verzerrung der Ergebnisse unwahrscheinlich und keine Anhaltspunkte für verzerrende Aspekte vorhanden waren, die die Beurteilung in einer qualitativen Weise verändern würden. Bei einem niedrigen Verzerrungspotential auf Studienebene wurde das Verzerrungspotential für jeden Endpunkt einzeln anhand endpunktspezifischer Aspekte bewertet.

### ***Methoden zur Synthese von Ergebnissen***

Ergebnisse zur Ableitung des medizinischen Zusatznutzens basieren auf den patientenrelevanten Endpunkten der eingeschlossenen Studie MEDALIST und stellen den Zusatznutzen von Luspatercept + BSC gegenüber Placebo + BSC dar. Im Rahmen von BSC war eine regelmäßige, bedarfsgerechte Transfusionstherapie mit EK in Kombination mit Eisenchelatoren erlaubt, die den derzeitigen Behandlungsstandard im vorliegenden Anwendungsgebiet abbildet. Die vorliegenden Ergebnisse zum medizinischen Zusatznutzen beziehen sich somit auf den Vergleich von Luspatercept gegenüber einer adäquaten, supportiven Transfusionstherapie.

Eine detaillierte Beschreibung der Studie MEDALIST anhand der Items 2b bis 14 des *Consolidated Standards of Reporting Trial (CONSORT)*-Statements unter Berücksichtigung des Studienziels, der Methodik der Studie und der Ergebnisdarstellung findet sich in Anhang 4-E.

Die folgenden Endpunkte wurden als patientenrelevant eingestuft und zur Ableitung des Zusatznutzens herangezogen:

- **Mortalität**
  - Gesamtüberleben (*overall survival*, OS)
- **Morbidität**
  - Transfusionsfreiheit
  - Reduktion der Transfusionslast, gemessen als erythroides Ansprechen (HI-E, *hematologic improvement-erythroid*) gemäß den Kriterien der *International Working Group (IWG)* aus dem Jahr 2006
  - Symptomatik gemessen mit den Symptomskalen des *European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module (EORTC QLQ-C30)*: Fatigue, Übelkeit und Erbrechen, Schmerz, Dyspnoe, Schlaflosigkeit, Appetitminderung, Obstipation, Diarrhoe, finanzielle Schwierigkeiten
  - Hospitalisierungen
- **Lebensqualität**
  - gesundheitsbezogene Lebensqualität gemessen mit den Funktionsskalen des EORTC QLQ-C30: körperliche Funktion, Rollenfunktion, emotionale Funktion, kognitive Funktion, soziale Funktion, gesundheitlicher Status insgesamt
  - krankheitsspezifische Lebensqualität gemessen mit den Domänen des *Quality of Life-E (QoL-E)*: allgemeiner Gesundheitsstatus, körperliches Wohlbefinden, funktionales Wohlbefinden, soziales/familiäres Wohlbefinden, sexuelles Wohlbefinden, Fatigue, MDS-bezogene Störungen

- **Sicherheit**

- jegliche unerwünschte Ereignisse (UE) (jegliche UE umfassen auch Laborparameter, die nicht per se patientenrelevant sind, und werden daher in der Ableitung des Zusatznutzens nicht berücksichtigt), schwerwiegende UE (SUE), schwere UE (definiert als UE mit Schweregrad  $\geq 3$  gemäß den *Common Terminology Criteria for Adverse Events (CTCAE)* des *National Cancer Institute (NCI)*), Therapieabbrüche aufgrund von UE

#### *Zusatzanalysen*

Zur besseren Interpretierbarkeit der Ergebnisse wurden Zusatzanalysen, insbesondere auch unter Berücksichtigung der Ereignisse bis Woche 48 bzw. dem letzten verfügbaren Datenschnitt vom 01.07.2019, sowie Auswertungen zu weiteren klinisch relevanten Parametern (z. B. mittlere Änderung des Hämoglobin (Hb)-Werts im Vergleich zur Baseline) berichtet.

#### *Sensitivitätsanalysen*

Die Robustheit der Ergebnisse wurde, sofern geeignet, mittels Sensitivitätsanalysen überprüft.

Hospitalisierungen aufgrund von UE und/oder Transfusionen wurden als Sensitivitätsanalysen der Gesamthospitalisierungen, d. h. Hospitalisierungen jeglicher Ursache, dargestellt.

Für die Gesamtraten an jeglichen UE, SUE und schweren UE (definiert als UE mit CTCAE-Schweregrad  $\geq 3$ ) wurden zusätzlich Auswertungen als Sensitivitätsanalysen dargestellt, bei denen UE unberücksichtigt blieben, die bereits über die Symptomskalen des EORTC QLQ-C30-Fragebogens als Symptome der Anämie erhoben wurden, sowie UE, die eine stationäre Aufnahme bedingt haben und bereits über den Endpunkt Hospitalisierungen erhoben wurden.

#### *Subgruppenanalysen*

Zusätzlich wurden Subgruppenanalysen für alle patientenrelevanten und für die Ableitung des Zusatznutzens herangezogenen Endpunkte durchgeführt. Die Auswahl der Subgruppenmerkmale erfolgte unter Berücksichtigung relevanter Patientencharakteristika und Behandlungskriterien sowie den im Studienprotokoll a priori festgelegten Analysen [1].

**Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen**

Im vorliegenden Anwendungsgebiet der Niedrigrisiko-MDS tritt bei ca. 70 – 80 % der Patienten als Erstmanifestation eine Anämie infolge eines Erythrozyten-Reifungs-Defekts auf, welche auch als Kernsymptom angesehen werden kann [2-4]. Bei den meisten Patienten führt die Anämie zu einer erheblichen Krankheitslast, Verringerung der Leistungsfähigkeit und Einschränkungen in der Lebensqualität. Im Gegensatz zu fortgeschrittenen Stadien (Hochrisiko-MDS), bei denen die Verlängerung des Gesamtüberlebens und Verzögerung der Progression in eine AML im Vordergrund steht, stehen bei Patienten mit Niedrigrisiko-MDS, die vergleichsweise lange leben, andere Therapieziele im Vordergrund. Das primäre Therapieziel der Niedrigrisiko-MDS ist die Behandlung der Zytopenie(n), insbesondere der Anämie, um die damit assoziierte Morbidität und Symptomlast zu minimieren und die Leistungsfähigkeit, Lebensqualität und Autonomie der Patienten zu erhalten oder zu verbessern [2, 5, 6]. Für den größten Anteil an Patienten mit Niedrigrisiko-MDS stehen zu diesem Zweck weder potentiell kurative noch zielgerichtete Therapieoptionen zur Verfügung. Für diese Patienten kommt lediglich eine supportive Therapie in Frage.

***Die Transfusionstherapie mit EK, ggf. in Kombination mit Eisenchelatoren, stellt die derzeit einzige verfügbare Therapieoption im vorliegenden Anwendungsgebiet dar***

Zentraler Bestandteil der Therapie einer Anämie – für Patienten im Anwendungsgebiet von Luspatercept sogar die einzige zugelassene Behandlungsoption – ist die supportive Behandlung mit EK-Transfusionen, welche in Abhängigkeit des klinischen Zustands (dynamisch bestimmt durch Symptome, Alter, Erkrankungsstadium, individuelle Patientenbiologie, Komorbiditäten) des Patienten verabreicht werden [2].

Durch die Verabreichung von fremden, funktionstüchtigen Erythrozyten kann der Mangel an körpereigenen, reifen Erythrozyten temporär kompensiert, der Hb-Wert transient erhöht und die Symptomatik der Anämie vorübergehend gemildert werden. Die Transfusionstherapie mit EK vermag jedoch nicht eine anhaltende, stabile Korrektur des Hb-Spiegels und somit andauernde Vermeidung der Anämie-Symptomatik zu erreichen. Darüber hinaus birgt sie das Risiko teils schwerwiegender Folgekomplikationen, wie die Eisenüberladung, geht mit erheblichen psychosozialen und zeitlichen Belastungen für den Patienten einher und beeinflusst die Überlebensprognose negativ [7, 8]. Die Transfusionstherapie mit EK stellt daher nur eine unzureichende therapeutische Maßnahme dar.

Das Erreichen der Therapieziele im vorliegenden Anwendungsgebiet, der Verringerung der Krankheitslast und der Erhalt oder die Verbesserung der Lebensqualität, ist eng mit einer Reduktion oder vollständigen Vermeidung von EK-Transfusionen verknüpft [2].

***Transfusionsbedingte Belastungen aus klinischer Sicht***

EK-Transfusionen sind mit akut und langfristig auftretenden Risiken und unvermeidlichen Komplikationen verbunden, wie der transfusionsbedingten, sekundären Häm siderose. Mit zunehmender Dauer der Transfusionstherapie ergeben sich Komplikationen durch die Zufuhr des in den EK enthaltenen Eisens, da der menschliche Körper die durch die Transfusionen aufgenommenen Mengen an Eisen nicht in ausreichendem Maße ausscheiden kann. Die zunehmende Eisenüberladung des Organismus führt schließlich zur toxischen Ablagerung von überschüssigem Eisen in verschiedenen Organen, insbesondere im Herz, in der Leber und in endokrinen Organen, und den damit verbundenen, teils schwerwiegenden Komplikationen [7, 9]. Aufgrund der regelmäßigen Exposition, dem Vorliegen von Komorbiditäten, des fortgeschrittenen Alters, einer oftmals eingeschränkten Immunfunktion und einer vorliegenden Neutropenie als weitere Manifestation der MDS ist das Risiko für weitere Transfusionsnebenwirkungen und -komplikationen (bspw. transfusionsassoziierte Lungenschäden und Volumenüberladung, Herzinsuffizienz, Infektionen und allergische Reaktionen) bei MDS-Patienten besonders hoch [8].

***Transfusionsbedingte Belastungen aus Sicht der Patienten***

Neben den klinischen Risiken und Folgekomplikationen einer Transfusionstherapie können regelmäßige Transfusionen sich auf verschiedenen Ebenen auch negativ auf das Wohlbefinden der Patienten auswirken. Durch die Verabreichung von fremden funktionstüchtigen Erythrozyten kann der Mangel an körpereigenen reifen Erythrozyten temporär kompensiert, der Hb-Wert transient erhöht und die Symptomatik der Anämie vorübergehend gemildert werden. Die Transfusionstherapie mit EK vermag jedoch nicht eine anhaltende, stabile Korrektur des Hb-Spiegels und somit andauernde Vermeidung der Anämie-Symptomatik zu erreichen. Nach Verabreichung der EK-Transfusion fühlen sich die Patienten aufgrund hoher Hb-Werte initial zwar gut, jedoch fallen die Hb-Werte schnell wieder ab, sodass Anämie-Symptome wieder auftreten können. Das regelmäßige Absinken des Hb-Werts ist mit dem Risiko von Anämie-Symptomen (z. B. Fatigue) und Anämie-bedingten Schäden (bspw. Verschlechterung von Komorbiditäten und des Gesundheitszustands) verbunden [5, 10, 11]. Diese Schwankungen des Hb-Werts sind für den Patienten in entsprechenden Energieniveaus spürbar, welche sich negativ auf den Allgemeinzustand und die Lebensqualität auswirken [5, 12]. MDS-Patienten beschreiben das Auftreten der Belastungsintoleranzen als ein „Leben mit einem 50 Pfund-Rucksack auf den Schultern“ oder „einer sehr schweren Decke“, die jedes Aufstehen erschwert [13-15].

Des Weiteren ist die Verabreichung von EK-Transfusionen häufig mit diversen psychosozialen Problemen und einem nicht unerheblichen Zeitaufwand verbunden, der die Teilnahme der Patienten an sozialen und beruflichen Aktivitäten stark einschränken kann [16]. Um sowohl die häufigen Arzttermine wahrzunehmen als auch an anderen Aktivitäten teilzunehmen sind transfusionsabhängige MDS-Patienten häufig auf die Hilfe von Familienmitgliedern oder anderen Personen angewiesen [8, 12, 15]. Somit besteht nicht nur ein enormer zeitlicher Aufwand, sondern auch eine Einschränkung der Autonomie [11, 13].

Unter Berücksichtigung dieser Limitationen der Transfusionstherapie ist es nicht überraschend, dass laut einer Befragung von MDS-Patienten in den USA, Großbritannien und Deutschland ein Zustand ohne Transfusionen oder zumindest mit weniger Transfusionen einer Transfusionsabhängigkeit deutlich bevorzugt wird [17].

***Im vorliegenden Anwendungsgebiet besteht ein großer therapeutischer Bedarf, der nun durch Luspatercept gedeckt werden kann***

Unter Berücksichtigung der mangelnden Therapiealternativen und bedeutsamen Limitationen und Risiken einer Transfusionstherapie besteht für die Patienten im vorliegenden Anwendungsgebiet ein großer, ungedeckter therapeutischer Bedarf nach einem Arzneimittel, das das Erreichen einer Transfusionsfreiheit ermöglicht und zu einer besseren Kontrolle der Erkrankung, insbesondere der Anämie, führt.

Luspatercept ist der erste Vertreter der neuen Wirkstoffklasse der Erythrozyten-Reifungs-Aktivator (ERA)<sup>1</sup>, der zur aktiven Behandlung von erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, indiziert ist. Die Therapie mit Luspatercept wirkt gezielt gegen den der Anämie zugrundeliegenden Erythrozyten-Reifungs-Defekt und korrigiert damit die Anämie, die Kernsymptomatik bei Niedrigrisiko-MDS. Durch die Verstärkung der gesunden Erythropoese kann eine anhaltende Verbesserung der Anämie erreicht werden, die sich in einer Transfusionsfreiheit bzw. klinisch relevanten Reduktion der Transfusionslast widerspiegelt. Luspatercept verbessert die Anämie, das Kernproblem der Niedrigrisiko-MDS, und ermöglicht eine Vermeidung transfusionsbedingter Komplikationen. Die hohe Wirksamkeit bei gleichzeitig guter Verträglichkeit wird den Therapiezielen im Niedrigrisiko-MDS gerecht.

---

<sup>1</sup> im Englischen: *erythrocyte maturation agent* (EMA)

### ***Zusammenfassung der Ergebnisse zum medizinischen Nutzen und Zusatznutzen von Luspatercept***

Der medizinische Nutzen und medizinische Zusatznutzen wurde anhand der Ergebnisse der RCT MEDALIST anhand von patientenrelevanten Endpunkten nachgewiesen. Da es sich bei Luspatercept um ein Arzneimittel zur Behandlung von seltenen Leiden gemäß § 35a Absatz 1 Satz 11 SGB V handelt, gilt der medizinische Zusatznutzen bereits durch die Zulassung als belegt.

Die Behandlung mit Luspatercept bei erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, bewirkt gegenüber Placebo, jeweils in Kombination mit einer bedarfsgerechten Behandlung mit EK-Transfusionen, ggf. in Kombination mit Eisenchelatoren, eine **Verbesserung der Anämie und Aufrechterhaltung der Lebensqualität sowie Autonomie** bei gleichzeitiger **Vermeidung von Transfusionen** und somit auch einer **Minimierung transfusionsassoziierter Risiken und Folgekomplikationen**. Dies ist insbesondere begründet auf:

- *einer schnell einsetzenden und langfristigen, klinisch relevanten Vermeidung von EK-Transfusionen:*
  - konsistent über alle betrachteten klinisch relevanten Zeiträume zeigte sich unter Luspatercept deutlich häufiger und schneller eine vollständige Vermeidung von EK-Transfusionen. Diese ist zum einen auf das Fehlen von Anämie-Symptomen zurückzuführen, die Anlass zu einer erneuten Transfusion darstellen würden, und entspricht zum anderen einer Vermeidung von transfusionsassozierten Folgekomplikationen:
    - Transfusionsfreiheit über mindestens acht Wochen:
      - relatives Risiko (RR) = 2,88 (95 %-KI: [1,56; 5,31])
      - HR = 3,92 (95 %-KI: [2,12; 7,25])
    - Transfusionsfreiheit über mindestens 12 Wochen:
      - RR = 3,56 (95 %-KI: [1,59; 7,99])
      - HR = 3,56 (95 %-KI: [1,75; 7,23])
    - Transfusionsfreiheit über mindestens 16 Wochen:
      - RR = 4,80 (95 %-KI: [1,51; 15,26])
      - HR = 4,25 (95 %-KI: [1,81; 9,98])

- Transfusionsfreiheit über mindestens 24 Wochen:
  - RR = 9,93 (95 %-KI: [1,36; 72,64])
  - HR = 5,78 (95 %-KI: [2,06; 16,17])
- darüber hinaus zeigte sich ein starker Vorteil hinsichtlich der Langfristigkeit der Vermeidung von Transfusionen, gemessen als längste Dauer eines transfusionsfreien Zeitraums: HR = 0,46 (95 %-KI: [0,33; 0,63])
- *einer langfristigen Verbesserung der Erythropoese:*
  - die Verbesserung des klinischen Zustands des Patienten findet Ausdruck in einer großen, statistisch signifikanten und klinisch relevanten Stabilisierung des Hb-Werts, die über den gesamten Studienverlauf anhielt und die effektive Verstärkung der Erythropoese und Vermeidung der Anämie widerspiegelt:
    - LS-MWD = 0,54 (95 %-KI: [0,34; 0,74])
    - SMD = 0,61 (95 %-KI: [0,33; 0,89])
- *einer deutlichen Überlegenheit hinsichtlich des Erreichens eines erythroiden Ansprechens (HI-E):*
  - auch das klinisch relevante Therapieziel HI-E, das die Verbesserung der MDS-bedingten erythroiden Zytopenie (die Anämie) unter Berücksichtigung von sowohl klinisch relevanten Anhebung des Hb-Spiegels als auch Senkung der Transfusionslast abbildet, wurde unter Luspatercept sowohl häufiger als auch schneller erreicht:
    - RR = 4,47 (95 %-KI: [2,38; 8,41])
    - HR = 4,37 (95 %-KI: [2,48; 7,70])
  - darüber hinaus zeigte sich unter Luspatercept auch eine deutliche, klinisch relevante zeitliche Entlastung durch Verringerung der Häufigkeit notwendiger Transfusionstermine – im Mittel entspricht diese Entlastung einer Verringerung der Transfusionstermine von fast 19 Tagen auf knapp 13 Tage pro Jahr, also etwa einer Woche weniger Fehltagen bei der Arbeit bzw. aufzubringende Zeit für die Therapie:
    - LS-MWD = -1,78 (95 %-KI: [-2,46; -1,11])
    - SMD = -0,62 (95 %-KI: [-0,91; -0,32])

- *einer über den gesamten Studienzeitraum stabil aufrecht erhaltenen, guten Lebensqualität und somit Vermeidung der weiteren Beeinträchtigung der Lebensqualität durch den Krankheitsverlauf:*
  - die über den EORTC QLQ-C30 erhobene gesundheitsbezogene Lebensqualität blieb auch unter Luspatercept bei Minimierung der supportiven Maßnahmen, insbesondere der Transfusionstherapie, konstant gut
  - der Effekt der Vermeidung von Transfusionen auf den Alltag der Patienten zeigte sich insbesondere im signifikant höheren Anteil an Patienten unter Luspatercept, die zu Woche 24 keine Beeinträchtigung des Alltags durch die Transfusionsabhängigkeit wahrnahmen; der Anteil im Luspatercept-Arm betrug 44,5 % im Vergleich zu 20,8 % im Placebo-Arm (p = 0,003)
- *einem gleichzeitig günstigen Sicherheits- und Verträglichkeitsprofil, das zu keiner Einschränkung in der Behandlung führt.*

Die Ergebnisse aus MEDALIST zum medizinischen Nutzen und Zusatznutzen sind in Tabelle 4-1 zusammengefasst.

Tabelle 4-1: Zusammenfassung der Therapieeffekte von Luspatercept

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
<b>Mortalität</b>			
<b>Gesamtüberleben</b> (bis zum Datenschnitt vom 01.07.2019)	L: 27 (17,6)/153; mOS = N.E. P: 18 (23,7)/76; mOS = N.E.	HR = 0,75 [0,41; 1,38]	0,357
<b>Morbidität</b>			
<b>Transfusionsfreiheit</b> (Woche 1 – 24)			
≥ 8 Wochen	L: 58 (37,9)/153 P: 10 (13,2)/76	RR = 2,88 [1,56; 5,31]	< 0,001
≥ 12 Wochen	L: 43 (28,1)/153 P: 6 (7,9)/76	RR = 3,56 [1,59; 7,99]	< 0,001
≥ 16 Wochen	L: 29 (19,0)/153 P: 3 (3,9)/76	RR = 4,80 [1,51; 15,26]	0,002
≥ 24 Wochen	L: 20 (13,1)/153 P: 1 (1,3)/76	RR = 9,93 [1,36; 72,64]	0,003

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
<i>Regulation der Erythropoese</i> <i>Hb-Wert [g/dl]/16 Wochen</i>	L: 152 (99,3)/153 P: 75 (98,7)/76	LS-MWD = 0,54 [0,34; 0,74]	< 0,001
<b>Reduktion der Transfusionslast</b> (HI-E modifiziert gemäß IWG 2006 <sup>(4)</sup> ; Woche 1 – 24)			
HI-E (gesamt)	L: 81 (52,9)/153 P: 9 (11,8)/76	RR = 4,47 [2,38; 8,41]	< 0,001
HI-E I <sup>(5)</sup>	L: 52 (48,6)/107 P: 8 (14,3)/56	RR = 3,40 [1,74; 6,65]	< 0,001
HI-E II <sup>(6)</sup>	L: 29 (63,0)/46 P: 1 (5,0)/20	RR = 12,61 [1,84; 86,27]	< 0,001
<i>Reduktion der Eisenüberladung</i> (Woche 1 – 24)			
<i>sFerritin</i>	L: 153 (100)/153 P: 76 (100)/76	LS-MWD = -169,05 [-303,60; -34,50]	0,014
<i>Bedarf an Eisenchelatoren</i>	L: 128 (83,7)/153 P: 68 (89,5)/76	LS-MWD = -41,0 [-120,3; 38,2]	0,309
<b>Symptomatik</b> (EORTC QLQ-C30; Woche 1 – 24)			
Fatigue	L: 32 (21,5)/149 P: 24 (31,6)/76	RR = 0,68 [0,43; 1,07]	0,099
Übelkeit und Erbrechen	L: 17 (11,4)/149 P: 5 (6,6)/76	RR = 1,73 [0,67; 4,52]	0,254
Schmerz	L: 25 (16,8)/149 P: 14 (18,4)/76	RR = 0,91 [0,50; 1,65]	0,758
Dyspnoe	L: 24 (16,1)/149 P: 16 (21,1)/76	RR = 0,77 [0,43; 1,35]	0,360
Schlaflosigkeit	L: 28 (18,8)/149 P: 18 (23,7)/76	RR = 0,79 [0,47; 1,34]	0,390
Appetitminderung	L: 22 (14,8)/149 P: 9 (11,8)/76	RR = 1,25 [0,60; 2,57]	0,548
Obstipation	L: 31 (20,8)/149 P: 13 (17,1)/76	RR = 1,22 [0,68; 2,18]	0,509
Durchfall	L: 11 (7,4)/149 P: 6 (7,9)/76	RR = 0,94 [0,36; 2,43]	0,891
Finanzielle Schwierigkeiten	L: 13 (8,7)/149 P: 9 (11,8)/76	RR = 0,74 [0,33; 1,65]	0,458

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
<b>Hospitalisierungen</b> (Woche 1 – 24)	L: 34 (22,2)/153 P: 17 (22,4)/76	RR = 0,99 [0,60; 1,64]	0,977
<b>Lebensqualität</b>			
<b>EORTC QLQ-C30</b> (Woche 1 – 24)			
Körperliche Funktion	L: 25 (16,8)/149 P: 18 (23,7)/76	RR = 0,71 [0,41; 1,21]	0,214
Rollenfunktion	L: 30 (20,1)/149 P: 18 (23,7)/76	RR = 0,85 [0,51; 1,42]	0,539
Emotionale Funktion	L: 18 (12,1)/149 P: 11 (14,5)/76	RR = 0,83 [0,42; 1,68]	0,613
Kognitive Funktion	L: 29 (19,5)/149 P: 14 (18,4)/76	RR = 1,06 [0,59; 1,88]	0,851
Soziale Funktion	L: 26 (17,4)/149 P: 16 (21,1)/76	RR = 0,83 [0,47; 1,45]	0,512
Gesundheitlicher Status insgesamt	L: 31 (20,8)/149 P: 12 (15,8)/76	RR = 1,32 [0,72; 2,42]	0,367
<b>QoL-E</b> (Woche 1 – 24)			
Körperliches Wohlbefinden	L: 110 (73,8)/149 P: 53 (69,7)/76	SMD = -0,23 [-0,51; 0,05]	0,125
Funktionales Wohlbefinden	L: 108 (72,5)/149 P: 53 (69,7)/76	SMD = -0,19 [-0,47; 0,09]	0,184
Soziales/familiäres Wohlbefinden	L: 102 (68,5)/149 P: 48 (63,2)/76	SMD = -0,21 [-0,51; 0,08]	0,070
Sexuelles Wohlbefinden	L: 105 (70,5)/149 P: 51 (67,1)/76	SMD = 0,00 [-0,29; 0,28]	0,976
Fatigue	L: 110 (73,8)/149 P: 53 (69,7)/76	SMD = -0,34 [-0,62; -0,07]	0,015
MDS-bezogene Störungen	L: 109 (73,2)/149 P: 53 (69,7)/76	SMD = -0,09 [-0,36; 0,19]	0,492
Allgemeines Wohlbefinden	L: 100 (67,1)/149 P: 48 (63,2)/76	SMD = -0,29 [-0,58; 0,00]	0,015
Gesamtsummenskala	L: 100 (67,1)/149 P: 48 (63,2)/76	SMD = -0,24 [-0,53; 0,06]	0,062
Behandlungseffekt	L: 107 (71,8)/149 P: 53 (69,7)/76	SMD = -0,22 [-0,50; 0,06]	0,099

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
<b>Sicherheit</b>			
<i>Jegliche UE (Woche 1 – 24)</i>	L: 145 (94,8)/153 P: 70 (92,1)/76	RR = 1,03 [0,95; 1,11]	0,429
<i>SOC: Gastrointestinale Erkrankungen</i>	L: 78 (51,0)/153 P: 25 (32,9)/76	RR = 1,55 [1,08; 2,21]	0,010
<i>SOC: Erkrankungen des Nervensystems</i>	L: 60 (39,2)/153 P: 17 (22,4)/76	RR = 1,75 [1,10; 2,79]	0,011
<i>SOC: Augenerkrankungen</i>	L: 12 (7,8)/153 P: 1 (1,3)/76	RR = 5,96 [0,79; 45,00]	0,045
<i>PT: Schwindelgefühl</i>	L: 23 (15,0)/153 P: 4 (5,3)/76	RR = 2,86 [1,02; 7,96]	0,031
<i>PT: Asthenie</i>	L: 28 (18,3)/153 P: 6 (7,9)/76	RR = 2,32 [1,00; 5,36]	0,037
<b>SUE</b> (Woche 1 – 24)	L: 40 (26,1)/153 P: 16 (21,1)/76	RR = 1,24 [0,75; 2,07]	0,400
<b>Schwere UE</b> (CTCAE ≥ 3; Woche 1 – 24)	L: 54 (35,3)/153 P: 27 (35,5)/76	RR = 0,99 [0,69; 1,44]	0,972
SOC: Erkrankungen des Nervensystems	L: 8 (5,2)/153 P: 0 (0,0)/76 <sup>(7)</sup>	RR = 8,50 [0,50; 145,34]	0,043
<b>Therapieabbrüche aufgrund von UE</b> (Woche 1 – 24)	L: 12 (7,8)/153 P: 4 (5,3)/76	RR = 1,49 [0,50; 4,47]	0,472
Datenschnitt vom 08.05.2018 (außer anders angegeben). Ergänzende, nicht zur Ableitung des Zusatznutzen herangezogene Analysen ( <i>kursiv dargestellt</i> )			
<p><sup>(1)</sup> Anzahl Studienteilnehmer mit einem Ereignis.</p> <p><sup>(2)</sup> Anzahl analysierter Studienteilnehmer. Für die Endpunkte Gesamtüberleben, Transfusionsfreiheit, Reduktion der Transfusionslast, Reduktion der Eisenüberladung und Hospitalisierungen wurde die ITT-Population herangezogen, die alle randomisierten Studienteilnehmer entsprechend ihrer Zuteilung umfasst. Symptomatik und Lebensqualität wurden in der HRQoL-Population analysiert, welche alle Studienteilnehmer umfasst, die zur Baseline und zu mindestens einem Zeitpunkt im Studienverlauf eine Erhebung aufwiesen. Die Sicherheitsendpunkte wurden anhand der Safety-Population analysiert, welche alle randomisierten Studienteilnehmer umfasst, die mindestens eine Dosis der Studienmedikation erhalten haben. Studienteilnehmer wurden dem Behandlungsarm zugeordnet, deren Behandlung sie auch tatsächlich erhalten haben.</p> <p><sup>(3)</sup> Unter BSC wird insbesondere eine bedarfsgerechte Behandlung mit EK-Transfusionen und Eisenchelatoren verstanden (siehe auch Tabelle 4-13).</p> <p><sup>(4)</sup> In MEDALIST wurde ein HI-E für Patienten mit Baseline-Transfusionslast weniger als vier EK-Einheiten pro acht Wochen nur erreicht, wenn gleichzeitige zur Erhöhung des Hb-Werts auch eine Transfusionsfreiheit in den betrachteten acht Wochen vorlag.</p> <p><sup>(5)</sup> HI-E I = erythroides Ansprechen für Patienten mit einer Transfusionslast zur Baseline von mindestens vier EK-Einheiten pro acht Wochen, definiert als Reduktion der Transfusionslast um mindestens vier EK-Einheiten pro acht Wochen.</p> <p><sup>(6)</sup> HI-E II = erythroides Ansprechen für Patienten mit einer Transfusionslast zur Baseline von weniger als vier EK-Einheiten pro acht Wochen, definiert als Erhöhung des Hb-Werts um ≥ 1,5 g/dl bei gleichzeitiger Transfusionsfreiheit über acht Wochen.</p> <p><sup>(7)</sup> Nullzellenkorrektur zur Berechnung der Effektschätzer wurde durchgeführt.</p>			

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
AML = akute myeloische Leukämie; BSC = Best-Supportive-Care; CTCAE = <i>Common Terminology Criteria for Adverse Events</i> ; EK = Erythrozytenkonzentrat; EORTC QLQ-C30 = <i>European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module</i> ; Hb = Hämoglobin; HI-E = erythroides Ansprechen ( <i>hematologic improvement-erythroid</i> ); HR = Hazard Ratio; HRQoL = gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> ); ITT = Intention-to-Treat; IWG = <i>International Working Group</i> ; KI = Konfidenzintervall; L = Luspatercept; N.E. = nicht erreicht oder nicht berechnet; P = Placebo; PT = bevorzugter Begriff ( <i>preferred term</i> ); QoL-E = <i>Quality of Life-E</i> ; RR = relatives Risiko; sFerritin = Serum-Ferritin; SMD = standardisierte Mittelwertdifferenz; SOC = Systemorganklasse ( <i>system organ class</i> ); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis Quelle: [1, 18-20]			

Bezogen auf die einzelnen Endpunktkategorien ist das Ausmaß des Zusatznutzens für Luspatercept zur Behandlung von erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, maßgeblich wie folgt begründet:

### **Mortalität**

#### *Gesamtüberleben*

Insgesamt traten bis zum Datenschnitt vom 01.07.2019 wenig Todesfälle auf (17,6 % der Patienten (27/153) im Luspatercept-Arm und 23,7 % der Patienten (18/76) im Placebo-Arm), das mediane Gesamtüberleben wurde nicht erreicht und statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen lagen nicht vor (siehe Tabelle 4-21).

#### *Zusammenfassende Bewertung der Mortalität*

Insgesamt zeigten sich für die Endpunktkategorie Mortalität, gemessen als Gesamtüberleben, keine signifikanten Unterschiede zwischen Luspatercept und Placebo, jeweils in Kombination mit BSC.

### **Morbidität**

In der Endpunktkategorie Morbidität wurde der Zusatznutzen anhand der patientenrelevanten Endpunkte Transfusionsfreiheit, Reduktion der Transfusionslast (gemessen als HI-E gemäß IWG 2006), Symptomatik (gemessen anhand des EORTC QLQ-C30-Fragebogens) und Hospitalisierungen bewertet. Klinisch relevante Parameter der Anämie (Hb-Wert) und Eisenüberladung (Serum-Ferritin (sFerritin), Bedarf an Eisenchelatoren), welche im vorliegenden Anwendungsgebiet einen besonderen Stellenwert haben, werden ergänzend berichtet.

### *Transfusionsfreiheit*

Im vorliegenden Anwendungsgebiet sind Patienten aufgrund mangelnder alternativer Therapieoptionen auf Transfusionen mit EK angewiesen. Die Transfusionstherapie ist zwar notwendig, um die Anämie-Symptome temporär zu vermeiden, birgt aber das Risiko teils schwerwiegender Folgekomplikationen, wie die Eisenüberladung, und vermag keine stabile Verbesserung der Anämie zu erreichen. Aufgrund dieser Limitationen geht die Transfusionstherapie mit erheblichen psychosozialen und zeitlichen Belastungen für den Patienten einher und stellt nur eine unzureichende therapeutische Maßnahme dar.

Transfusionsbedingte Belastungen können sowohl akut (z. B. Zeitaufwand, psychosoziale Effekte und potentielle akute Nebenwirkungen) als auch langfristig (z. B. Folgekomplikationen und notwendige Komedikationen aufgrund einer transfusionsbedingten Eisenüberladung) spürbar sein und sich u. a. in bedeutsamen Einschränkungen im alltäglichen Leben sowie einer Verringerung des Wohlbefindens bis hin zu depressiven Verstimmungen widerspiegeln [16]. Laut einer Befragung von MDS-Patienten in Deutschland, den USA und Großbritannien würden viele Patienten sogar Lebenszeit „opfern“, um eine Transfusionsfreiheit zu erreichen, was die immense körperliche und psychische Belastung einer chronischen Transfusionstherapie und die Relevanz des Therapieziels einer Transfusionsfreiheit unterstreicht [17]. Da Transfusionen bei MDS-Patienten in Abhängigkeit des klinischen Zustands, d. h. in Abhängigkeit von Symptomen, verabreicht werden, entspricht eine Transfusionsfreiheit nicht nur einer Vermeidung akuter und langfristiger Transfusionskomplikationen, sondern auch einem Fehlen von Anämie-Symptomen, also einer Verbesserung der Grunderkrankung. Das Erreichen einer Transfusionsfreiheit stellt somit einen patientenrelevanten Endpunkt und ein bedeutendes Therapieziel im vorliegenden Anwendungsgebiet dar. Die Patientenrelevanz dieses Endpunkts wurde auch bereits seitens des Instituts für Wirtschaftlichkeit im Gesundheitswesen (IQWiG) und des Gemeinsamen Bundesausschusses (G-BA) in vorangegangenen Verfahren anerkannt [21-23].

Unter Luspatercept erreichten 37,9 % der Patienten (58/153) eine Transfusionsfreiheit über mindestens acht Wochen. Dagegen blieben unter Placebo nur 13,2 % der Patienten (10/76) über mindestens acht Wochen transfusionsfrei. Insgesamt erreichten im Luspatercept-Arm fast dreimal mehr Patienten eine Transfusionsfreiheit über mindestens acht Wochen (RR = 2,88; 95 %-KI: [1,56; 5,31]). Bezogen auf eine Transfusionsfreiheit über  $\geq 24$  Wochen zeigte sich die Überlegenheit von Luspatercept gegenüber Placebo noch deutlicher: 13,1 % der Patienten (20/153) unter Luspatercept und nur 1,3 % der Patienten (1/76) unter Placebo erreichten eine Transfusionsfreiheit über diesen langen Zeitraum. Dementsprechend erreichten Patienten unter Luspatercept dieses Therapieziel fast zehnmal häufiger als unter Placebo (RR = 9,93; 95 %-KI: [1,36; 72,64]). Dies entspricht einer absoluten Risikodifferenz (Luspatercept – Placebo) von 24,8 %-Punkten (bezogen auf eine Transfusionsfreiheit über mindestens acht Wochen; 95 %-KI: [13,94; 35,56]) bzw. 11,8 %-Punkten (bezogen auf eine Transfusionsfreiheit über  $\geq 24$  Wochen; 95 %-KI: [5,83; 17,68]) zugunsten der mit Luspatercept behandelten Patienten. Unabhängig davon, wie lange die Minstdauer der Transfusionsfreiheit definiert war (8, 12, 16 oder 24 Wochen), ergab sich in allen Fällen ein statistisch signifikanter ( $p < 0,05$ ) und im Ausmaß großer Unterschied zwischen den beiden Behandlungsarmen.

Die ergänzend dargestellten Ergebnisse der Zeit bis zur Transfusionsfreiheit bestätigen darüber hinaus, dass nicht nur mehr Patienten unter Luspatercept eine Transfusionsfreiheit erreichten, sondern dass dieser Effekt auch sehr schnell und im Vergleich zu Placebo signifikant früher eintrat (bezogen auf eine Transfusionsfreiheit über mindestens acht Wochen: HR = 3,92; 95 %-KI: [2,12; 7,25]; bezogen auf eine Transfusionsfreiheit über  $\geq 24$  Wochen: HR = 5,78; 95 %-KI: [2,06; 16,17]). Die mediane Zeit bis zur Transfusionsfreiheit über mindestens acht Wochen in den ersten 24 Wochen bzw. 48 Wochen der Behandlung bei alleiniger Berücksichtigung von Patienten, die dieses Ereignis auch erreichten, betrug einen Tag bzw. zwei Tage unter Luspatercept und 17 Tage bzw. 22,5 Tage unter Placebo [1].

#### Dauer transfusionsfreier Intervalle und multiples Ansprechen

Weitere Analysen zur Dauer der transfusionsfreien Intervalle zeigen, dass unabhängig vom Erreichen einer Transfusionsfreiheit das Risiko einer erneuten EK-Transfusion unter Luspatercept um fast 55 % reduziert wird (HR = 0,46; 95 %-KI: [0,33; 0,63]). Dies bedeutet, dass mit Luspatercept behandelte Patienten deutlich seltener transfundiert werden müssen und der Zeitraum zwischen zwei aufeinanderfolgenden Transfusionszeitpunkten signifikant verlängert wird.

Der positive Effekt von Luspatercept auf die Verlängerung von transfusionsfreien Intervallen spiegelt sich auch in den dargestellten Swimmer-Plots wider. In dieser Abbildung wird zudem eine Besonderheit der Behandlung mit Luspatercept deutlich: Selbst bei einem erneuten Transfusionsbedarf nach dem Erreichen einer *ersten* Transfusionsfreiheit über mindestens acht Wochen können Patienten anschließend erneut ansprechen, d. h. werden wieder transfusionsfrei. Unter Placebo erzielten die meisten Patienten, falls sie überhaupt ansprachen, hingegen nur ein einmaliges Ansprechen; mehr als zwei Intervalle einer Transfusionsfreiheit wurden nicht beobachtet.

Luspatercept führt zu einer anhaltenden Förderung der Erythropoese. Damit ergibt sich ein erheblicher Vorteil für auf Luspatercept ansprechende Patienten: transfusionsfreie Intervalle über mindestens acht Wochen werden über einen längeren Zeitraum als den studiendefinierten primären Endpunkt (*erste* Transfusionsfreiheit über mindestens acht Wochen) erreicht. Aus der bisherigen Erfahrung mit transfusionspflichtigen Patienten mit MDS ist die „erstmalige erneute Transfusion nach Erreichen einer Transfusionsfreiheit“ üblicherweise gleichbedeutend mit „Wirkungsverlust“ (siehe z. B. Studien zur Wirksamkeit von Lenalidomid [24, 25] und Epoetin-alfa [26]). Dieses Paradigma ist mit Luspatercept nicht mehr gültig.

Insgesamt zeigten sich hinsichtlich des Erreichens einer Transfusionsfreiheit folgende Effekte:

Unter Luspatercept

- erreichten deutlich mehr Patienten eine langfristige Transfusionsfreiheit, die zudem deutlich früher eintrat,
- erreichten mehr Patienten mehrfach einer Transfusionsfreiheit und
- hielt die Transfusionsfreiheit deutlich länger an

als unter Placebo.

#### Verbesserung der Erythropoese

Der Hb-Wert spiegelt die Funktionsfähigkeit der Erythropoese wider, wodurch sich das Ausmaß einer Anämie, d. h. der Mangel an Erythrozyten, bemessen lässt. Die stabile Anhebung des Hb-Werts ist somit von zentraler Bedeutung und eine Vermeidung von Transfusionen kann nur erreicht und aufrechterhalten werden, solange keine Anämie-Symptome auftreten, d. h. ein adäquater Hb-Wert nicht unterschritten wird. Darüber hinaus steht ein niedriger Hb-Wert mit einer erhöhten Mortalität im Zusammenhang [27].

Die ergänzend dargestellten Ergebnisse zum Hb-Spiegel belegen somit, dass Luspatercept effektiv die Erythropoese verbessert, die Erythrozyten-Reifung aktiviert und somit eine Verbesserung des Hb-Spiegels und der Anämie bewirkt. Dieser Effekt setzte bereits während des ersten Behandlungszyklus mit Luspatercept ein und war über den gesamten Studienzeitraum hinweg stabil. Im vorliegenden Kontext ist diese Regulation der Erythropoese insbesondere relevant, als dass nur durch eine stabile Erhöhung des Hb-Werts eine Transfusionsfreiheit erreicht werden kann. Anders ausgedrückt bedeutet dies, dass eine Transfusionsfreiheit nicht auf Kosten des Hb-Werts und der Anämie erreicht wird, sondern die effektive Aktivierung der Erythrozyten-Reifung widerspiegelt.

### *Reduktion der Transfusionslast*

Neben dem Erreichen einer kompletten Transfusionsfreiheit stellt insbesondere für Patienten mit einer hohen Transfusionslast bereits deren Reduktion ein patientenrelevantes Therapieziel dar. Ein restriktiver Einsatz von Blutprodukten entspricht auch den Empfehlungen der Bundesärztekammer (BÄK) in deren Querschnitts-Leitlinie und Richtlinie zum Umgang mit Blutprodukten, nach denen im Rahmen der patientenindividuellen Hämotherapie (*Patient Blood Management*) die Transfusionstherapie erst eingesetzt werden soll, wenn andere Maßnahmen nicht geeignet sind [28, 29]. Durch Minimierung der Exposition von MDS-Patienten mit Fremdblut im Rahmen der Reduktion der Transfusionslast werden nicht nur die Empfehlungen der BÄK umgesetzt. Das Therapieziel einer Reduktion der Transfusionslast geht auch mit einer Verbesserung der erythroiden Zytopenie (Anämie aufgrund der MDS-Grundkrankung) und einer Reduktion transfusionsassoziiierter, akuter wie langfristiger Risiken einher. Zur Bewertung klinisch relevanter Effekte in MDS-Studien wurden durch die internationale Arbeitsgruppe medizinischer Experten (IWG) im Jahr 2006 Kriterien des erythroiden Ansprechens (HI-E) etabliert.

### Erythroides Ansprechen (HI-E) gemäß IWG-Kriterien aus dem Jahr 2006

Unter Luspatercept zeigten sich statistisch signifikante und klinisch relevante Effekte hinsichtlich einer Reduktion der Transfusionslast gemessen als HI-E gemäß den etablierten Kriterien der IWG aus dem Jahr 2006. In Abhängigkeit der Transfusionslast entspricht ein HI-E einer klinisch relevanten Reduktion der transfundierten EK-Einheiten um mindestens vier EK-Einheiten pro acht Wochen und/oder einer stabilen Erhöhung des Hb-Werts über acht Wochen um  $\geq 1,5$  g/dl (in MEDALIST wurde eine Erhöhung des Hb-Werts nur bei gleichzeitiger Transfusionsfreiheit als Ansprechen gewertet). Sowohl hinsichtlich eines HI-E insgesamt (RR = 4,47; 95 %-KI: [2,38; 8,41]), als auch bei Betrachtung des HI-E innerhalb der Populationen mit unterschiedlicher Transfusionslast vor Behandlungsbeginn, zeigten sich konstant statistisch signifikante Vorteile für Luspatercept, die in ihrem Ausmaß einen bedeutenden Vorteil belegen.

Darüber hinaus trat ein HI-E bei mit Luspatercept behandelten Patienten nicht nur häufiger, sondern auch deutlich früher auf als unter Placebo (HI-E (gesamt): HR = 4,37; 95 %-KI: [2,48; 7,70]).

### Mittlere Änderung der Transfusionslast und Transfusionshäufigkeit

Auswertungen zur mittleren Änderung der Transfusionslast und Transfusionshäufigkeit bestätigen zudem die deutlichen Effekte hinsichtlich einer Reduktion von EK-Transfusionen. So benötigt ein transfusionsabhängiger Patient mit einer durchschnittlichen Transfusionslast von 11 EK-Einheiten/16 Wochen, welche im Mittel alle drei Wochen verabreicht werden, unter Luspatercept etwa ein Drittel weniger EK-Einheiten (LS-MWD = -3,53; 95 %-KI: [-4,70; -2,36]; SMD = -0,71; 95 %-KI: [-1,00; -0,41];  $p < 0,001$ ) und muss weniger häufig Transfusionen erhalten (Verlängerung des Zeitraums zwischen zwei Transfusionszeitpunkten um fast das Doppelte; LS-MWD = -1,78; 95 %-KI: [-2,46; -1,11]; SMD = -0,62; 95 %-KI: [-0,91; -0,32];  $p < 0,001$ ).

### *Reduktion der Eisenüberladung*

Bei chronisch transfundierten Patienten ergeben sich mit zunehmender Dauer der Transfusions-therapie Komplikationen durch die Zufuhr des in den EK enthaltenen Eisens, da der menschliche Körper die durch die Transfusionen aufgenommenen Mengen an Eisen nicht in ausreichendem Maße ausscheiden kann. Die zunehmende Eisenüberladung des Organismus führt schließlich zur toxischen Ablagerung von überschüssigem Eisen in verschiedenen Organen, insbesondere im Herz, in der Leber und in endokrinen Organen (sekundäre Hämosiderose), und den damit verbundenen, teils schwerwiegenden Komplikationen [9, 30]. Ab einer bestimmten Transfusionslast (i. d. R. ab 20 erhaltenen EK-Transfusionen) bzw. einem sFerritin-Spiegel > 1.000 ng/ml sollte zur Vermeidung von Siderose-bedingten Schäden eine Eisenchelationstherapie in Erwägung gezogen werden [2]. Effektiver als die medikamentöse Eisenchelation ist jedoch die Vermeidung der übermäßigen Eisenzufuhr, also Vermeidung von Transfusionen.

Die ergänzend dargestellten Ergebnisse zur Reduktion der Eisenüberladung spiegeln das klinisch relevante Potential der Eiseneinsparung durch die reduzierte Eisenaufnahme über EK-Transfusionen bei Erreichen einer Transfusionsfreiheit bzw. eines erythroiden Ansprechens wider. So zeigte sich unter Luspatercept eine statistisch signifikant höhere Reduktion des sFerritin-Spiegels, welcher zur Verlaufskontrolle des Eisenhaushalts herangezogen wird, als unter Placebo (LS-MWD = -169,05; 95 %-KI: [-303,60; -34,50]). Auch für den Bedarf an Eisenchelatoren zeigten sich bereits positive Effekte hinsichtlich einer Reduktion der täglichen Dosis unter Luspatercept im Vergleich zu Placebo. Da die Dosierung von Eisenchelatoren in Abhängigkeit des sFerritin-Spiegels erfolgt, sind Dosismodifikationen vor allem im langfristigen Verlauf zu erwarten.

### *Symptomatik*

Untersuchungen zur Verbesserung der Symptomatik, gemessen als Verbesserung der Symptomskalen des EORTC QLQ-C30-Fragebogens um die validierte klinische Relevanzschwelle (*minimal clinically important difference*, MCID) von zehn Punkten, zeigten keine klinisch relevanten Unterschiede zwischen den beiden Behandlungsarmen. Die Symptomatik war somit zwischen Luspatercept bei gleichzeitiger Vermeidung von EK-Transfusionen und Placebo unter Beibehaltung der konventionellen Transfusionstherapie vergleichbar. Durch die adäquate Behandlung der Patienten im Rahmen der BSC war das Auftreten von Symptomen auch nicht zu erwarten gewesen.

### *Hospitalisierungen*

Hinsichtlich des Endpunkts Hospitalisierungen zeigten sich keine relevanten Unterschiede zwischen den beiden Behandlungsarmen. Hauptgrund für Hospitalisierungen waren in beiden Behandlungsarmen UE; einzelne UE, die vorrangig zu stationären Aufnahmen führten, konnten nicht ausgemacht werden. Auch bei Betrachtung der stationären Aufnahmen aufgrund von UE und/oder Transfusionen zeigten sich keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen.

### *Zusammenfassende Bewertung der Morbidität*

Im vorliegenden Anwendungsgebiet der Niedrigrisiko-MDS tritt bei ca. 70 – 80 % der Patienten als Erstmanifestation eine Anämie infolge eines Erythrozyten-Reifungs-Defekts auf, welche auch als Kernsymptom angesehen werden kann und im Zentrum aller Komplikationen steht [2-4]. Die Anämie beeinflusst den klinischen Zustand des Patienten maßgeblich und macht die Transfusionstherapie notwendig. Entsprechend ist eine Verbesserung der Anämie durch eine Korrektur des Erythrozyten-Reifungs-Defekts und somit eine Verbesserung des klinischen Zustands, der sich unmittelbar in einer Vermeidung oder zumindest Reduktion der Transfusionslast sowie stabilen Erhöhung der Hb-Werte widerspiegelt, das höchste Therapieziel. Langfristig kann dadurch auch der Eisenhaushalt normalisiert werden und das Gesamtüberleben verlängert werden.

Zusammengefasst zeigten sich für Luspatercept in der Kategorie Morbidität ausschließlich positive Ergebnisse und eine deutliche Überlegenheit gegenüber Placebo hinsichtlich des Erreichens einer Transfusionsfreiheit und einer Reduktion der Transfusionslast. Diese patientenrelevanten Therapieziele konnten unter Luspatercept nicht nur häufiger, sondern auch deutlich schneller erreicht werden als unter Placebo. Der Verbesserung des klinischen Zustands der Patienten liegt eine effektive Korrektur des Erythrozyten-Reifungs-Defekts und Verbesserung der Erythropoese zugrunde, welche sich in stabil erhöhten Hb-Werten – auch bei Reduktion bzw. Vermeidung von EK-Transfusionen – widerspiegelt. Eine Vermeidung von Transfusionen entspricht somit einer Stabilisierung des Hb-Werts und Vermeidung von Symptomen, die Anlass für eine EK-Transfusion darstellen würden, und bedeutet zugleich eine Vermeidung von transfusionsbedingten Nebenwirkungen und Folgekomplikationen. Das Potential von Luspatercept hinsichtlich einer Reduktion der sekundären Eisenüberladung durch Einsparung an EK-Transfusionen, d. h. langfristiger Vermeidung einer sekundären Hämosiderose und damit verbundenen Komplikationen, zeichnet sich zudem bereits in den Ergebnissen zum sFerritin-Spiegel ab. Statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen hinsichtlich Symptomatik und Hospitalisierungen lagen nicht vor.

Unter Luspatercept zeigte sich eine bedeutsame Verbesserung der Anämie und deren Kontrolle. Insgesamt liegt somit ein erheblicher Vorteil von Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, hinsichtlich der Endpunktkategorie Morbidität vor.

### ***Lebensqualität***

Da es sich bei MDS um eine hämatologisch-onkologische Erkrankung handelt, wurde die Lebensqualität in MEDALIST über den generischen Fragebogen EORTC QLQ-C30 sowie den krankheitsspezifischen Fragebogen QoL-E erhoben, der insbesondere den Einfluss der Transfusionstherapie auf die Lebensqualität transfusionsabhängiger MDS-Patienten erfasst.

Die durch die ineffektive Erythropoese bedingte Anämie kann die Lebensqualität von MDS-Patienten maßgeblich beeinträchtigen [8, 17, 31-33]. Die Erhaltung oder Verbesserung der Lebensqualität ist entsprechend ein relevantes Therapieziel im vorliegenden Anwendungsgebiet [2, 5, 6, 34]. In MEDALIST war in beiden Behandlungsarmen eine adäquate supportive Therapie, insbesondere die bedarfsgerechte Transfusionstherapie mit EK, ggf. in Kombination mit Eisenchelatoren, erlaubt, die geeignet ist, die Beeinträchtigungen der Lebensqualität durch die Anämie zu minimieren.

Das Therapieziel, die Lebensqualität aufrecht zu erhalten, wurde unter Luspatercept auch bei einer Minimierung der supportiven Maßnahmen, insbesondere der Transfusionstherapie, erreicht. Die Lebensqualität war über den gesamten Studienverlauf hinweg konstant, klinisch relevante Unterschiede zwischen den beiden Behandlungsarmen lagen nicht vor.

Zu betonen ist der direkte Einfluss der Vermeidung von Transfusionen unter Luspatercept auf die Wahrnehmung der Beeinträchtigungen durch die Transfusionsabhängigkeit im Alltag. Fast die Hälfte aller mit Luspatercept behandelten Patienten (44,5 % (49/110)) gaben zu Woche 24 an, im Alltag durch die Transfusionsabhängigkeit überhaupt nicht beeinträchtigt zu sein. Im Placebo-Arm waren es nur 20,5 % (11/53) der Patienten, die diese Einschätzung teilten. Die Anteile zur Baseline lagen bei 14,8 % und 25,0 % im Luspatercept- und Placebo-Arm. Der Unterschied zu Woche 24 war statistisch signifikant und entsprach einem RR von 2,15 (95 %-KI: [1,22; 3,78]) zugunsten von Luspatercept.

### ***Zusammenfassende Bewertung der Lebensqualität***

Insgesamt zeigten sich für die Endpunktkategorie Lebensqualität keine klinisch relevanten Unterschiede zwischen den beiden Behandlungsarmen. Die Lebensqualität wurde in beiden Behandlungsarmen über den gesamten Studienverlauf stabil aufrechterhalten und war vergleichbar. Die Vermeidung von Transfusionen durch Luspatercept schlägt sich in einer für die Patienten spürbaren Vermeidung alltäglicher transfusionsbedingter Beeinträchtigungen nieder.

### **Sicherheit**

In der vorliegenden RCT, in der die aktive und wirksame Substanz Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, verglichen wurde, ist vereinzelt mit mehr unerwünschten Ereignissen zu rechnen. Unter der Behandlung mit Luspatercept zeigten sich gegenüber Placebo aber keine statistisch signifikanten Unterschiede in den Gesamtraten an jeglichen UE, SUE, schweren UE oder Therapieabbrüchen aufgrund von UE, sodass von einer sehr guten Verträglichkeit von Luspatercept im Vergleich zu Placebo und mit keiner Einschränkung für die mit Luspatercept-behandelten Patienten auszugehen ist.

Statistisch signifikante, aber als geringfügig und damit nicht relevant zu klassifizierende, Unterschiede lagen bezogen auf die UE-Kategorie jegliche UE für die Systemorganklasse (**system organ class, SOC**) „Gastrointestinale Erkrankungen“, „Erkrankungen des Nervensystems“, „Augenerkrankungen“ und die beiden bevorzugten Begriffe (*preferred term, PT*) „Asthenie“ und „Schwindelgefühl“ sowie für die UE-Kategorie schwere UE für die SOC „Erkrankungen des Nervensystems“ vor. Bezogen auf schwere UE und SUE gab es auch auf Ebene der SOC und PT keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen. Therapieabbrüche aufgrund von UE traten nur vereinzelt auf, statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen lagen auf Ebene der SOC und PT ebenfalls nicht vor.

In Abschnitt 4.3.1.3.1 sind UE von besonderem Interesse ergänzend dargestellt. MDS sind chronische Erkrankungen, die üblicherweise progredient verlaufen und zu einem erheblichen Prozentsatz in akute Leukämien übergehen. UE von besonderem Interesse waren daher definiert als prämaligne Erkrankungen und Malignitäten, erhoben in Form von *standardized MedDRA Queries (SMQ)*. Daneben wurde das Auftreten von Neutropenien und Thrombozytopenien erfasst. Diese sind relevante Risikofaktoren und typische Manifestationen von MDS und können zu Blutungen und Infektionen führen. Die Endpunkte Zeit bis zur Progression in eine AML sowie die Fortentwicklung der Erkrankung in ein Hochrisiko-MDS stellen weitere relevante Sicherheitsaspekte dar und wurden daher ebenfalls gesondert berichtet. Alle hier genannten Ereignisse traten insgesamt nur sehr selten auf, statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen lagen nicht vor. Ein negativer Einfluss von Luspatercept auf die Pathophysiologie und Progression der Erkrankung konnte nicht festgestellt werden.

### **Zusammenfassende Bewertung der Sicherheit**

Insgesamt zeigten sich für die Endpunktkategorie Sicherheit weder Hinweise auf einen größeren noch auf einen geringeren Nutzen von Luspatercept im Vergleich zum Placebo-Arm, jeweils in Kombination mit BSC. Das Sicherheits- und Verträglichkeitsprofil von Luspatercept ist somit als günstig einzustufen und führt zu keiner Einschränkung in der Behandlung.

## **Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen**

Die Ergebnisse zum medizinischen Nutzen und Zusatznutzen von Luspatercept beruhen auf der doppelblinden, randomisiert kontrollierten Phase-III-Studie MEDALIST. Die Aussagesicherheit der Ergebnisse ist aufgrund der hohen Evidenzstufe der Studie (1b) sowie des geringen Verzerrungspotentials auf Studien- und Endpunktebene als hoch einzustufen. Entsprechend lässt sich basierend auf den Ergebnissen aus MEDALIST ein Hinweis auf einen Zusatznutzen ableiten. Folgende Aspekte sind bei der Bewertung des Zusatznutzens relevant:

- Bei MDS handelt es sich um seltene Erkrankungen des blutbildenden Systems im Knochenmark mit einer erheblichen Krankheitslast und Beeinträchtigung der Lebensqualität. Dabei spielen sowohl erkrankungsbedingte als auch therapiebedingte, d. h. transfusionsbedingte klinische und psychosoziale Komplikationen eine große Rolle. Im Zentrum aller Komplikationen steht die chronische Anämie infolge eines Erythrozyten-Reifungs-Defekts, die auch als Kernproblem der Erkrankung angesehen werden kann. Diese bestimmt den klinischen Zustand der Patienten maßgeblich und macht eine Therapie notwendig.
- Für Patienten im Anwendungsgebiet von Luspatercept gibt es einen großen therapeutischen Bedarf nach einem Arzneimittel, das den Erythrozyten-Reifungs-Defekt korrigiert und dadurch die Anämie verbessert, die Transfusionslast senkt oder vollständig vermeidet und zu einer besseren Kontrolle der Erkrankung und Aufrechterhaltung der Lebensqualität und Autonomie der Patienten führt. Derzeit steht als einzige zugelassene Therapieoption die supportive Transfusionstherapie mit EK, ggf. in Kombination mit Eisenchelatoren, zur Verfügung. Diese kann den Mangel an körpereigenen reifen Erythrozyten zwar temporär kompensieren, den Hb-Wert transient erhöhen und die Symptomatik der Anämie vorübergehend mildern, korrigiert jedoch den Erythrozyten-Reifungs-Defekt und die ineffektive Erythropoese nicht. Eine anhaltende, stabile Korrektur des Hb-Spiegels und somit andauernde Vermeidung der Anämie-Symptomatik ist somit nicht zu erreichen. Darüber hinaus geht die Transfusionstherapie mit EK mit akuten Nebenwirkungen und langfristigen Folgekomplikationen einher.
- Aus Sicht der Patienten stellen regelmäßige Transfusionen eine Beeinträchtigung des alltäglichen Lebens dar. Eine Transfusionsunabhängigkeit oder zumindest Reduktion der Transfusionslast wird von MDS-Patienten ausdrücklich gewünscht. Langfristig können durch die Reduktion der Transfusionslast auch transfusionsassoziierte Folgekomplikationen vermieden und die Morbidität und Mortalität des Patienten gesenkt werden.
- Luspatercept ist der erste Vertreter der neuen Wirkstoffklasse der ERA und damit der erste innovative Therapieansatz, der gezielt gegen den der Anämie zugrundeliegenden Erythrozyten-Reifungs-Defekt wirkt und damit die Anämie, die Kernsymptomatik bei Niedrigrisiko-MDS verbessert. Damit deckt Luspatercept den großen therapeutischen Bedarf im vorliegenden Anwendungsgebiet.

- Mit Luspatercept können MDS-Patienten im Anwendungsgebiet:
  - schneller, häufiger und anhaltender eine Transfusionsfreiheit erreichen als mit Placebo. Dies bedeutet zum einen das Fehlen von Symptomen, die einen Anlass zu einer Transfusion darstellen, und zum anderen die Vermeidung von Folgekomplikationen, die im Rahmen einer Transfusionstherapie auftreten können. Für die Patienten geht eine Vermeidung von Transfusionen zudem mit einem Zeitgewinn und einer Vermeidung der transfusionsbedingten Belastung im Alltag einher.
  - die Anämie durch eine Korrektur des Erythrozyten-Reifungs-Defekts und Förderung der Erythropoese verbessern.
  - die generische Lebensqualität, die aufgrund von MDS oft beeinträchtigt ist, stabil aufrechterhalten. Die Anämie wird durch Luspatercept so gut kontrolliert und verbessert, dass die Patienten auch ohne Transfusionen keine relevante Beeinträchtigung ihrer Lebensqualität durch die Anämie verspüren.
  - eine spürbare Erleichterung hinsichtlich der transfusionsbedingten Störungen im Alltag erreichen.
  - all diese Wirksamkeitsparameter bei einem insgesamt günstigen Sicherheits- und Verträglichkeitsprofil erzielen, das zu keiner Einschränkung von mit Luspatercept behandelten Patienten führt.

Grundsätzlich gilt entsprechend § 5 Absatz 7 des 5. Kapitels der VerfO für Luspatercept zur Behandlung von erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von MDS mit Ringsideroblasten, mit sehr niedrigem, niedrigem oder intermediärem Risiko, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, als belegt. Aufgrund der hohen Evidenzstufe und niedrigen Verzerrung auf Studien- und Endpunktebene ist die Aussagesicherheit der Ergebnisse hoch. Die deutlichen, bisher nicht erreichten Vorteile in der Wirksamkeit bei vergleichbarer Lebensqualität und keiner Einschränkung durch UE belegen den Zusatznutzen von Luspatercept im Vergleich zu Placebo.

**Insgesamt lässt sich so das Ausmaß des Zusatznutzens unter Angabe der Aussagekraft der Nachweise quantifizieren und führt zu einem Hinweis auf einen erheblichen Zusatznutzen für Luspatercept.**

## 4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

### 4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

Die Benennung der Vergleichstherapie in Modul 4 muss zur Auswahl der zweckmäßigen Vergleichstherapie im zugehörigen Modul 3 konsistent sein.

*Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.*

In der vorliegenden Nutzenbewertung soll das Ausmaß des Zusatznutzens von Luspatercept (Reblozyl®) für erwachsene Patienten mit transfusionsabhängiger Anämie aufgrund von MDS mit Ringsideroblasten, mit sehr niedrigem, niedrigem oder intermediärem Risiko, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, anhand von patientenrelevanten Endpunkten untersucht werden.

Bei Luspatercept handelt es sich um ein Arzneimittel zur Behandlung eines seltenen Leidens gemäß § 35a Absatz 1 Satz 11 SGB V, sodass bereits durch die Zulassung der medizinische Zusatznutzen als belegt gilt. Nachweise über den medizinischen Nutzen und medizinischen Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie (zVT) müssen nicht vorgelegt werden.

Zur Ableitung des Ausmaßes des Zusatznutzens werden Ergebnisse aus randomisiert kontrollierten Studien (Evidenzstufe 1b gemäß § 5 Absatz 6 Verfo) hinsichtlich patientenrelevanter Endpunkte aus den Kategorien Mortalität, Morbidität, Lebensqualität und Sicherheit berücksichtigt. Da die Aussagekraft von Ergebnissen aus RCT generell höher einzustufen ist als aus nicht randomisierten oder nicht vergleichenden Studien, wird auf eine Darstellung von Ergebnissen aus Studien mit niedrigerer Evidenzstufe verzichtet.

#### 4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister/ einer Studienergebnisdatenbank erfolgen, während ein Kongressabstrakt allein in der Regel nicht für eine Studienbewertung ausreicht.

*Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar. Erstellen Sie dabei für unterschiedliche Themen der Recherche (z. B. unterschiedliche Fragestellungen) jeweils eine separate Übersicht.*

Um relevante Studien für die Beantwortung der in Abschnitt 4.2.1 formulierten Fragestellung zu identifizieren, wurden Ein- und Ausschlusskriterien hinsichtlich der Patientenpopulation, der Intervention, der erhobenen Endpunkte und des Studiendesigns definiert. Eine Übersicht der Ein- und Ausschlusskriterien findet sich in Tabelle 4-2.

Studien, die die genannten Kriterien nicht erfüllen, wurden ausgeschlossen. Studien, die die Kriterien nur teilweise erfüllten, wurden dann eingeschlossen, wenn mindestens 80 % der Studienpopulation die Kriterien erfüllte oder eine den Kriterien entsprechende Auswertung von Teilpopulation(en) möglich war [35].

Tabelle 4-2: Kriterien für den Ein- und Ausschluss von Studien in die Nutzenbewertung

Einschlusskriterien		Ausschlusskriterien	
E1	Erwachsene Patienten mit transfusionsabhängiger Anämie aufgrund von MDS mit Ringsideroblasten, mit sehr niedrigem, niedrigem oder intermediärem Risiko gemäß IPSS-R oder vergleichbarer Risikoklassifizierung, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind	A1	Population nicht wie in E1 definiert (d. h. Patientenpopulation, die nicht der Zulassung von Luspatercept entspricht)
E2	Luspatercept (1,0 mg/kg bis maximal 1,75 mg/kg Körpergewicht s. c. alle drei Wochen)	A2	Intervention nicht wie in E2 definiert
E3	Mindestens ein patientenrelevanter Endpunkt (Mortalität, Morbidität, Lebensqualität, Sicherheit) wurde berichtet (siehe auch Abschnitt 4.2.5.2.2)	A3	Keine patientenrelevanten Endpunkte berichtet
E4	Studie im RCT-Design oder Meta-Analyse aus RCT	A4	Keine Studie im RCT- Design oder daraus abgeleitete Meta-Analyse
E5	Studiendauer $\geq$ 24 Wochen	A5	Studiendauer $<$ 24 Wochen
E6	Berichterstattung liefert ausreichende Informationen zur Beurteilung von Methodik und Ergebnissen in Anlehnung an die CONSORT-Kriterien (z. B. Studienbericht, Volltextpublikation in Fachzeitschriften oder ausführliche Ergebnisdarstellung in Studienregister verfügbar) Einschließlich Meta-Analysen auf Grundlage der Studienrohdaten Publikation in Deutsch oder Englisch	A6	Berichterstattung liefert keine ausreichenden Informationen zur Beurteilung von Methodik und Ergebnissen (z. B. Konferenzbeitrag) Meta-Analysen auf Grundlage bereits publizierter Studiendaten (Sekundärliteratur) Publikation in anderer Sprache als Deutsch oder Englisch
CONSORT = <i>Consolidated Standards of Reporting Trials</i> ; ESA = Erythropoese-stimulierende Agenzien ( <i>Erythropoiesis-stimulating agents</i> ); IPSS-R = <i>International Prognostic Scoring System-revised</i> ; MDS = Myelodysplastische Syndrome; RCT = Randomisiert kontrollierte Studie ( <i>randomized controlled trial</i> ); s. c. = subkutan			

Die Fragestellung der vorliegenden Nutzenbewertung bezieht sich auf das gemäß Zulassung definierte Anwendungsgebiet von Luspatercept und die Selektionskriterien wurden dementsprechend gewählt.

Das **Selektionskriterium 1 (Patientenpopulation)** umfasst die relevante Patientenpopulation gemäß Zulassung von Luspatercept [36]. Die zu untersuchende **Intervention (Selektionskriterium 2)** entspricht der Dosierung und Darreichungsform gemäß Fachinformation von Luspatercept [36]. Für die Nutzenbewertung berücksichtigte Studien sollten entsprechend **Selektionskriterium 3 (Endpunkte)** mindestens einen patientenrelevanten Endpunkt aus einer der Kategorien Mortalität, Morbidität, Lebensqualität oder Sicherheit berichten (vgl. auch Abschnitt 4.2.5.2). Um eine hohe Aussagekraft der Ergebnisse und minimale Ergebnisunsicherheit sicherzustellen, wurden für die Ableitung des Zusatznutzens ausschließlich Studien im

RCT-Design (Evidenzgrad 1b) oder Meta-Analysen aus RCT herangezogen (**Selektionskriterium 4: Studiendesign**). Um relevante Behandlungseffekte mit ausreichender Sicherheit im Kontext der zugrundeliegenden chronischen Erkrankung nachweisen zu können, wurden nur Studien mit einer Studiendauer von mindestens 24 Wochen eingeschlossen (**Selektionskriterium 5: Studienlänge**). Das **Selektionskriterium 6 (Ergebnisdarstellung)** umfasst nur Studien, zu denen ausreichende Informationen verfügbar waren, um eine adäquate Bewertung und Einordnung der Ergebnisse sicherzustellen.

### 4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

#### 4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

#### 4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE (inklusive „in-process & other non-indexed citations“) und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

*Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.*

Am 18.05.2020 wurde eine bibliografische Recherche für das zu bewertende Arzneimittel Luspatercept in den Datenbanken MEDLINE, EMBASE und Cochrane (*Cochrane Central Register of Controlled Trials*) entsprechend der in Abschnitt 4.2.1 formulierten Fragestellung und den in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien durchgeführt. Die Suche wurde in jeder Datenbank einzeln und mit einer jeweils adaptierten Suchstrategie durchgeführt, welche im Anhang 4-A detailliert dokumentiert ist. In MEDLINE und EMBASE bestand der Suchblock aus einer Suche nach der Intervention, welcher anschließend mittels eines validierten Suchfilters auf RCT eingeschränkt wurde. Für die Suche in der Cochrane Datenbank wurde der Suchblock nach der Intervention auf die Kategorie „Trials“ eingeschränkt.

Die Recherchestrategie wurde weder zeitlich noch sprachlich eingeschränkt.

#### **4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken**

Eine Suche in öffentlich zugänglichen Studienregistern/ Studienergebnisdatenbanken ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern / Studienergebnisdatenbanken vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern/ Studienergebnisdatenbanken muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern sowie Studienergebnisdatenbanken immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern/ Studienergebnisdatenbanken [clinicaltrials.gov](http://www.clinicaltrials.gov) ([www.clinicaltrials.gov](http://www.clinicaltrials.gov)), EU Clinical Trials Register (EU-CTR, [www.clinicaltrialsregister.eu](http://www.clinicaltrialsregister.eu)), International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal), Suchportal der WHO, Clinical Data Suchportal der European Medicines Agency (<https://clinicaldata.ema.europa.eu>) sowie dem Arzneimittel-Informationssystem (AMIS, <http://www.pharmnet-bund.de/dynamic/de/am-info-system/index.html>) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern / Studienergebnisdatenbanken (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern/ Studienergebnisdatenbanken anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister/ Studienergebnisdatenbank einzeln und mit einer für das jeweilige Studienregister/ Studienergebnisdatenbank adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Für Clinical Data (Suchportal der European Medicines Agency) und das Arzneimittel-Informationssystem (AMIS) genügt hingegen die Suche nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

*Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern/ Studienergebnisdatenbanken die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.*

Die Suche nach Studien mit dem zu bewertenden Arzneimittel Luspatercept in Studienregistern wurde am 18.05.2020 in den Studienregistern *clinicaltrials.gov*, *EU Clinical Trials Register* (EU-CTR) und das Suchportal der *International Clinical Trials Registry Platform* (ICTRP) entsprechend der in Abschnitt 4.2.1 definierten Fragestellung und den Ein- und Ausschlusskriterien gemäß Abschnitt 4.2.2 durchgeführt. Die Suche wurde in jedem Studienregister einzeln und mit einer jeweils adaptierten Suchstrategie durchgeführt. Eine detaillierte Dokumentation der Suchen in den Studienregistern *clinicaltrials.gov*, EU-CTR und dem Suchportal ICTRP befindet sich im Anhang 4-B.

Um die Vollständigkeit der Ergebnisdarstellung der identifizierten, relevanten Studien zu überprüfen, wurde anschließend in den Studienergebnisdatenbanken der EMA (Clinical Data) und des Arzneimittelinformationssystems (AMIS) nach den aus der bibliografischen Literaturrecherche und der Recherche in Studienregistern identifizierten, relevanten Studien gezielt gesucht. Als Suchbegriffe wurden verschiedene Varianten der Intervention und Studienidentifikationsnummern verwendet.

Die Recherchestrategie wurde weder zeitlich, sprachlich noch anhand des Status der Studie eingeschränkt.

#### 4.2.3.4 Suche auf der Internetseite des G-BA

Die Internetseite des G-BA ist grundsätzlich zu durchsuchen, um sicherzustellen, dass alle vorliegenden Daten zu Studienmethodik und –ergebnissen von relevanten Studien in die Bewertung einfließen.

Auf der Internetseite des G-BA werden Dokumente zur frühen Nutzenbewertung nach §35a SGB V veröffentlicht. Diese enthalten teilweise anderweitig nicht veröffentlichte Daten zu Studienmethodik und –ergebnissen<sup>2</sup>. Solche Daten sind dabei insbesondere in den Modulen 4 der Dossiers pharmazeutischer Unternehmer, in IQWiG-Nutzenbewertungen sowie dem Beschluss des G-BA einschließlich der Tragenden Gründe und der Zusammenfassenden Dokumentation zu erwarten.

---

<sup>2</sup> Köhler M, Haag S, Biester K, Brockhaus AC, McGauran N, Grouven U, Kölsch H, Seay U, Hörn H, Moritz G, Staack K, Wieseler B. Information on new drugs at market entry: retrospective analysis of health technology assessment reports, journal publications, and registry reports. *BMJ* 2015;350:h796

Die Suche auf der Internetseite des G-BA muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche auf der G-BA Internetseite immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird. Die Suche ist dann sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie durchzuführen. Es genügt die Suche nach Einträgen zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

*Beschreiben Sie nachfolgend das Vorgehen für die Suche. Benennen Sie die Wirkstoffe und die auf der Internetseite des G-BA genannten zugehörigen Vorgangsnummern, zu denen Sie eine Suche durchgeführt haben.*

*Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.*

Die Internetseite des G-BA wurde am 18.05.2020 nach der Intervention (Luspatercept), Indikation („Myelodysplastische Syndrome“, „MDS“) sowie den Studiennamen und -nummern (MEDALIST, NCT02631070, „ACE-536-MDS-001“) durchgeführt, die im Rahmen der bibliografischen Literaturrecherche und/oder der Suche in Studienregistern und Studienergebnisdatenbanken als relevant identifiziert wurden.

#### **4.2.3.5 Selektion relevanter Studien**

*Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2, 4.2.3.3 und 4.2.3.4 beschriebenen Rechenschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.*

Die Selektion relevanter Studien aus den Ergebnissen der bibliografischen Literaturrecherche sowie der Suche in Studienregistern wurde anhand der folgenden Schritte nacheinander von zwei Personen unabhängig voneinander durchgeführt. Sofern sich bezüglich des Ein- und Ausschlusses von Treffern abweichende Beurteilungen ergaben, wurden diese im abschließenden Konsensverfahren durch die beiden am Verfahren beteiligten Personen geklärt.

Als relevant identifizierte Studien wurden durch gezielte Suche in den Studienergebnisdatenbanken (EMA und AMIS) sowie auf der Internetseite des G-BA auf Vollständigkeit hinsichtlich der Verfügbarkeit relevanter Informationen und Studienergebnisse überprüft.

### **Bibliografische Literaturrecherche**

Die Selektion relevanter Studien aus den Ergebnissen der bibliografischen Literaturrecherche erfolgte in einem vierstufigen Prozess:

1. Entfernung von Duplikaten
2. Ausschluss von Referenzen, bei denen klar erkennbar war, dass es sich um keine Berichterstattungen handelt, welche ausreichende Informationen zur Beurteilung der Methodik und Ergebnissen liefern
3. Ausschluss nicht relevanter Ergebnisse durch Titel- und Abstract-Screening anhand der in Abschnitt 4.2.2 definierten Ein- und Ausschlusskriterien; Ergebnisse, die nicht eindeutig anhand des Titels und/oder Abstracts ausgeschlossen werden konnten, wurden im 4. Schritt überprüft; ausgeschlossene Meta-Analysen wurden zusätzlich auf die Darstellung relevanter Studien geprüft
4. Ausschluss nicht relevanter Ergebnisse nach Volltext-Screening

### **Suche in Studienregistern**

Zur Selektion relevanter Studien aus den Ergebnissen der Suche in Studienregistern wurden in einem ersten Schritt solche Studienregistereinträge ausgeschlossen, die den in Abschnitt 4.2.2 festgelegten Einschlusskriterien nicht entsprachen. Grundlage dieser Einschätzung waren alle in den Studienregistern hinterlegten Informationen. Die Selektion relevanter Studien erfolgte für jedes Studienregister einzeln.

#### **4.2.4 Bewertung der Aussagekraft der Nachweise**

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie der behandelnden Personen
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

*Beschreiben Sie die für die Bewertung der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

Die eingeschlossenen, relevanten Studien wurden hinsichtlich ihrer Aussagekraft anhand der zur Verfügung stehenden Informationen (Studienberichte und Publikationen) bewertet.

### **Bewertung des Verzerrungspotentials**

Die Bewertung des Verzerrungspotentials erfolgte sowohl endpunktübergreifend als auch endpunktspezifisch. In einem ersten Schritt wurde das Verzerrungspotential der RCT endpunktübergreifend anhand der folgenden Komponenten bewertet:

#### ***Verzerrungsaspekte der Ergebnisse auf Studienebene:***

- Erzeugung der Randomisierungssequenz
- Verdeckung der Gruppenzuteilung
- Verblindung des Patienten sowie der behandelnden Personen
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Zusätzliche Bewertung des Verzerrungspotentials der RCT mit Einstufung als „niedrig“ oder „hoch“.

Bei einem hohen Verzerrungspotential auf Studienebene ist davon auszugehen, dass auch die Endpunkte hoch verzerrt sind, sodass auch hier eine Einstufung des Verzerrungspotentials als „hoch“ vorgenommen wurde. In Ausnahmen kann das Verzerrungspotential auf Endpunktebene aber auch trotz hoher Verzerrung auf Studienebene „niedrig“ sein; solche Fälle bedurften einer entsprechenden Begründung.

Bei einem niedrigen Verzerrungspotential auf Studienebene wurde für jeden Endpunkt einzeln das Verzerrungspotential auf Endpunktebene gemäß den folgenden Aspekten geprüft:

#### ***Verzerrungsaspekte der Ergebnisse auf Endpunktebene***

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Die Einstufungen der Verzerrungspotentiale auf Studien- und Endpunktebene dienen der Diskussion heterogener Studienergebnisse und der Bewertung von deren Aussagesicherheit. Ein Ausschluss von Ergebnissen ausschließlich aufgrund eines hohen Verzerrungspotentials fand nicht statt.

Eine detaillierte Bewertung der Verzerrungspotentiale ist im Anhang 4-F dargestellt.

## 4.2.5 Informationssynthese und -analyse

### 4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)<sup>3</sup>. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND<sup>4</sup> bzw. STROBE-Statements<sup>5</sup> folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

*Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.*

Die Darstellung des Designs und der Methodik der eingeschlossenen Studien erfolgte anhand der Empfehlungen des CONSORT-Statements 2010 (Items 2b bis 14), welches Angaben zur Methodik (Studienhypothese, Teilnehmerselektion, Intervention und Behandlung im Placebo-Arm, Endpunkte, Randomisierung, Verblindung und statistische Annahmen) und den Ergebnissen (Patientenfluss, Rekrutierung) der Studie enthält. Eine detaillierte Beschreibung inklusive eines Flow-Charts, die den Patientenfluss veranschaulicht, befindet sich im Anhang 4-E.

### 4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 4.3.2 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

<sup>3</sup> Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c332.

<sup>4</sup> Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Publ Health* 2004; 94(3): 361-366.

<sup>5</sup> Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.

*Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.*

*Sofern zur Berechnung von Ergebnissen von Standardverfahren und –software abgewichen wird (insbesondere beim Einsatz spezieller Software oder individueller Programmierung), sind die Berechnungsschritte und ggf. verwendete Software explizit abzubilden. Insbesondere der Programmcode ist in lesbarer Form anzugeben.*

#### **4.2.5.2.1 Patientencharakteristika**

In der MEDALIST-Studie wurden erwachsene Patienten mit transfusionsabhängiger Anämie aufgrund von MDS mit Ringsideroblasten, mit sehr niedrigem, niedrigem oder intermediärem Risiko, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, eingeschlossen. Eingeschlossene Patienten wurden anhand der folgenden Patientencharakteristika zu Studienbeginn charakterisiert (siehe Tabelle 4-14):

- Demographische Charakteristika:
  - Alter  
( $\leq 64$  Jahre, 65 – 74 Jahre,  $\geq 75$  Jahre)
  - Geschlecht  
(männlich, weiblich)
  - Ethnische Zugehörigkeit  
(schwarz oder afroamerikanisch, weiß, nicht erfasst, andere)
  - Gewicht  
( $< 70$  kg, 70 –  $< 85$  kg, 85 –  $< 100$  kg,  $\geq 100$  kg)
  - Größe [cm]
  - Body Mass Index (BMI)  
( $< 20$  kg/m<sup>2</sup>, 20 –  $< 25$  kg/m<sup>2</sup>, 25 –  $< 30$  kg/m<sup>2</sup>,  $\geq 30$  kg/m<sup>2</sup>)
  
- Erkrankungscharakteristika zur Baseline:
  - Zeit seit ursprünglicher MDS-Diagnose  
( $\leq 2$  Jahre,  $> 2$  – 5 Jahre,  $> 5$  Jahre)
  - Ringsideroblasten-Status  
( $\geq 15$  % der Gesamtzahl erythroider Zellen im KM)
  - WHO-Klassifikation der MDS  
(MDS RARS, MDS RCMD, andere)
  - IPSS-R-Klassifizierung  
(sehr niedrig, niedrig, intermediär, hoch)

- Serum-EPO  
( $< 100$  U/l,  $100 - < 200$  U/l,  $200 - 500$  U/l,  $> 500$  U/l)
- Mutation der *Splicing factor 3B subunit 1* (SF3B1)  
(mutiert, nicht mutiert)
- *Eastern Cooperative Oncology Group* (ECOG) Performance-Status  
(0, 1, 2)
  
- Transfusionslast zur Baseline:
  - Durchschnittlicher Bedarf an EK-Transfusionen [EK-Einheiten pro acht Wochen]
  - Transfusionslast in den letzten acht bzw. 16 Wochen vor der ersten Dosis Studienmedikation  
(mindestens vier bis weniger als sechs EK-Einheiten pro acht Wochen, weniger als vier EK-Einheiten pro acht Wochen, weniger als sechs EK-Einheiten pro acht Wochen, mindestens sechs EK-Einheiten pro acht Wochen)
  - Hb-Wert [g/dl]
  
- Behandlungsanamnese der MDS:
  - Vorherige ESA-Therapie  
(ja, nein)
  - Vorherige Therapie mit Granulozyten-Kolonie-stimulierendem Faktor (*granulocyte colony-stimulating factor*, G-CSF)/Granulozyten-Makrophagen-Kolonie-stimulierendem Faktor (*granulocyte-macrophage colony-stimulating factor*, GM-CSF)  
(ja, nein)
  - Vorherige Therapie mit Eisenchelatoren  
(ja, nein)
  
- Laborwert-Charakteristika zur Baseline:
  - sFerritin-Spiegel [ $\mu\text{g/l}$ ]
  - Nierenfunktion (gemessen als Kreatinin-Clearance)  
( $< 40$  ml/min,  $40 - < 60$  ml/min,  $\geq 60$  ml/min)
  - Thrombozytenzahl [ $10^9/\text{l}$ ]
  - Absolute Neutrophilenzahl (*absolute neutrophil count*, ANC)  
( $< 0,5 \times 10^9/\text{l}$ ,  $0,5 - < 1,0 \times 10^9/\text{l}$ ,  $\geq 1,0 \times 10^9/\text{l}$ )
  
- Komorbiditäten:
  - Kardiovaskuläre Komorbiditäten (dargestellt als *Major Adverse Cardiac Events* (MACE)-Ereignisse)  
(ja, nein)
  - Komorbiditätsscore (ähnlich dem *Charlson Comorbidity Index*)

#### 4.2.5.2.2 Patientenrelevante Endpunkte

Gemäß Definition des IQWiG wird als patientenrelevant verstanden, „wie ein Patient fühlt, seine Funktionen und Aktivitäten wahrnehmen kann oder ob er überlebt“ [35].

Im vorliegenden Anwendungsgebiet sind Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS umfasst. Die Anämie (Mangel an funktionstüchtigen Erythrozyten) infolge eines Erythrozyten-Reifungs-Defekts ist mit 70 – 80 % die häufigste Erstmanifestation bei Patienten mit Niedrigrisiko-MDS. Bei MDS-Patienten mit Ringsideroblasten ist die Anämie und damit einhergehende Transfusionsabhängigkeit besonders ausgeprägt (vgl. auch Modul 3 B). Im Gegensatz zu fortgeschrittenen Stadien (Hochrisiko-MDS) ist die Überlebensprognose bei Patienten mit Niedrigrisiko-MDS vergleichsweise gut (medianes Überleben: 3,0 – 8,8 Jahre) und auch das Risiko einer Progression in eine AML ist relativ gering (mediane Zeit bis 25 % der Patienten in eine AML übergehen:  $\geq 3,0$  Jahre) [37]. Daher ist bei Patienten mit Niedrigrisiko-MDS das primäre Therapieziel die Behandlung der Zytopenie(n) und somit Verbesserung des klinischen Zustands, um die damit assoziierte Morbidität (z. B. Dyspnoe, Fatigue, Tachykardie und Kopfschmerzen aufgrund einer Anämie, Infektionen aufgrund einer Neutropenie und Blutungsstörungen aufgrund einer Thrombozytopenie) und Beeinträchtigung der Lebensqualität zu minimieren und die Autonomie der Patienten zu erhalten [2, 3, 38].

Zentraler Bestandteil der Therapie einer Anämie ist die supportive Behandlung mit EK-Transfusionen, welche in Abhängigkeit des klinischen Zustands (dynamisch bestimmt durch Symptome, Alter, Erkrankungsstadium, individuelle Patientenbiologie, Komorbiditäten) des Patienten verabreicht werden [2] – für Patienten im Anwendungsgebiet von Luspatercept ist dies sogar die einzige zugelassene Behandlungsoption. Die Transfusionstherapie ist zwar notwendig, um die Anämie-Symptome temporär zu vermeiden, birgt aber das Risiko teils schwerwiegender Folgekomplikationen und vermag keine stabile Verbesserung der Anämie zu erreichen. Aufgrund dieser Limitationen geht die Transfusionstherapie mit erheblichen psychosozialen und zeitlichen Belastungen für den Patienten einher und stellt nur eine unzureichende therapeutische Maßnahme dar:

- Die Transfusionsabhängigkeit ist ein negativer Prognosefaktor für das Gesamtüberleben [8, 39-43].
- Jede EK-Transfusion kann Nebenwirkungen und Folgekomplikationen verursachen, die das Befinden und den Gesundheitsstatus des Patienten weiter einschränken [44]. Das Risiko für Nebenwirkungen und Komplikationen ist bei MDS-Patienten aufgrund der regelmäßigen Exposition sowie aufgrund von Komorbiditäten, fortgeschrittenen Alters, eingeschränkter Immunfunktion und Neutropenie besonders hoch [8].
- Dauerhafte EK-Transfusionen können zu einer sekundären Eisenüberladung (sekundäre Hämosiderose) führen, die Organschädigungen, Organversagen und sogar zum Tod führen kann [3, 4, 9]. Eine Eisenüberladung geht zudem mit einer Verschlechterung der Überlebensprognose einher [27, 41, 43].

- Die Gefahr einer transfusionsbedingten sekundären Hämosiderose macht ab einer Transfusionslast von  $\geq 20$  EK-Einheiten bzw. einem sFerritin-Spiegel  $\geq 1.000$  ng/ml wiederum die zusätzliche Therapie mit Eisenchelatoren notwendig. Die Anwendung von Eisenchelatoren ist jedoch auch selbst mit dem Risiko von zum Teil schwerwiegenden Nebenwirkungen verbunden und kann die Lebensqualität der Patienten negativ beeinflussen [9, 45, 46]. Insbesondere das Auftreten von Nierenschäden (Deferoxamin, Deferasirox), Hepatopathien (Deferasirox) sowie Okulo- und Ototoxizität (Deferoxamin, Deferasirox) gestalten den Einsatz nicht ganz unproblematisch [3, 9, 47-49].
- Eine Transfusionstherapie geht mit großen zeitlichen aber auch psychosozialen Belastungen einher, die sich auch in Erhebungen zur Lebensqualität von MDS-Patienten widerspiegelt [16, 17, 32]. Ein großes Unbehagen gegenüber Fremdblut sowie die ständige Angst vor potentiellen Risiken spielen hier mit hinein.
- Die Transfusionstherapie mit EK vermag keine anhaltende, stabile Korrektur des Hb-Spiegels und somit andauernde Vermeidung der Anämie-Symptomatik zu erreichen. Nach Verabreichung der EK-Transfusion fühlen sich die Patienten aufgrund hoher Hb-Werte initial zwar gut, jedoch fallen die Hb-Werte schnell wieder ab, sodass Anämie-Symptome wieder auftreten können. Das regelmäßige Absinken des Hb-Werts ist mit dem Risiko von Anämie-Symptomen (z. B. Fatigue) und Anämie-bedingten Schäden (bspw. Verschlechterung von Komorbiditäten und des Gesundheitszustands) verbunden. Diese Schwankungen des Hb-Werts sind für den Patienten in entsprechenden Energieniveaus spürbar, welche sich negativ auf den Allgemeinzustand und die Lebensqualität auswirken [5, 12].

Unter Berücksichtigung der limitierten Behandlungsoptionen und den teils schwerwiegenden Nachteilen einer supportiven Transfusionstherapie besteht somit ein großer medizinischer Bedarf an wirksamen Therapien, die die folgenden Therapieziele erreichen:

- **Behandlung der Anämie, d. h. Verbesserung der Erythropoese:** Durch den Erythrozyten-Reifungs-Defekt leiden MDS-Patienten an einer chronischen Anämie, die zudem die Transfusionsabhängigkeit bedingt. Eine Korrektur des Erythrozyten-Reifungs-Defekts kann die Anämie behandeln, den klinischen Zustand des Patienten verbessern und den Transfusionsbedarf senken. Die Verbesserung der Erythropoese und Anämie spiegelt sich in einer stabilen Erhöhung des Hb-Werts wider.
- **Vermeidung von EK-Transfusionen und transfusionsbedingten Komplikationen:** Die Indikation für eine EK-Transfusion erfolgt in Abhängigkeit des klinischen Zustands des Patienten. Eine Transfusionsfreiheit entspricht somit dem Fehlen von Anämie-Symptomen, die Transfusionen notwendig machen würden. Zudem werden durch die Vermeidung von EK-Transfusionen gleichzeitig auch transfusionsassoziierte akute und langfristige Nebenwirkungen und Folgekomplikationen vermieden (bspw. allergische Reaktionen, sekundäre Eisenüberladung und Siderose-bedingte Morbidität, Ausbildung von Alloantikörpern und hämolytische Transfusionsreaktionen). Auch wenn der Einsatz von Eisenchelatoren das Risiko von Siderose-bedingten Schäden deutlich senken kann [50, 51], ist der Eisenhaushalt bei einigen Patienten nur unzureichend regulierbar, eine Vermeidung der ursächlichen

Transfusionen wäre zielführender. Das Erreichen einer Transfusionsfreiheit stellt somit das Hauptziel der Behandlung der symptomatischen Anämie bei Patienten mit Niedrigrisiko-MDS dar. Unabhängig davon profitieren Patienten bereits von der Reduktion der Transfusionslast und -häufigkeit, da dadurch die direkten und verzögert einsetzenden Folgen der Transfusionstherapie ebenfalls gemildert und reduziert werden [38].

- **Aufrechterhaltung der Lebensqualität und Autonomie der Patienten:** Die Grunderkrankung an sich, darunter Symptome der Anämie als auch psychische Belastungen wie die Unsicherheit des Krankheitsverlaufs, kann sich negativ auf das Wohlbefinden der Patienten auswirken. Hinzu kommen zeitliche und psychosoziale Belastungen der Transfusionstherapie, die zu einer weiteren Beeinträchtigung und Verringerung der Autonomie führen können. Ziel ist es daher, die Lebensqualität und Autonomie der Patienten aufrecht zu erhalten [2]

Die in der Studie MEDALIST erhobenen patientenrelevanten Endpunkte sind im Nachfolgenden beschrieben und gliedern sich entsprechend der Endpunktkategorien Mortalität, Morbidität, Lebensqualität und Sicherheit (siehe Tabelle 4-3).

Tabelle 4-3: Für die Nutzenbewertung herangezogene patientenrelevante Endpunkte (Hauptanalyse)

Endpunktkategorie	Endpunkt
Mortalität	Gesamtüberleben Zeit bis zum Tod jeglicher Ursache
Morbidität	Transfusionsfreiheit Anteil Patienten mit einer Transfusionsfreiheit über $\geq 8$ , $\geq 12$ , $\geq 16$ bzw. $\geq 24$ Wochen
	Reduktion der Transfusionslast Anteil der Patienten mit einem HI-E
	Symptomatik Anteil der Patienten mit einer Verbesserung der Symptomatik um $\geq 10$ Punkte gemessen anhand des EORTC QLQ-C30-Fragebogens
	Hospitalisierungen Anteil Patienten mit mindestens einer Hospitalisierung
Lebensqualität	Gesundheitsbezogene Lebensqualität gemessen anhand des generischen Fragebogens EORTC QLQ-C30 Anteil der Patienten mit einer Verbesserung der gesundheitsbezogenen Lebensqualität um $\geq 10$ Punkte gemessen anhand des EORTC QLQ-C30-Fragebogens
	Krankheitsspezifische Lebensqualität gemessen anhand des Fragebogens QoL-E Mittlere Änderung der krankheitsspezifischen Lebensqualität gemessen anhand des QoL-E-Fragebogens im Vergleich zur Baseline

Endpunktkategorie	Endpunkt
Sicherheit	Gesamtraten UE Anteil Patienten mit mindestens einem UE (einschließlich Aufteilung nach SOC und PT) Anteil Patienten mit mindestens einem SUE (einschließlich Aufteilung nach SOC und PT) Anteil Patienten mit mindestens einem schweren UE (CTCAE $\geq$ 3) (einschließlich Aufteilung nach SOC und PT) Anteil Patienten mit einem Therapieabbruch aufgrund von UE (einschließlich Aufteilung nach SOC und PT)
<small>CTCAE = Common Terminology Criteria for Adverse Events; EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module; HI-E = erythroides Ansprechen (hematologic improvement-erythroid); PT = bevorzugter Begriff (preferred term); QoL-E = Quality of Life-E; SOC = Systemorgan-klasse (system organ class); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis</small>	

## Mortalität

### *Gesamtüberleben*

#### *Patientenrelevanz:*

Das Gesamtüberleben ist allgemein als patientenrelevanter Endpunkt akzeptiert, die Patientenrelevanz wurde zudem vom G-BA im Beratungsgespräch zu Luspatercept bestätigt [52].

#### *Validität und Operationalisierung:*

Das Gesamtüberleben stellt einen harten, etablierten Endpunkt dar, der als valide einzustufen ist.

#### Hauptanalyse:

Operationalisiert wurde das Gesamtüberleben als Zeit zwischen der Randomisierung und dem Todeszeitpunkt, bzw. dem Datum der Zensurierung für Patienten, die zum Zeitpunkt der Analyse noch lebten, die Studie abgebrochen hatten oder während der Verlaufsbeobachtung verloren gegangen sind.

#### *Darstellung im Dossier:*

- **Hauptanalyse:** Gesamtüberleben (Zeit bis zum Tod jeglicher Ursache); Woche eins bis zum Datenschnitt vom 01.07.2019

## Morbidität

### *Transfusionsfreiheit*

#### *Patientenrelevanz:*

Das Erreichen einer Transfusionsfreiheit spiegelt die effektive Verbesserung der Erythropoese und Vermeidung von Anämie wider. Durch die Korrektur des Erythrozyten-Reifungs-Defekts kann die chronische Anämie bei MDS gezielt behandelt und somit auch die typischen Anämie-Symptome gemildert oder eliminiert werden. EK-Transfusionen sind bei vollständiger Transfusionsfreiheit nicht mehr notwendig. In diesem Zustand werden also sowohl die Symptome der Anämie als auch die Nebenwirkungen und Komplikationen der Transfusionstherapie vermieden. Daher ist das Erreichen einer Transfusionsfreiheit eines der zentralen, patientenrelevanten Ziele der Therapie bei Niedrigrisiko-MDS und wird auch für die Bemessung des Therapieerfolgs in klinischen Studien empfohlen [3, 38, 53]. Der G-BA und das IQWiG haben den Endpunkt Transfusionsfreiheit in vorangegangenen Verfahren bereits als patientenrelevant anerkannt [21-23].

Auch die BÄK weist sowohl in ihrer Querschnitts-Leitlinie zur Therapie mit Blutkomponenten und Plasmaderivaten als auch in der Richtlinie Hämotherapie auf die Relevanz eines restriktiven Einsatzes von EK-Transfusionen hin. Wenn möglich, sind kausale oder andere, gleichwertige Therapien einer Transfusionstherapie immer vorzuziehen. Die Eignung anderer Maßnahmen ist vor der Substitutionsbehandlung mit EK zu prüfen [28, 29]. Diese Handlungsanweisungen betonen die äußerste Relevanz der Vermeidung und Verminderung von EK-Transfusionen.

Gemäß Angaben der internationalen Arbeitsgruppe klinischer Experten der MDS (IWG) aus dem Jahr 2006 wird eine Transfusionsfreiheit über mindestens acht Wochen als klinisch relevanter Therapieerfolg angesehen, welcher als primärer Endpunkt in der Studie MEDALIST umgesetzt wurde [38].

#### *Validität und Operationalisierung:*

Das Erreichen einer Transfusionsfreiheit ist ein etablierter Endpunkt im Anwendungsgebiet der Niedrigrisiko-MDS, der seit der ersten Veröffentlichung der IWG zur Bemessung klinisch relevanter Therapieerfolge fester Bestandteil dieser Ansprechkriterien ist [38, 53, 54]. Auch im Rahmen von Zulassungsverfahren wurde dieser Endpunkt seitens der EMA bereits zur Bewertung des medizinischen Nutzens herangezogen [55]. Der Endpunkt Transfusionsfreiheit wird somit als valide und zur Ableitung eines Zusatznutzens geeignet eingestuft.

Hauptanalyse:

Auswertungen zur Transfusionsfreiheit beziehen sich entsprechend der zur Studienplanung und -durchführung geltenden Fassung der IWG-Kriterien aus dem Jahr 2006 auf einen transfusionsfreien Zeitraum von mindestens acht Wochen. Zusätzlich wird das Erreichen einer Transfusionsfreiheit über mindestens 12 Wochen, entsprechend der Präspezifikation im Studienbericht, und mindestens 16 (entsprechend der überarbeiteten IWG-Kriterien aus dem Jahr 2018) und 24 Wochen (entsprechend einer vollständigen Transfusionsfreiheit im betrachteten primären Analysezeitraum) zur Bewertung des langfristigen Therapieeffekts dargestellt.

Zusatzanalysen:

Auswertungen über den Studienzeitraum bis Woche 48 bzw. bis zum Datenschnitt vom 01.07.2019 werden aufgrund der eingeschränkten Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalysen (Zeit bis zum Erreichen einer Transfusionsfreiheit) dargestellt. Die Zeit bis zum Erreichen einer Transfusionsfreiheit war definiert als Zeitraum zwischen dem Behandlungsbeginn und dem Beginn des ersten transfusionsfreien Zeitraums über mindestens 8, 12, 16 bzw. 24 Wochen. Patienten, die nach 48 Wochen keine Transfusionsfreiheit über mindestens 8, 12, 16 bzw. 24 Wochen erreicht haben, wurden zensiert.

Analysen zur **längsten Dauer eines transfusionsfreien Zeitraums** werden ergänzend dargestellt. In diese Analyse fließen die Transfusionsereignisse der Patienten ohne Einschränkung der Mindestlänge einer Transfusionsfreiheit ein. Die längste Dauer eines transfusionsfreien Intervalls war definiert als längster zusammenhängender Zeitraum zwischen zwei aufeinanderfolgenden Transfusionszeitpunkten, in dem keine EK-Einheit transfundiert wurde.

**Swimmer-Plots** der Transfusionszeitpunkte und Intervalle mit einer Transfusionsfreiheit von mindestens acht Wochen werden ergänzend berichtet, um die Länge transfusionsfreier Intervalle, die Häufigkeiten eines Ansprechens und die Transfusionszeitpunkte zu visualisieren.

Zusätzlich wurde der **Hb-Wert** als klinisch relevanter Parameter im Kontext der Transfusions-therapie untersucht. Der Hb-Wert dient der Bewertung der anhaltenden Regulation der Erythro-poese und somit Korrektur des der Anämie zugrundeliegenden Mechanismus. Der Hb-Wert korreliert mit dem klinischen Zustand des Patienten, welcher für die Indikation von Transfu-sionen entscheidend ist [2, 5]. Erst eine stabile Verbesserung des Hb-Werts ermöglicht das Erreichen einer Transfusionsfreiheit. Das Absinken des Hb-Werts würde erneute Transfusionen notwendig machen. Somit demonstriert die stabile Verbesserung des Hb-Wertes auch, dass eine Transfusionsfreiheit nicht auf Kosten einer Verschlechterung der Erythropoese und Anämie, gemessen als Verringerung des Hb-Werts, erreicht wird. Die Änderung des Hb-Werts wurde operationalisiert als mittlere Änderung im Vergleich zur Baseline. Der Baseline-Hb-Wert war definiert als niedrigster Wert innerhalb der 35 Tage vor Behandlungsbeginn nach Anwendung der 14/3-Tage-Regel. Diese Regel besagt, dass nur solche Hb-Werte berücksichtigt werden, die mindestens 14 Tage nach einer EK-Transfusion gemessen werden, es sei denn, dass eine weitere Transfusion innerhalb von drei Tagen nach der Hb-Wert-Messung stattfindet. In diesem Fall wird der Hb-Wert herangezogen, der innerhalb von 14 Tagen nach der ersten Transfusion gemessen wird, da davon auszugehen ist, dass diese erste Transfusion nur einen geringen Einfluss auf den Hb-Wert hatte.

*Darstellung im Dossier:*

- **Hauptanalyse:** Anteil der Patienten mit einer Transfusionsfreiheit; Woche 1 – 24
- **Zusatzanalyse 1:** Zeit bis zum Erreichen einer Transfusionsfreiheit; Woche eins bis zum Datenschnitt vom 01.07.2019
- **Zusatzanalyse 2:** Längste Dauer eines transfusionsfreien Zeitraums; Woche eins bis zum Datenschnitt vom 01.07.2019
- **Zusatzanalyse 3:** Swimmer-Plots Transfusionszeitpunkte und Intervalle mit einer Trans-fusionsfreiheit über mindestens acht Wochen; Woche eins bis zum Datenschnitt vom 01.07.2019
- **Zusatzanalyse 4:** Mittlere Änderung des Hb-Werts im Vergleich zur Baseline; Woche 1 – 24 und Woche 1 – 48

### ***Reduktion der Transfusionslast***

#### *Patientenrelevanz:*

Wie im vorherigen Abschnitt erläutert, ist das bedeutendste Therapieziel der Behandlung symptomatischer Anämien bei Patienten mit Niedrigrisiko-MDS das Erreichen einer Transfusionsfreiheit. Unabhängig davon profitieren Patienten bereits von der Reduktion der Transfusionslast, da dadurch die direkten und verzögert einsetzenden Belastungen und Folgen der Transfusionstherapie gemildert werden [2, 38]. So konnte gezeigt werden, dass bereits eine Reduktion der Transfusionslast als eine deutliche Verbesserung des Wohlbefindens wahrgenommen wird [17]. Darüber hinaus entspricht eine Reduktion der Transfusionslast einer proportionalen Reduktion der Risiken, die durch Transfusionsreaktionen und sekundäre Eisenüberladung entstehen [30, 41]. Diese sind insbesondere vor dem chronischen Hintergrund der Erkrankung langfristig von großer Bedeutung.

Somit stellt auch eine Reduktion der Transfusionslast einen patientenrelevanten Endpunkt dar.

#### *Validität und Operationalisierung:*

Die Reduktion der Transfusionslast bis hin zur Transfusionsfreiheit und eine Verbesserung der Hb-Werte wurde im Rahmen der IWG-Kriterien zur Bewertung von klinisch relevanten Therapieerfolgen als erythroides Ansprechen (HI-E) etabliert. Insbesondere bei Patienten mit hoher Transfusionslast ist allein die Reduktion der notwendigen EK-Einheiten bereits als klinisch relevanter Endpunkt anzusehen [38, 53, 54]. Das HI-E gilt im Anwendungsgebiet der Niedrigrisiko-MDS als valider, patientenrelevanter Endpunkt [4, 38].

Gemäß der zum Zeitpunkt der Studienplanung und -durchführung aktuellen IWG-Kriterien aus dem Jahr 2006 ist eine relevante Reduktion der Transfusionslast als hämatologische Verbesserung im Sinne eines HI-E wie folgt definiert:

- Für Patienten mit einer Transfusionslast von mindestens vier EK-Einheiten pro acht Wochen vor Behandlungsbeginn: Verringerung der Anzahl transfundierter EK-Einheiten um mindestens vier EK-Einheiten pro acht Wochen im Vergleich zu dem Acht-Wochen-Zeitraum vor Behandlungsbeginn (nur Transfusionen, die bei einem Hb-Spiegel von  $\leq 9,0$  g/dl verabreicht wurden, werden berücksichtigt)

oder

- für Patienten mit einer Transfusionslast von weniger als vier EK-Einheiten pro acht Wochen vor Behandlungsbeginn: mittlere Steigerung des Hb-Spiegels um  $\geq 1,5$  g/dl über mindestens acht Wochen (bei einem Wert von  $< 11$  g/dl vor Behandlungsbeginn) (in MEDALIST wurde eine Erhöhung des Hb-Werts nur bei gleichzeitiger Transfusionsfreiheit als Ansprechen gewertet) [38].

Hauptanalyse:

Im Nutzendossier dargestellt werden Auswertungen zu den Anteilen an Patienten, die während der ersten 24 Wochen der Behandlungsphase ein HI-E gemäß den hier genannten IWG-Kriterien aus dem Jahr 2006 erreichen. Separate Auswertungen zum jeweiligen HI-E-Kriterium in Abhängigkeit der Transfusionslast zur Baseline werden ergänzend dargestellt.

Zusatzanalysen:

Auswertungen über den Studienzeitraum bis Woche 48 bzw. bis zum Datenschnitt vom 01.07.2019 werden aufgrund der eingeschränkten Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalysen (Zeit bis zum Erreichen eines HI-E) dargestellt. Die Zeit bis zum Erreichen eines HI-E war definiert als Zeitraum zwischen Behandlungsbeginn und dem Beginn des ersten Zeitraums über mindestens acht Wochen mit einem HI-E. Patienten, die zum Zeitpunkt des Datenschnitts kein HI-E über mindestens acht Wochen erreicht haben, wurden zu diesem Zeitpunkt zensiert.

Die **mittlere Änderung der Transfusionslast**, definiert als durchschnittliche Änderung der transfundierten EK-Einheiten in jeglichem 16-Wochen-Intervall im Vergleich zur Baseline, wird ergänzend zu den Auswertungen des HI-E abgebildet. Die Baseline-Transfusionslast bezieht sich auf die Transfusionshistorie der 16 Wochen vor Behandlungsbeginn.

Als weiteres Maß zur Bewertung der Entlastung des Patienten durch reduzierte EK-Transfusionen werden die Ergebnisse zur **mittleren Änderung der Transfusionshäufigkeit** berichtet. Der zentrale Unterschied zwischen Transfusionslast und Transfusionshäufigkeit besteht darin, dass die Transfusionslast die Belastung durch die absolute Menge der EK-Einheiten in einem bestimmten Zeitraum betrachtet, während durch die Transfusionshäufigkeit vornehmlich die zeitliche Belastung durch die Anzahl der Transfusionszeitpunkte abgebildet wird. Die mittlere Änderung der Transfusionshäufigkeit ist definiert als Anzahl an Transfusionszeitpunkten pro 16-Wochen-Intervall im Vergleich zur Baseline. Die Baseline-Transfusionshäufigkeit bezieht sich auf die Transfusionshistorie der 16 Wochen vor Behandlungsbeginn.

*Darstellung im Dossier:*

- **Hauptanalyse:** Anteil der Patienten mit einem HI-E; Woche 1 – 24
- **Zusatzanalyse 1:** Zeit bis zum Erreichen eines HI-E; Woche eins bis zum Datenschnitt vom 01.07.2019
- **Zusatzanalyse 2:** Mittlere Änderung der Transfusionslast im Vergleich zur Baseline; Woche 1 – 24 und Woche 1 – 48
- **Zusatzanalyse 3:** Mittlere Änderung der Transfusionshäufigkeit im Vergleich zur Baseline; Woche 1 – 24 und Woche 1 – 48

### ***Reduktion der Eisenüberladung***

Ergebnisse zu den Endpunkten Reduktion der Eisenüberladung gemessen als sFerritin-Spiegel und Bedarf an Eisenchelatoren werden zusätzlich dargestellt. Die Reduktion der Eisenüberladung stellt einen in der vorliegenden Indikation klinisch relevanten Endpunkt dar und wird daher zum besseren Verständnis der Wirksamkeit und Sicherheit von Luspatercept berichtet, jedoch nicht zur Ableitung des Zusatznutzens herangezogen.

#### *Klinische Relevanz:*

Die durch die regelmäßigen EK-Transfusionen bedingte erhöhte Eisenaufnahme führt unbehandelt zu Eisenablagerungen in verschiedenen Organen wie Herz, Leber und endokrinen Organen. Dadurch kommt es im Laufe der Zeit zu schweren Organschäden bis hin zum Tod (Siderose-bedingte Morbidität und Mortalität) [3, 4, 30, 41]. Um einer Eisenüberladung adäquat entgegen zu wirken, werden transfusionsabhängige MDS-Patienten in Abhängigkeit der Eisenlast (entsprechend des Ausmaßes und der Dauer der Transfusionstherapie ( $\geq 20$  EK-Einheiten) bzw. dem sFerritin-Spiegel ( $\geq 1.000$  ng/ml)) begleitend zur Transfusionstherapie mit einer Eisenchelattherapie behandelt [2]. In der klinischen Studie MEDALIST wurde eine adäquate Eisenchelattherapie im Rahmen der BSC sichergestellt. Symptome einer Eisenüberladung waren entsprechend nicht zu erwarten, sodass die Auswertung dieser Symptome nicht sinnvoll ist. Stattdessen wird eine Bewertung anhand des klinisch relevanten Parameters sFerritin vorgenommen.

Der sFerritin-Spiegel dient als klinisch relevanter Parameter zur Verlaufsbeobachtung der Eisenüberladung und zur Identifizierung von Trends hinsichtlich der Eisenlast im Körper [2]. Im Gegensatz zur Symptomatik der Eisenüberladung ist mittels des sFerritin-Spiegels eine Bewertung des Eisenhaushalts auch schon vor dem Vorliegen von Organschädigungen und Symptomen möglich. Eine adäquate Kontrolle des Eisenhaushalts geht zudem mit einer verbesserten Morbidität und Mortalität einher [41, 56]. Die ergänzend berichteten Auswertungen zum Bedarf an Eisenchelatoren spiegeln zudem den Einfluss der Eisenbelastung auf die Notwendigkeit weiterer, mit zusätzlichen Nebenwirkungen und Komplikationen verbundenen Behandlungen wider. Da Dosisanpassungen der Eisenchelatoren erst verzögert in Abhängigkeit der Eisenparameter erfolgt, ist im Rahmen einer klinischen Studie maximal eine Bewertung von Trends möglich [48].

#### *Operationalisierung:*

Die Reduktion der Eisenüberladung wird anhand des klinisch relevanten Parameters **sFerritin** bewertet. Da der sFerritin-Spiegel täglichen Schwankungen durch Infektionen oder entzündlichen Prozessen unterliegt, wurde die Änderung des sFerritin-Spiegels als mittlere Änderung in jeglichem 16-Wochen-Intervall im Vergleich zu den 16 Wochen vor Behandlungsbeginn (= Baseline-Wert) operationalisiert.

Als weitere Auswertung wird der Bedarf an Eisenchelatoren als **mittlere Änderung der täglichen Dosis Eisenchelator** in den 16-Wochen-Intervallen Woche 9 – 24 bzw. 33 – 48 im Vergleich zur Baseline (mittlere tägliche Dosis Eisenchelator in den 16 Wochen vor Behandlungsbeginn) berichtet.

*Darstellung im Dossier:*

- **Analyse 1:** Mittlere Änderung des sFerritin-Spiegels im Vergleich zur Baseline; Woche 1 – 24 und Woche 1 – 48
- **Analyse 2:** Mittlere Änderung der täglichen Dosis Eisenchelator im Vergleich zur Baseline; Woche 9 – 24 und Woche 33 – 48

*Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens**Patientenrelevanz:*

Die Erhebung krankheitsbedingter Symptome anhand des EORTC QLQ-C30-Fragebogens stellt einen patientenrelevanten Endpunkt dar, da er die für die Patienten spürbaren Auswirkungen der Erkrankung sowie der Therapie direkt abbildet.

*Validität und Operationalisierung:*

Der EORTC QLQ-C30 ist ein validierter und etablierter Fragebogen zur Erhebung der Symptomatik und Lebensqualität von Patienten mit onkologischen Erkrankungen und ist somit auch für Patienten mit Niedrigrisiko-MDS geeignet. Dies wurde bereits mehrfach in klinischen Studien an Patienten mit behandlungsbedürftiger MDS bestätigt [57]. Vom G-BA wurde der EORTC QLQ-C30-Fragebogen bereits in mehreren Nutzenbewertungsverfahren anerkannt, seine Validität wurde zudem im Beratungsgespräch zu Luspatercept bestätigt [52, 58-61].

Über den EORTC QLQ-C30 (Version 3.0) werden anhand der einzelnen Symptomskalen folgende Symptome erhoben:

- Fatigue,
- Dyspnoe,
- Übelkeit und Erbrechen,
- Schmerz,
- Schlaflosigkeit,
- Appetitminderung,
- Obstipation,
- Diarrhö und
- finanzielle Schwierigkeiten.

Die einzelnen Items wurden auf Skalen von 1 – 4 durch die Patienten bewertet. Entsprechend der Auswertungskonventionen des Handbuchs des Fragebogens wurde für jede Skala zunächst ein Rohwert entsprechend dem Mittelwert der verwertbaren Items gebildet (z. B. Mittelwert der Item-Scores 10, 12 und 18 für die Symptomskala Fatigue), welcher anschließend in einen standardisierten Skalenwert von 0 – 100 transformiert wurde:

$$\text{standardisierter Skalenwert} = ((\text{roher Wert} - 1) / \text{Wertebereich}) \times 100$$

Der Wertebereich ist definiert als Differenz des möglichen Maximalwerts und Minimalwerts. Für alle Symptomskalen entspricht der Wertebereich bei einer möglichen Spanne von 1 bis 4 dementsprechend dem Wert „3“ (= 4 – 1). Höhere Werte entsprechen einer Verschlechterung der Symptome [62].

Als Baseline-Wert wurde der am Tag 1 des ersten Behandlungszyklus erhobene Wert herangezogen. Bei Patienten, die innerhalb von 14 Tagen vor Behandlungsbeginn einen Fragebogen ausgefüllt hatten, wurde dieser als Baseline-Wert herangezogen.

Änderungen der Symptomskalen um fünf bis zehn Punkte werden als klinisch relevant eingestuft, wobei der G-BA im Rahmen der Nutzenbewertung Auswertungen im Sinne eines konservativen Ansatzes über Änderungen um zehn Punkte heranzieht [59, 63]. Dies wurde auch im Beratungsgespräch zu Luspatercept bestätigt [52, 61].

#### Hauptanalyse:

Entsprechend der validierten und anerkannten MCID von zehn Punkten wurden für den EORTC QLQ-C30 Responderanalysen hinsichtlich einer klinisch relevanten Verbesserung der Symptomatik um  $\geq 10$  Punkte während der ersten 24 Wochen der Behandlungsphase durchgeführt.

#### Zusatzanalysen:

Auswertungen über den Studienzeitraum bis Woche 48 werden aufgrund der eingeschränkten Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalysen (**Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte**) dargestellt

Auswertungen zur Änderung der Symptomatik im Vergleich zur Baseline für alle Erhebungszeitpunkte werden ergänzend berichtet und mittels grafischer Darstellung visualisiert. Unterschiede zwischen den Behandlungsarmen zu den einzelnen Erhebungszeitpunkten werden mittels ANOVA (*analysis of variance*) berechnet und bei Vorliegen einer statistischen Signifikanz ( $p < 0,05$ ) in den Grafiken kenntlich gemacht.

Darüber hinaus werden Auswertungen über den gesamten Studienverlauf mittels gemischten Modells für wiederholte Messungen (*mixed-model repeated measure*, MMRM) berechnet. Basierend auf den Ergebnissen des MMRM wurden zur Bewertung klinisch relevanter Änderungen zusätzlich standardisierte Mittelwertdifferenzen (SMD) in Form von Hedges' g dargestellt.

*Darstellung im Dossier:*

- **Hauptanalyse:** Anteil der Patienten mit einer Verbesserung der Symptomatik um mindestens zehn Punkte; Woche 1 – 24
- **Zusatzanalyse 1:** Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte; Woche 1 – 48
- **Zusatzanalyse 2:** Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten im Vergleich zur Baseline; Woche 1 – 48
- **Zusatzanalyse 3:** Mittlere Änderung der Symptomatik im Studienverlauf im Vergleich zur Baseline (MMRM-Analyse); Woche 1 – 24 und Woche 1 – 48

***Hospitalisierungen****Patientenrelevanz:*

Hospitalisierungen haben direkte Auswirkungen auf die Wahrnehmung von Funktionen und Aktivitäten der Patienten und sind infolgedessen patientenrelevant. Zudem werden Patienten in der Regel nur dann stationär aufgenommen, wenn Krankheitssymptome von einem solchen Ausmaß oder einer solchen Schwere aufgetreten sind, dass ihre Versorgung in der ambulanten Behandlung nicht mehr gewährleistet werden kann.

*Operationalisierung und Validität:*

Der Endpunkt Hospitalisierungen ist ein weiteres Maß für die Beeinträchtigung von Patienten durch ihre Krankheit. Streng genommen spiegeln nur Hospitalisierungen aufgrund von MDS diese erkrankungsbezogene Belastung wider. Eine Abgrenzung zwischen stationären Aufnahmen aufgrund von MDS und anderen Ursachen ist basierend auf den vorliegenden Informationen nicht sicher möglich. Näherungsweise werden Hospitalisierungen aufgrund von UE und/oder Transfusionen als erkrankungsbezogene Ereignisse herangezogen.

Hauptanalyse:

Es werden Auswertungen zu Hospitalisierungen jeglicher Ursache (Gesamthospitalisierungen) dargestellt, die seit Randomisierung bis Woche 24 aufgetreten sind. Die Gründe für die Hospitalisierungen wurden während der Studie entsprechend der Kategorien

- UE,
- elektive Behandlung einer vorbestehenden Erkrankung,
- nicht protokollbasierte Untersuchung,
- Behandlung, die vor Unterzeichnung der Einverständniserklärung geplant war,

- protokollbedingte Behandlung,
- sozialer, technischer oder praktischer Grund in Abwesenheit eines UE und
- Transfusion

durch den Prüfarzt erhoben.

Da unter Gesamthospitalisierungen auch solche stationären Aufnahmen zusammengefasst wurden, die z. B. aufgrund von Behandlungen durchgeführt wurden, die bereits vor Unterzeichnung der Einverständniserklärung geplant waren, und somit nicht unmittelbar mit der Behandlung der MDS im Rahmen der Studie im Zusammenhang stehen, wurden in einer ergänzenden Sensitivitätsanalyse nur Hospitalisierungen aufgrund von UE und/oder Transfusionen als Ereignis berücksichtigt.

#### Zusatzanalysen:

Auswertungen über den Studienzeitraum bis zum Datenschnitt vom 01.07.2019 werden aufgrund der eingeschränkten Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalyse (Zeit bis zur ersten Hospitalisierung) dargestellt.

#### *Darstellung im Dossier:*

- **Hauptanalyse:** Anteil der Patienten mit mindestens einer Hospitalisierung; Woche 1 – 24
- **Sensitivitätsanalyse:** Anteil der Patienten mit mindestens einer Hospitalisierung aufgrund von UE und/oder Transfusionen; Woche 1 – 24
- **Zusatzanalyse 1:** Zeit bis zur ersten Hospitalisierung; Woche eins bis zum Datenschnitt vom 01.07.2019

## **Lebensqualität**

### *Patientenrelevanz:*

Die Erhebung der Lebensqualität ist ein patientenrelevanter Endpunkt, da die Auswirkungen der Erkrankung sowie der Behandlung auf das Befinden des Patienten direkt abgebildet werden. Zur Erhebung der Lebensqualität empfiehlt der G-BA die Verwendung eines generischen als auch krankheitsspezifischen Fragebogens [52].

### ***Lebensqualität gemessen anhand des generischen EORTC QLQ-C30-Fragebogens***

#### *Validität und Operationalisierung:*

Wie bereits erläutert, ist der EORTC QLQ-C30 ein validierter und etablierter Fragebogen zur Erhebung der Symptomatik und Lebensqualität von Patienten mit onkologischen Erkrankungen und ist als solcher auch für Patienten mit Niedrigrisiko-MDS geeignet. Dieser generische Fragebogen wurde auch vom G-BA im Beratungsgespräch zu Luspatercept sowie in mehreren Nutzenbewertungsverfahren als valide und geeignet anerkannt [52, 58-61].

Die Lebensqualität wurde anhand der einzelnen Funktionsskalen des EORTC QLQ-C30 (Version 3.0) bewertet:

- körperliche Funktion,
- Rollenfunktion,
- emotionale Funktion,
- kognitive Funktion,
- soziale Funktion und
- globaler Gesundheitsstatus

Die einzelnen Items wurden auf Skalen von 1 – 4 bzw. 1 – 7 für den globalen Gesundheitsstatus durch die Patienten bewertet. Entsprechend der Auswertungskonventionen des Handbuchs des Fragebogens und analog zu den Symptomskalen wurde für jede Skala zunächst ein roher Wert entsprechend dem Mittelwert der verwertbaren Items gebildet. Die Umwandlung des rohen Wertes für den globalen Gesundheitsstatus in einen standardisierten Skalenwert von 0 – 100 erfolgte analog zu der Vorgehensweise der Symptomskalen. Für die Funktionsskalen wurde die Transformation der rohen Werte zu standardisierten Skalenwerten wie folgt berechnet:

$$\text{standardisierter Skalenwert} = 1 - ((\text{Rohwert} - 1) / \text{Wertebereich}) \times 100.$$

Der Wertebereich ist definiert als Differenz des möglichen Maximalwerts und Minimalwerts. Für alle Funktionsskalen entspricht der Wertebereich bei einer möglichen Spanne von eins bis vier dementsprechend dem Wert „drei“ (= 4 – 1). Der globale Gesundheitsstatus wurde auf eine Skala von eins bis sieben bewertet, sodass hier ein Wert von „sechs“ (= 7 – 1) in den Wertebereich einfließt. Höhere Werte entsprechen einer Verbesserung der Lebensqualität bzw. einem verbesserten Gesundheitsstatus [62].

Als Baseline-Wert wurde der am Tag eins des ersten Behandlungszyklus erhobene Wert herangezogen. Bei Patienten, die innerhalb von 14 Tagen vor Behandlungsbeginn einen Fragebogen ausgefüllt hatten, wurde dieser als Baseline-Wert herangezogen.

Hauptanalyse:

Analog zur Darstellung der Ergebnisse der Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens wurden für den globalen Gesundheitsstatus und die Funktionskalen des EORTC QLQ-C30 Responderanalysen hinsichtlich einer klinisch relevanten Verbesserung der Lebensqualität bzw. des Gesundheitsstatus um die MCID von zehn Punkten durchgeführt [63]. Eine Änderung der Skalen um mindestens zehn Punkte entspricht der klinischen Relevanzschwelle, die auch vom G-BA bereits in verschiedenen Nutzenbewertungen herangezogen und im Beratungsgespräch zu Luspatercept bestätigt wurde [52, 59, 61].

Zusatzanalysen:

Auswertungen über den Studienzeitraum bis Woche 48 werden aufgrund der eingeschränkten Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalyse (**Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte**) dargestellt.

Auswertungen zur **mittleren Änderung der gesundheitsbezogenen Lebensqualität** im Vergleich zur Baseline für alle Erhebungszeitpunkte werden ergänzend berichtet und mittels grafischer Darstellung visualisiert. Unterschiede zwischen den Behandlungsarmen zu den einzelnen Erhebungszeitpunkten wurden mittels ANOVA berechnet und bei Vorliegen einer statistischen Signifikanz ( $p < 0,05$ ) in den Grafiken kenntlich gemacht.

Darüber hinaus wurden Auswertungen über den gesamten Studienverlauf mittels MMRM berechnet. Basierend auf den Ergebnissen des MMRM wurden zur Bewertung klinisch relevanter Änderungen zusätzlich SMD in Form von Hedges' g dargestellt.

*Darstellung im Dossier:*

- **Hauptanalyse:** Anteil der Patienten mit einer Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte; Woche 1 – 24
- **Zusatzanalyse 1:** Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte; Woche eins bis zum Datenschnitt vom 01.07.2019
- **Zusatzanalyse 2:** Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten im Vergleich zur Baseline; Woche 1 – 48
- **Zusatzanalyse 3:** Mittlere Änderung der gesundheitsbezogenen Lebensqualität im Studienverlauf im Vergleich zur Baseline (MMRM-Analyse); Woche 1 – 24 und Woche 1 – 48

**Lebensqualität gemessen anhand des krankheitsspezifischen Fragebogens QoL-E***Validität und Operationalisierung:*

Der QoL-E ist der erste Fragebogen, der speziell für Patienten mit MDS zur Erhebung der krankheitsspezifischen Lebensqualität entwickelt wurde [57, 64]. Dies unterscheidet ihn von zuvor in MDS-Studien eingesetzten Fragebögen, bei denen es sich ausschließlich um generische Instrumente, wie z. B. den *Functional Assessment of Cancer Therapy* (FACT), handelt.

Der QoL-E beinhaltet 29 Items, von denen 27 Items zu insgesamt sechs Skalen zusammengefasst werden. Die sechs Skalen beschreiben neben den Gesundheitsdimensionen körperliches, funktionelles, soziales und sexuelles Wohlbefinden auch krankheitsbezogene Dimensionen (Fatigue und MDS-bezogene Störungen) [64]. Dementsprechend wird insbesondere auch der krankheitsspezifischen Belastung durch z. B. EK-Transfusionen Rechnung getragen.

Aus den einzelnen Dimensionen lassen sich Summenskalen des allgemeinen Wohlbefindens (Summe aller Dimensionen ohne MDS-spezifische Störungen), des Behandlungseffekts (*Treatment Outcome Index*, TOI; Summe der physischen, funktionellen und MDS-spezifischen Dimensionen) sowie eine Gesamtsummenskala (Summe aller Dimensionen) ableiten (siehe Tabelle 4-4).

Tabelle 4-4: Übersicht der Dimensionen und Summenskalen des QoL-E Fragebogens

Dimension des QoL-E-Fragebogens	Anzahl Items	Item-Nummer (Version 3)	Summenskalen		
			Allgemeines Wohlbefinden	Behandlungseffekt (TOI)	Gesamtsummenskala
Physisches Wohlbefinden (QoL-FIS)	4	3a-d	X	X	X
Funktionelles Wohlbefinden (QoL-FUN)	3	4a-b, 5	X	X	X
Soziales und Familienleben (QoL-SOC)	4	6a-c, 7	X		X
Sexuelles Wohlbefinden (QoL-SEX)	2	8, 14f	X		X
Fatigue (QoL-FAT)	7	9, 10, 11a-d, 12	X		X
MDS-spezifische Störungen (QoL-SPEC)	7	13, 14a-e, 14g		X	X

MDS = Myelodysplastische Syndrome; TOI = Treatment Outcome Index

Die Entwicklung und Validierung des Fragebogens erfolgte auf Grundlage einer Datenbasis aus sieben Studien mit insgesamt 199 Niedrigrisiko-MDS-Patienten, die zwischen 2002 und 2010 in Italien, den USA und Bulgarien durchgeführt worden sind [64]. Die Erhebung des QoL-E erfolgte in diesen Studien als Querschnittsanalyse parallel zu anderen klinischen Endpunkten und Lebensqualitätserhebungen wie dem EORTC QLQ-C30.

Die gute psychometrische Güte des Instruments lässt sich anhand der folgenden typischen Kenngrößen zur Charakterisierung der Validität darstellen [64]:

- **Konstruktvalidität:** Es wurden starke Zusammenhänge zwischen der ersten Version des QoL-E und den Domänen des FACT-*General* (FACT-G) ermittelt ( $r \geq 0,71$ ), insbesondere für die Gesundheitsdimensionen körperliches, funktionelles, soziales und allgemeines Wohlbefinden sowie für den Summenscore Behandlungseffekt. Darüber hinaus wurden gute Korrelationen zwischen den QoL-E-Domänen und relevanten klinischen Parametern der Krankheitsschwere, wie z. B. Hb-Wert und Transfusionsstatus, bestimmt und der starke Zusammenhang mit FACT-G in der großen Validierungskohorte belegt [64].
- **Reliabilität:** Eine interne Konsistenz von Cronbachs  $\alpha$  von  $\geq 0,70$  wurde über alle Domänen des QoL-E ermittelt. Ebenfalls konnte eine ausreichende Test-Retest-Reliabilität ( $> 0,70$ ) gezeigt werden [64].
- **Floor- und Ceiling-Effekte:** Es wurden keine Floor- und Ceiling-Effekte beobachtet. Die Mittelwerte aller Domänen lagen zwischen 43,6 und 68,1 Punkten auf einer Skala von 0 – 100 (95 %-KI: [39,3; 71,34]) [64].

Es handelt sich somit um ein valides und geeignetes Instrument zur Erhebung der krankheits-spezifischen Lebensqualität bei Patienten mit Niedrigrisiko-MDS.

Jede Dimension oder Summenskala wird auf eine standardisierte Skala von 0 – 100 transformiert, sodass höhere Werte einer verbesserten Lebensqualität bzw. einem verbesserten Gesundheitsstatus entsprechen. Die Transformation der Werte auf diese standardisierte Skala erfolgt in vier Schritten:

1. Re-Skalierung der Rohwerte, sodass für alle Items ein höherer Wert einer besseren Lebensqualität entspricht (z. B. Item eins: höchster Rohwert („vier“) entspricht schlechtester Lebensqualität; Re-Skalierung Item eins = 5 – roher Wert).
2. Transformation der re-skalierten Werte in standardisierte Werte von 0 – 100:

$$\text{standardisierter Wert} = \frac{\text{re-skaliertes Wert} - \text{niedrigster möglicher Wert}}{\text{Wertebereich}} \times 100$$

Die Scores der Dimensionen werden als Mittelwert der einzelnen standardisierten Item-Werte berechnet. Dimensionen, für die  $\leq 50$  % der Items beantwortet wurden, werden als fehlende Werte behandelt. Alle Items werden gleichermaßen gewichtet.

Summenskalen werden als Mittelwerte der jeweiligen Dimensionen berechnet. Fehlten die Werte von zumindest einer der enthaltenen Dimensionen, wurde auch die resultierende Summenskala als fehlend angesehen. Eine Ausnahme stellt die Summenskala allgemeines Wohlbefinden dar, welche auch berechnet wurde, wenn für die sexuelle Dimension keine Werte vorlagen.

Als Baseline-Wert wurde der am Tag eins des ersten Behandlungszyklus erhobene Wert herangezogen. Bei Patienten, die innerhalb von 14 Tagen vor Behandlungsbeginn einen Fragebogen ausgefüllt hatten, wurde dieser als Baseline-Wert herangezogen.

#### Hauptanalyse:

Auswertungen während der primären Behandlungsphase (Woche 1 – 24) wurden mittels MMRM berechnet. Basierend auf den Ergebnissen des MMRM werden zur Bewertung klinisch relevanter Änderungen zusätzlich SMD in Form von Hedges' g berichtet.

#### Zusatzanalysen:

Auswertungen über den Studienzeitraum bis Woche 48 werden aufgrund der eingeschränkten Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend berichtet.

Zur vollumfänglichen Bewertung der krankheitsspezifischen Lebensqualität werden ergänzend Auswertungen zur Änderung der krankheitsspezifischen Lebensqualität im Vergleich zur Baseline für alle Erhebungszeitpunkte berechnet und mittels grafischer Darstellung visualisiert. Unterschiede zwischen den Behandlungsarmen zu den einzelnen Erhebungszeitpunkten werden mittels ANOVA berechnet und bei Vorliegen einer statistischen Signifikanz ( $p < 0,05$ ) in den Grafiken kenntlich gemacht.

#### *Darstellung im Dossier:*

- **Hauptanalyse:** Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf im Vergleich zur Baseline (MMRM-Analyse); Woche 1 – 24
- **Zusatzanalyse 1:** Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf im Vergleich zur Baseline (MMRM-Analyse); Woche 1 – 48
- **Zusatzanalyse 2:** Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten im Vergleich zur Baseline; Woche 1 – 24 und Woche 1 – 48

## **Sicherheit**

### *Unerwünschte Ereignisse*

#### *Patientenrelevanz:*

Die Betrachtung von UE ist bei der Bewertung des medizinischen Nutzens und Zusatznutzens unerlässlich, um die Verträglichkeit und Sicherheit eines Arzneimittels im Behandlungskontext zu bewerten, und stellt somit einen patientenrelevanten Endpunkt dar. Neben den Inzidenzen von UE wird der Schweregrad und die Reversibilität bei der Bewertung berücksichtigt, da dies direkten Einfluss auf die Implikation des UE für den Patienten hat.

### *Operationalisierung und Validität:*

Gemäß § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, **der Verringerung von Nebenwirkungen** oder einer Verbesserung der Lebensqualität. Auswertungen zur Sicherheit, d. h. dem Auftreten von Nebenwirkungen, stellen valide Endpunkte dar, um den Nutzen eines Arzneimittels vollumfänglich zu bewerten,

Bei der Erhebung der UE wurden alle Ereignisse ab Behandlungsbeginn bis Woche 24 berücksichtigt, und zwar unabhängig von einem möglichen direkten Kausalzusammenhang mit der Studienmedikation.

Die Kodierung der UE wurde mittels des *Medical Dictionary for Regulatory Activities* (MedDRA) Version 20.0 durchgeführt. In den Darstellungen wurden jeweils die UE mit höchstem Schweregrad berücksichtigt.

### Hauptanalysen:

Im Rahmen der Nutzenbewertung werden die Inzidenzen der folgenden UE-Kategorien bewertet:

- Gesamtrate jeglicher UE,
- Gesamtrate SUE,
- Gesamtrate schwerer UE, definiert als UE mit NCI-CTCAE Version 4.03 Grad  $\geq 3$  [65] und
- Gesamtrate der Therapieabbrüche aufgrund von UE.

Zusätzlich werden die UE dieser Kategorien differenziert auf Ebene der SOC und PT dargestellt.

Bei der Erhebung von UE wurden Ereignisse berücksichtigt, welche als Symptomatik der Anämie über den EORTC QLQ-C30-Fragebogen sowie als stationäre Einweisungen über die Hospitalisierungen bereits erhoben und dargestellt wurden. Um eine Doppelbewertung dieser Morbiditätsereignisse in der Betrachtung der UE zu vermeiden, blieben solche UE-Ereignisse, welche Anämie-Symptome darstellen oder eine stationäre Aufnahme des Patienten bedingten, in einer ergänzenden Auswertung zu den Gesamtraten an jeglichen UE, SUE und schweren UE (CTCAE  $\geq 3$ ) in Form von Sensitivitätsanalysen unberücksichtigt.

Fatigue und Dyspnoe sind Symptome der Anämie, die gleichzeitig über den EORTC QLQ-C30-Fragebogen und als UE erhoben worden sein konnten. Um eine Doppelbewertung in den Kategorien Morbidität und Nebenwirkungen auszuschließen, wurden entsprechende Ereignisse in einer Sensitivitätsanalyse zu den Gesamtraten an jeglichen UE, SUE und schweren UE (CTCAE  $\geq 3$ ) nicht berücksichtigt. Hierfür relevante PT wurden durch Vergleich der entsprechenden Fragen im EORTC QLQ-C30 mit den Begriffen niedrigster Ebene (*lowest level terms*,

LLT) in MedDRA identifiziert (siehe Tabelle 4-5). Das Symptom Dyspnoe wurde im EORTC QLQ-C30-Fragebogen über eine Frage erhoben. Da davon auszugehen ist, dass Patienten mit Anämie unter Kurzatmigkeit sowohl Dyspnoe an sich als auch das typische Symptom der Anämie, die Belastungsdyspnoe verstehen, wurden beide PT als erkrankungsbezogene Ereignisse ausgewählt.

Tabelle 4-5: Übersicht der über den EORTC QLQ-C30-Fragebogen erhobenen Symptome der Anämie auf Ebene der UE

Symptomskala	Zugehöriges Einzelitem	LLT	PT
Fatigue	10. Mussten Sie sich ausruhen?	Erschöpfung	Ermüdung
	12. Fühlten Sie sich schwach?	Fühlte sich schwach	Schwindelgefühl
	18. Waren Sie müde?	Müdigkeit	Ermüdung
Dyspnoe	8. Waren Sie kurzatmig?	Kurzatmigkeit Belastungsdyspnoe	Dyspnoe Belastungsdyspnoe
LLT = Begriff niedrigster Ebene ( <i>lowest level term</i> ); PT = bevorzugter Begriff ( <i>preferred term</i> )			

#### Zusatzanalysen:

Auswertungen über den gesamten Studienzeitraum bis zum Datenschnitt vom 01.07.2019 werden aufgrund der eingeschränkten Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalysen (**Zeit bis zum ersten UE**) berichtet.

Luspatercept greift als Erythrozyten-Reifungs-Aktivator in die Erythropoese ein. Um auszuschließen, dass dieser Eingriff in die Hämatopoese bösartige Aberrationen zur Folge hat, wurde das Auftreten anderen prämaligener und maligner Erkrankungen als **UE von besonderem Interesse**, jeweils erhoben als SMQ, untersucht. Die Ergebnisse werden unter Berücksichtigung schwerer (CTCAE  $\geq 3$ ) und schwerwiegender Ereignisse dargestellt.

Im Beratungsgespräch mit dem G-BA wurde neben den bereits genannten Endpunkten auf die Patientenrelevanz der Endpunkte Infektionen und Blutungen hingewiesen. Diese Symptome sind keine typischen Nebenwirkungen einer isolierten Anämie, können bei MDS-Patienten jedoch mitunter als Folge einer gleichzeitigen Neutropenie bzw. Thrombozytopenie auftreten [2]. Das vorliegende Anwendungsgebiet umfasst nur die Behandlung von Anämie aufgrund von MDS, sodass weder Infektionen noch Blutungen relevante Endpunkte im Anwendungsgebiet darstellen. Das Auftreten von Neutropenien und Thrombozytopenien wird dennoch ergänzend unter der Kategorie Sicherheit als UE von besonderem Interesse adressiert, da diese Ereignisse sich als weitere Manifestationen der MDS niederschlagen können und in die Risikostratifizierung gemäß IPSS-R der Patienten mit einfließen. Schwere Thrombozytopenien stellen darüber hinaus unabhängige prognostische Faktoren für das Überleben dar [5, 66]. Die Ergebnisse werden unter Berücksichtigung schwerer (CTCAE  $\geq 3$ ) und schwerwiegender Ereignisse dargestellt.

Aufgrund der chronischen, progressiven Pathophysiologie der MDS, stellen im vorliegenden Anwendungsgebiet das Voranschreiten der Erkrankung, die **Progression in ein Hochrisiko-MDS bzw. in eine AML**, bedeutende Sicherheitsaspekte von Arzneimitteln im Anwendungsgebiet dar. Eine Progression in ein Hochrisiko-MDS war definiert als Ereignisse des PT „Anämie refraktär mit Blastenüberschuss“ oder „Myelodysplastisches Syndrom“. Da es sich im vorliegenden Fall um Patienten mit Niedrigrisiko-MDS handelt, welche eine niedrige Progressionsrate aufweisen (mediane Zeit bis zur Progression in eine AML bei 25 % der Patienten mit Niedrigrisiko-MDS vs. Hochrisiko-MDS: 3,2 – N.E. Jahre vs. 0,73 – 1,4 Jahre nach IPSS-R [37]), stellt die Verzögerung einer Progression kein primäres Therapieziel dar. Zweck der gesonderten Bewertung dieser Ereignisse ist, dass ein Voranschreiten der Erkrankung und/oder Auftreten von Folgeerkrankungen durch die Arzneimitteltherapie ausgeschlossen werden kann. Dementsprechend werden diese Ergebnisse als klinisch relevante Zusatzanalysen gesondert berichtet.

Die Zeit bis zur Progression in eine AML war operationalisiert als Zeit zwischen der Randomisierung und der Diagnose einer AML gemäß WHO-Klassifikation, definiert als Blastenanteil im peripheren Blut oder Knochenmark von  $\geq 20\%$ . Patienten ohne Progression in eine AML wurden zum letzten Datum zensiert, an dem bekannt war, dass keine Progression vorlag.

Ergänzend werden die Ergebnisse zur Inzidenz der Progression in ein Hochrisiko-MDS dargestellt.

#### *Darstellung im Dossier:*

- **Hauptanalyse 1:** Gesamtraten UE; Woche 1 – 24
- **Sensitivitätsanalyse:** Gesamtraten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen; Woche 1 – 24
- **Hauptanalyse 2:** Inzidenzen von UE auf Ebene der SOC und PT; Woche 1 – 24
- **Zusatzanalyse 1:** Zeit bis zum ersten UE; Woche eins bis zum Datenschnitt vom 01.07.2019
- **Zusatzanalyse 2:** Zeit bis zum ersten UE auf Ebene der SOC und PT; Woche eins bis zum Datenschnitt vom 01.07.2019
- **Zusatzanalyse 3:** UE von besonderem Interesse
  - Gesamtraten UE von besonderem Interesse; Woche 1 – 24
  - Zeit bis zum ersten UE von besonderem Interesse; Woche eins bis zum Datenschnitt vom 01.07.2019
- **Zusatzanalyse 4:** Zeit bis zur Progression in eine AML; Woche eins bis zum Datenschnitt vom 01.07.2019

- **Zusatzanalyse 5:** Progression in ein Hochrisiko-MDS; Woche eins bis zum Datenschnitt vom 01.07.2019

### **Verwendete statistische Methodik**

Die Ergebnisse zu den einzelnen Endpunkten wurden entsprechend der Operationalisierungen und Vorgaben des Studienberichts und des statistischen Analyseplans ausgewertet und dargestellt. Davon ausgenommen ist die Wahl des Betrachtungszeitraums, worauf sich auch kleine Abweichungen zu den im Studienbericht dargestellten Ergebnissen begründen lassen. Hintergrund für dieses Vorgehen ist im folgenden Absatz „Analysezeitraum“ erläutert. Zusätzliche Analysen ergaben sich darüber hinaus aus den Vorgaben der Dossiervorlage.

Bei Analysen, deren Zielvariable über bestimmte Intervalle betrachtet wird, wird – sofern nicht anders angegeben – jegliches zusammenhängendes Intervall bestimmter Länge mit der größtmöglichen Änderung herangezogen, das innerhalb des Beobachtungszeitraums aufgetreten ist (bspw. mittlere Änderung des Hb-Werts in jeglichem 16-Wochen-Intervall während der primären Behandlungsphase bis Woche 24).

### **Analysezeitraum**

Die Studie MEDALIST ist in zwei Behandlungsphasen aufgeteilt. Die erste Behandlungsphase (Woche 1 – 24) umfasst alle eingeschlossenen und randomisierten Patienten und ist somit direkt geeignet, um Aussagen zum medizinischen Nutzen und Zusatznutzen von Luspatercept gegenüber dem derzeitigen Behandlungsstandard, einer regelmäßigen, bedarfsgerechten Transfusionstherapie in Kombination mit Eisenchelatoren, abzuleiten.

In die zweite Behandlungsphase (ab Woche 25) durften nur Patienten übergehen, die nach Ermessen des Prüfarztes von der Behandlung in den ersten 24 Wochen profitiert haben. Diejenigen Patienten, für die der Prüfarzt keinen Nutzen der Therapie feststellen konnte oder die eine Progression der Erkrankung aufwiesen, wurden von der Teilnahme dieser zweiten Behandlungsphase ausgeschlossen und gingen direkt in die Posttreatment-Follow-up-Phase über. Somit kommt es in der zweiten Studienphase zu einer Imbalance in den Anteilen der Patienten, die bis Woche 48 (Luspatercept vs. Placebo: 51 % vs. 15,8 %) und darüber hinaus behandelt wurden. Zum Zeitpunkt des Datenschnitts vom 07.01.2019 hatten alle Patienten im Placebo-Arm und 68,6 % der Patienten im Luspatercept-Arm die Behandlung abgebrochen. Daraus resultiert eine Unsicherheit hinsichtlich der Interpretierbarkeit (siehe Abschnitt 4.3.1.2).

Die Auswertungen der Ergebnisse der Hauptanalysen beziehen sich dementsprechend auf die primäre Behandlungsphase (Woche 1 – 24), die bereits zum Datenschnitt vom 08.05.2018 vollständig für alle Patienten vorlag. Neue Erkenntnisse hinsichtlich des Zusatznutzens von Luspatercept basierend auf den späteren Datenschnitten aus dem Jahr 2019 sind nicht zu erwarten. Auch die Analysen für den Endpunkt Gesamtüberleben werden vor dem Hintergrund

der Erkrankung als reif angesehen. Für ein vollständiges Bild zu Wirksamkeits- und Sicherheitsaspekten werden die Ergebnisse über den Studienverlauf bis zum letzten verfügbaren Datenschnitt (01.07.2019) ergänzend als Zusatzanalysen, sofern möglich in Form von Ereigniszeitanalysen, berichtet.

### **Analysepopulationen**

Gemäß Vorgabe des statistischen Analyseplans und Studienberichts wurden alle Auswertungen der Endpunktkategorien Mortalität und Morbidität, mit Ausnahme der Symptomatik, bezogen auf die ITT-Population berechnet. Die ITT-Population umfasst alle randomisierten Patienten entsprechend ihrer Gruppeneinteilung und zwar unabhängig davon, welche Studienmedikation die Patienten tatsächlich erhalten haben.

Für die patientenberichteten Endpunkte zur Erhebung der Lebensqualität und Symptomatik wurde gemäß statistischem Analyseplan die HRQoL-Population herangezogen, d. h. alle randomisierten Patienten, für die zur Baseline und zu mindestens einem Zeitpunkt während der Behandlungsphase eine Bewertung der Lebensqualität bzw. Symptomatik vorliegt.

Sicherheitsendpunkte wurden anhand der Safety-Population ausgewertet. Diese Analysepopulation umfasst alle randomisierten Patienten, die mindestens eine Dosis Studienmedikation erhalten haben. Patienten werden dem Behandlungsarm zugeordnet, dessen Behandlung sie tatsächlich erhalten haben.

Die Verteilung der Patienten auf die herangezogenen Analysepopulationen ist in Tabelle 4-6 zusammengefasst.

Tabelle 4-6: Analysepopulationen auf Studienebene aus MEDALIST

<b>Studie MEDALIST (Luspatercept + BSC vs. Placebo + BSC<sup>(1)</sup>)</b>	<b>Luspatercept + BSC<sup>(1)</sup> n (%)<sup>(5)</sup></b>	<b>Placebo + BSC<sup>(1)</sup> n (%)<sup>(5)</sup></b>
ITT-Population <sup>(2)</sup>	153 (100)	76 (100)
Safety-Population <sup>(3)</sup>	153 (100)	76 (100)
HRQoL-Population <sup>(4)</sup>	149 (97,4)	76 (100)

<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> ITT-Population umfasst alle randomisierten Patienten, unabhängig davon, ob sie die Studienmedikation erhalten haben.  
<sup>(3)</sup> Safety-Population umfasst alle randomisierten Patienten, die mindestens 1 Dosis Studienmedikation erhalten haben. Patienten werden der Studiengruppe zugeordnet, deren Behandlung sie tatsächlich erhalten haben.  
<sup>(4)</sup> HRQoL-Population umfasst alle Patienten der ITT-Population, für die eine Bewertung der Lebensqualität zur Baseline und mindestens ein Zeitpunkt während der 48 Wochen Behandlungsphase vorliegt.  
<sup>(5)</sup> Prozentualer Anteil bezogen auf alle randomisierten Patienten des Behandlungsarms.

BSC = Best-Supportive-Care; HRQoL = gesundheitsbezogene Lebensqualität (*health related quality of life*); ITT = Intention-to-Treat  
Quelle: [1]

### ***Analysen mit binären Messwerten***

Für dichotome Zielvariablen wurden zusätzlich zu den absoluten und prozentualen Anteilen das Odds Ratio (OR), das RR und die Risikodifferenz (RD) inklusive 95 %-KI mit Hilfe des Cochran-Mantel-Haenszel (CMH)-Tests berechnet. Sofern präspezifiziert, wurde der CHM-Test nach durchschnittlicher Transfusionslast und IPSS-R-Score zur Baseline stratifiziert.

### ***Analysen mit stetigen Messwerten***

Für stetige Variablen wurden Änderungen zur Baseline anhand von Mittelwerten inklusive SD berechnet und Unterschiede zwischen den Behandlungsarmen mittels ANCOVA (*analysis of covariance*) mit durchschnittlicher Transfusionslast und IPSS-R-Score zur Baseline als Kovariate bestimmt. Nur Patienten mit Werten zur Baseline und dem Auswertungszeitpunkt gingen in die Berechnungen mit ein. Zur Bestimmung klinisch relevanter Effekte wurden entsprechend den Empfehlungen des Methodenpapiers des IQWiG SMD in Form von Hedges' g berechnet [35]. Als Irrelevanzschwelle wurde je nach Richtung der Effektschätzer +0,2 oder -0,2 angewendet. Bei einem 95 %-KI, das vollständig ober- bzw. unterhalb dieser Irrelevanzschwelle liegt, wurde von einem klinisch relevanten Effekt ausgegangen.

Für mit Skalen erhobene, patientenberichtete Endpunkte wurden neben Responderanalysen, d. h. Auswertungen als dichotome Zielvariable mit prädefiniertem Ansprechkriterium, auch Auswertungen zur Änderung im Vergleich zur Baseline berechnet und dargestellt. Dabei wurden einerseits Unterschiede zwischen den Behandlungsarmen mittels ANOVA zu den einzelnen Erhebungszeitpunkten im Studienverlauf bestimmt und andererseits Auswertungen unter vollständiger Berücksichtigung aller über den Studienverlauf ermittelten Informationen mittels MMRM-Analyse berechnet. Änderungen im Vergleich zur Baseline flossen dabei als abhängige Variable und Werte zur Baseline, der Behandlungsarm, die Zeit seit Behandlungsbeginn und der Interaktionsterm Behandlung  $\times$  Zeit als Kovariate in das MMRM-Modell ein.

### ***Ereigniszeitanalysen***

Ereigniszeitanalysen wurden mittels Kaplan-Meier-Methode berechnet und Behandlungseffekte wurden durch stratifizierte Log-Rank-Tests verglichen. Die mediane Ereigniszeit sowie Hazard Ratios (HR), berechnet anhand von Cox-Modellen, wurden jeweils einschließlich der 95 %-KI und des Kaplan-Meier-Plots dargestellt. Bei Analysen zu verschiedenen Datenschnitten wurde der Kaplan-Meier-Plot zum jeweils letzten Datenschnitt dargestellt, da dieser alle relevanten Informationen enthält.

### 4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten nach der Knapp-Hartung-Methode mit der Paule-Mandel-Methode zur Heterogenitätsschätzung<sup>6</sup> erfolgen. Im Fall von sehr wenigen Studien ist die Heterogenität nicht verlässlich schätzbar. Liegen daher weniger als 5 Studien vor, ist auch die Anwendung eines Modells mit festem Effekt oder eine qualitative Zusammenfassung in Betracht zu ziehen. Kontextabhängig können auch alternative Verfahren wie z. B. Bayes'sche Verfahren oder Methoden aus dem Bereich der generalisierten linearen Modelle in Erwägung gezogen werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet<sup>7</sup> werden.

---

<sup>6</sup> Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. *Cochrane Database Syst Rev* 2015: 25-27.

<sup>7</sup> Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand geeigneter statistische Maße auf Vorliegen von Heterogenität<sup>8, 6</sup> erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik  $\geq 0,05$ ), soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

*Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

Es lag eine relevante Studie vor, sodass keine Meta-Analysen durchgeführt wurden.

#### 4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

*Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

---

<sup>8</sup> Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.

Zu folgenden Endpunkten wurden Sensitivitätsanalysen durchgeführt:

### **Hospitalisierungen**

Unter Gesamthospitalisierungen wurden stationäre Aufnahmen jeglicher Ursache zusammengefasst, darunter auch Hospitalisierungen, die nicht im direkten Zusammenhang mit der Studienmedikation und/oder zugrundeliegenden Erkrankung stehen. Dementsprechend wurden Hospitalisierungen aufgrund von UE und/oder Transfusionen als Sensitivitätsanalysen mit engerem Bezug zur Therapie der Erkrankung bzw. der Erkrankung an sich dargestellt.

### **UE**

Auswertungen zu UE ohne Berücksichtigung von Ereignissen, die Symptome der Anämie darstellen und bereits über den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens erhoben wurden, sowie UE, die stationäre Aufnahmen verursachten und bereits über Morbidität – Hospitalisierungen erhoben und berichtet wurden, werden als Sensitivitätsanalysen der Gesamtraten jeglicher UE, SUE und schwerer UE dargestellt.

Die Auswahl der PT, welche Symptome der Anämie darstellen und bei diesen Auswertungen unberücksichtigt blieben, erfolgte durch Zuordnung von LLT nach MedDRA für die Einzelitems der Symptomskalen des EORTC QLQ-C30-Fragebogens und anschließende Zusammenfassung auf Ebene der zugehörigen PT (siehe Abschnitt 4.2.5.2, Tabelle 4-5).

#### **4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren**

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. –stadium
- Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen (zu ergänzenden Kriterien zur Darstellung siehe Abschnitt 4.3.1.3.2).

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

*Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

Subgruppenanalysen dienen der Identifikation von Patientengruppen und/oder Behandlungscharakteristika, die hinsichtlich der Wirksamkeit und Sicherheit besonders von der Therapie mit dem zu bewertenden Arzneimittel im Vergleich zur zweckmäßigen Vergleichstherapie profitieren bzw. nicht profitieren. Um Unterschiede in der Effektivität der Behandlung in den spezifizierten Subgruppen zu untersuchen, wurden die Ergebnisse aller untersuchten, patientenrelevanten Endpunkte auf potentielle Effektmodifikationen hin geprüft. Hierzu wurden, mit Ausnahme der Subgruppenmerkmale Ethnische Zugehörigkeit, die im Studienprotokoll präspezifizierten Subgruppenmerkmale herangezogen.

Auf eine Darstellung der präspezifizierten Subgruppen „weiß“, „schwarz oder afroamerikanisch“, „andere“ und „nicht berichtet“ des Subgruppenmerkmals Ethnische Zugehörigkeit wurde verzichtet. Hintergrund ist, dass in den Subgruppen „schwarz oder afroamerikanisch“ (ein Patient im Luspatercept-Arm) und „andere“ (jeweils ein Patient in beiden Behandlungsarmen) insgesamt nur sehr wenig Patienten enthalten waren, die eine sinnvolle Interpretation nicht möglich machen (siehe auch Tabelle 4-14). Eine sinnvolle medizinische Interpretation von potentiellen Effektmodifikationen zwischen der Subgruppe „weiß“ und „nicht berichtet“ ist ebenfalls nicht möglich, da diese beiden Subgruppen nicht sinnvoll voneinander abgegrenzt werden können.

Darüber hinaus wurde das Subgruppenmerkmal Region post-hoc ergänzend analysiert, um die Ergebnisse für den Behandlungskontext gemäß der Verfo des G-BA besser beurteilen zu können. Auf eine Analyse nach Länder- oder Zentrumseffekten wurde verzichtet, da aussagekräftige Ergebnisse aufgrund von geringen Patientenzahlen pro Studienzentrum bzw. pro Land nicht zu erwarten und somit nicht geeignet sind, um Schlussfolgerungen für diese Subgruppen abzuleiten. Stattdessen wurden Subgruppenanalysen zu den übergeordneten, geografischen Regionen durchgeführt. Dies entspricht auch den Empfehlungen der EMA in der „Guideline on

*the investigation of subgroups in confirmatory trials*“ und der ICH (*International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use*)-Leitlinie zur Planung und Durchführung von multizentrischen klinischen Studien (ICH-E17: *General principles for planning and design of multi-regional clinical trials*), gemäß denen bei multizentrischen Studien mit vielen Studienzentren und beteiligten Ländern und daraus resultierenden kleinen Patientenzahlen pro Zentrum und Land eine Betrachtung der Regionen sinnvoller und somit zu präferieren ist [67, 68]. In MEDALIST eingeschlossene Patienten waren weitestgehend gleichmäßig auf über 60 Studienzentren in elf Ländern verteilt, sodass zudem nicht von einem Ungleichgewicht aufgrund von einzelnen, sehr großen Zentren oder Ländern auszugehen ist. Das Subgruppenmerkmal „Nordamerika“ fasst die in Kanada und den USA behandelten Patienten zusammen, während die Länder Belgien, Deutschland, Frankreich, Großbritannien, Italien, Niederlande, Schweden, Spanien und die Türkei unter dem Subgruppenmerkmal Europa zusammengefasst werden.

Die herangezogenen Grenzen für kategorisierte Subgruppenmerkmale waren im statistischen Analyseplan prädefiniert und entsprechen den üblicherweise verwendeten Kategorien für generische Subgruppenmerkmale (z. B. „Männer vs. Frauen“, „ $\leq 64$  vs.  $65-74$  vs.  $\geq 75$  Jahre“) bzw. den krankheitsspezifischen Grenzen gemäß Leitlinienempfehlungen und einschlägiger Literatur (z. B. „sehr niedriges vs. niedriges vs. intermediäres IPSS-R“).

### **Analysierte Subgruppen**

Für das vorliegende Nutzendossier wurden die folgenden Subgruppen und dazugehörigen Kategorien analysiert. Bei quantitativen Merkmalen sind zudem die zugrundeliegenden Rationale für die Wahl der Trennpunkte angegeben.

- Alter  
( $\leq 64$  Jahre,  $65 - 74$  Jahre,  $\geq 75$  Jahre)
- Geschlecht  
(männlich, weiblich)
- WHO-Klassifizierung (2008) der MDS zur Baseline  
(Refraktäre Anämie mit Ringsideroblasten (*refractory anaemia with ring sideroblasts*, RARS), Refraktäre Zytopenie mit multilineären Dysplasien (*refractory cytopenia with multi-lineage dysplasia*, RCMD))
- Zeit seit initialer MDS-Diagnose zur Baseline  
( $\leq 2$  Jahre, 2 bis 5 Jahre,  $> 5$  Jahre)

Gemäß Risikoklassifizierung nach IPSS-R entsprechen die beiden Grenzwerte zwei und fünf Jahre näherungsweise der Grenze der medianen Überlebenszeit zwischen Niedrigrisiko- und Hochrisiko-MDS-Patienten (zwischen 1,6 und drei Jahre) sowie der medianen Überlebenszeit von Patienten mit niedrigem IPSS-R-Score (5,3 Jahre) [37]

- ECOG Performance-Status  
(0 – 1, 2)
- Transfusionslast zur Baseline I  
( $< 6$  EK-Einheiten pro acht Wochen,  $\geq 6$  EK-Einheiten pro acht Wochen in den 16 Wochen vor Behandlungsbeginn)

Der Grenzwert von sechs EK-Einheiten pro acht Wochen entspricht der Transfusionslast, bei der eine Reduktion der Transfusionslast um  $\geq 4$  EK-Einheiten pro acht Wochen (HI-E gemäß IWG-Kriterien 2006) zu einer Transfusionsunabhängigkeit führt ( $< 2$  EK-Einheiten pro acht Wochen; siehe Einschlusskriterien MEDALIST).

- Transfusionslast zur Baseline II  
( $< 4$  EK-Einheiten pro acht Wochen,  $\geq 4$  bis  $< 6$  EK-Einheiten pro acht Wochen,  $\geq 6$  EK-Einheiten pro acht Wochen in den 16 Wochen vor Behandlungsbeginn)

Gemäß IWG-Kriterien (2006) wird eine Transfusionslast von vier EK-Einheiten pro acht Wochen als Cut-off für die Wahl des geeigneten Ansprechkriteriums im Rahmen eines erythroiden Ansprechens (HI-E) herangezogen [38]. Die weiteren Grenzwerte ergeben sich analog zur Transfusionslast zur Baseline I.

- Risikogruppe gemäß IPSS-R zur Baseline  
(sehr niedrig, niedrig, intermediär)
- Längste Dauer einer vorherigen ESA-Behandlung  
( $< 6$  Monate, 6 bis 12 Monate,  $> 12 - 24$  Monate,  $> 24$  Monate)

Die Dauer einer ESA-Therapie liegt im Median bei ca. 12 bzw. ca. 24 Monaten [69, 70], dementsprechend stellen die gewählten Grenzwerte die mediane bzw. halbe mediane Dauer einer typischen ESA-Therapie dar.

- Nierenfunktionsstatus (gemessen anhand der Kreatinin-Clearance)  
( $> 40 - 60$  ml/min,  $\geq 60$  ml/min)

Die vorgenommene Kategorisierung folgt der Einstufung der Nierenfunktion in einen Funktionsverlust ( $> 40 - 60$  ml/min) und keinen/geringgradigen Funktionsverlust ( $\geq 60$  ml/min).

- SF3B1-Mutationsstatus  
(mutiert, nicht mutiert)

- Serum-EPO-Spiegel zur Baseline  
( $< 100 \text{ U/l}$ ,  $100 - < 200 \text{ U/l}$ ,  $200 - 500 \text{ U/l}$ ,  $> 500 \text{ U/l}$ )

Im Rahmen des Nordic-Scores fließt der Serum-EPO-Spiegel entsprechend der Grenzen  $< 100 \text{ U/l}$  und  $> 500 \text{ U/l}$  in die Abschätzung des Therapieerfolgs ein [71]. Die ergänzende Unterteilung anhand des Grenzwerts von  $200 \text{ U/l}$  entspricht der Angabe in der Zulassung von ESA [72, 73].

- Thrombozytenzahl zur Baseline  
( $< 100 \times 10^9/l$ ,  $100 - 400 \times 10^9/l$ ,  $> 400 \times 10^9/l$ )

Thrombozytenzahlen  $< 100 \times 10^9/l$  entsprechen einer Thrombozytopenie, welche auch als Risikofaktor in die Berechnung des IPSS-R-Scores einfließt [37] und bei der Bewertung einer vollständigen Remission berücksichtigt wird [38]. Thrombozytenzahlen im Bereich von  $100 - 400 \times 10^9/l$  entsprechen dem normalen Referenzbereich, Thrombozytenzahlen  $> 400 \times 10^9/l$  entsprechen einer Thrombocytose (Erhöhung der Thrombozyten).

- Region  
(Nordamerika, Europa)

### Methodik von Subgruppenanalysen

Für alle Endpunkte wurde in einem ersten Schritt der jeweilige Interaktionstest für die definierten Subgruppenmerkmale hinsichtlich der Behandlung berechnet und tabellarisch dargestellt. Für alle Berechnungen wurden Analysen ohne Stratifizierungsfaktoren verwendet. Für Subgruppenmerkmale, die die folgenden Kriterien nicht erfüllten, wurden keine Interaktions-p-Werte berechnet, da aufgrund der geringen Patientenzahlen bzw. Ereignishäufigkeiten in den Subgruppen eine Interpretation der Ergebnisse nicht sinnvoll möglich gewesen wäre [35]:

- mindestens zehn Patienten in mindestens zwei Subgruppen eines Merkmals, bzw.
- für binäre Endpunkte: mindestens zehn Ereignisse in mindestens einer Subgruppe eines Merkmals.

Subgruppen mit einem Interaktions-p-Wert  $\geq 0,05$  werden als homogen hinsichtlich des betrachteten Behandlungseffekts beurteilt. Die Ergebnisse werden ergänzend in Anhang 4-G dargestellt, die Bewertung des Effekts erfolgt ausschließlich basierend auf den Ergebnissen der Gesamtpopulation.

Bei Vorliegen eines statistisch signifikanten Interaktionstests (Interaktions-p-Wert < 0,05) werden die Ergebnisse der jeweiligen Subgruppe detailliert in Abschnitt 4.3.1.3.2 dargestellt. Gemäß IQWiG-Methodenpapier impliziert ein statistisch signifikanter Interaktionstest das Vorliegen einer Effektmodifikation, d. h. ein zwischen den Subgruppen unterschiedlicher Effekt der Behandlung [35]. Bei Testung einer größeren Anzahl statistischer Hypothesen, insbesondere bei solchen ohne Adjustierung für multiples Testen, ist jedoch zu beachten, dass es bereits durch die angenommene Irrtumswahrscheinlichkeit, also per Definition bei 5 % der durchgeführten Tests, zu einem falsch positiven Ergebnis kommen kann. Bei den im vorliegenden Fall durchgeführten 602 Subgruppenanalysen (43 Hauptanalysen [zu acht Endpunkten] mit jeweils 14 Subgruppen) muss rein aus statistischer Sicht bereits mit ca. 30 falsch positiven Ergebnissen gerechnet werden. Aus einem positiven Interaktionstest kann demnach nicht direkt auf eine Effektmodifikation geschlossen werden, sondern muss vielmehr eine Plausibilisierung einer möglichen Effektmodifikation erfolgen (siehe auch Abbildung 4-1).

In einem ersten Schritt wird daher die Lage der Effektschätzer und deren 95 %-KI bezogen auf die einzelnen Subgruppenmerkmale und die Gesamtpopulation bewertet. Sofern die Lage der Effektschätzer und zugehörigen 95 %-KI zwischen den Kategorien des Subgruppenmerkmals und dem Effekt bezogen auf die Gesamtpopulation nicht relevant voneinander abweichen, wird auch nicht von einer relevanten Effektmodifikation ausgegangen und eine abweichende Bewertung des Zusatznutzens wird nicht durchgeführt.

Bei Effektschätzern und zugehörigen 95 %-KI, die bezüglich eines Subgruppenmerkmals relevant voneinander abweichen, d. h. eine Inkonsistenz der Ergebnisse beobachtet wird, wird in Anlehnung an die Vorgaben der *Guideline on the investigation of subgroups in confirmatory clinical trials* der EMA die Belastbarkeit dieser Ergebnisse anhand der folgenden Kriterien geprüft [67]:

- Vorliegen einer medizinischen Rationale für den beobachteten inkonsistenten Effekt
- Wiederholung der Inkonsistenz, u. a. über verschiedene Endpunkte hinweg

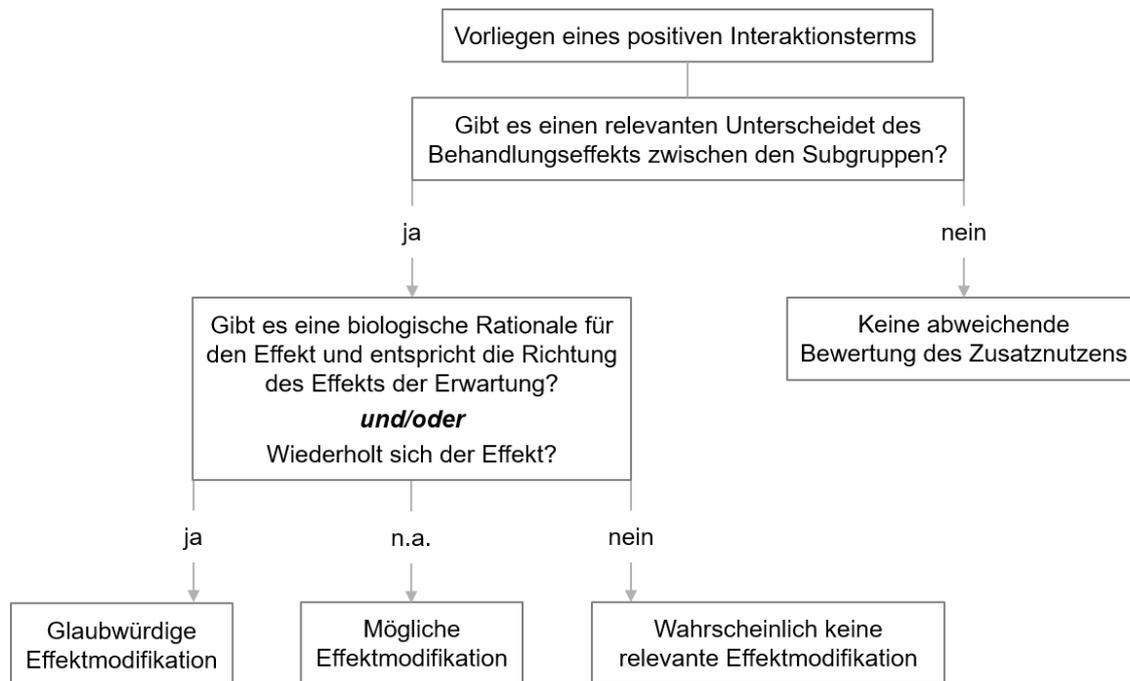


Abbildung in Anlehnung an den Entscheidungsbaum der EMA-Guideline Annex 1 [67]; n. a. = nicht abschätzbar

Abbildung 4-1: Plausibilisierung möglicher Effektmodifikationen in MEDALIST

Bei der Interpretation der Ergebnisse auf Ebene der Subgruppenmerkmale wird dementsprechend neben dem Ergebnis des Interaktionstests auch die beschriebene Unsicherheit durch multiples Testen sowie ein Zusammenhang mit einer möglichen klinischen Rationale und die Konsistenz des beobachteten Effekts über verschiedene Endpunkte hinweg berücksichtigt.

#### 4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Es besteht dabei internationaler Konsens, dass Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator (häufig als nicht adjustierte indirekte Vergleiche bezeichnet) regelhaft keine valide Analysemethoden darstellen<sup>9</sup>. Eine Ausnahme kann das Vorliegen von dramatischen Effekten sein. An Stelle von nicht adjustierten indirekten Vergleichen sollen je nach Datenlage einfache adjustierte indirekte Vergleiche<sup>10</sup> oder komplexere Netzwerk-Meta-Analysen (auch als „Mixed Treatment Comparison [MTC] Meta-Analysen“ oder „Multiple Treatment Meta-Analysen“ bezeichnet) für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche berechnet werden. Aktuelle Verfahren wurden beispielsweise von Lu und Ades (2004)<sup>11</sup> und Rücker (2012)<sup>12</sup> vorgestellt.

Alle Verfahren für indirekte Vergleiche gehen im Prinzip von den gleichen zentralen Annahmen aus. Hierbei handelt es sich um die Annahmen der Ähnlichkeit der eingeschlossenen Studien, der Homogenität der paarweisen Vergleiche und der Konsistenz zwischen direkter und indirekter Evidenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist<sup>13</sup>.

Das Ergebnis eines indirekten Vergleichs kann maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen. Als Brückenkomparatoren sind dabei insbesondere Interventionen zu berücksichtigen, für die sowohl zum bewertenden Arzneimittel als auch zur zweckmäßigen Vergleichstherapie mindestens eine direkt vergleichende Studie vorliegt (Brückenkomparatoren ersten Grades).

---

<sup>9</sup> Bender R, Schwenke C, Schmoor C, Hauschke D. Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR [online]. [Zugriff: 31.10.2016]. URL: [http://www.gmds.de/pdf/publikationen/stellungnahmen/120202\\_IQWIG\\_GMDS\\_IBS\\_DR.pdf](http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf).

<sup>10</sup> Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50(6): 683-691.

<sup>11</sup> Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

<sup>12</sup> Rücker G. Network meta-analysis, electrical networks and graph theory. *Res Synth Methods* 2012; 3(4): 312-324.

<sup>13</sup> Schöttker B, Lühmann D, Boukhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

Insgesamt ist es notwendig, die zugrunde liegende Methodik für alle relevanten Endpunkte genau und reproduzierbar zu beschreiben und die zentralen Annahmen zu untersuchen<sup>14, 15, 16</sup>

*Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:*

- *Benennung aller potentiellen Brückenkomparatoren ersten Grades und ggf. Begründung für die Auswahl.*
- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayes'schen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten, die Art der Untersuchung der Konvergenz der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Ähnlichkeit der eingeschlossenen Studien.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenzannahme im Netzwerk.*
- *Bilden Sie den Code des Computerprogramms inklusive der einzulesenden Daten in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Ein indirekter Vergleich wurde nicht durchgeführt.

---

<sup>14</sup> Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

<sup>15</sup> Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study *BMJ* 2011; 343 :d4909

<sup>16</sup> Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. *Res Synth Methods* 2013; 4(4): 291-323.

### 4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2)
- Weitere Untersuchungen (Abschnitt 4.3.2.3)

Falls für die Bewertung des Zusatznutzens mehrere Komparatoren (z.B. Wirkstoffe) herangezogen werden, sind die Aussagen zum Zusatznutzen primär gegenüber der Gesamtheit der gewählten Komparatoren durchzuführen (z. B. basierend auf Meta-Analysen unter gemeinsamer Betrachtung aller direkt vergleichender Studien). Spezifische methodische Argumente, die gegen eine gemeinsame Analyse sprechen (z. B. statistische oder inhaltliche Heterogenität), sind davon unbenommen. Eine zusammenfassende Aussage zum Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie ist in jedem Fall erforderlich.

#### 4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

##### 4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

###### 4.3.1.1.1 Studien des pharmazeutischen Unternehmers

*Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:*

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*

- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

*Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer, Angabe zu geplanten und durchgeführten Datenschnitten und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.*

Tabelle 4-7: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer ggf. Datenschnitt	Therapiearme
MEDALIST	ja	ja	laufend	02/2016 – voraussichtlich 11/2020 1. Datenschnitt: 08.05.2018 2. Datenschnitt: 07.01.2019 3. Datenschnitt: 01.07.2019 Finaler Datenschnitt: voraussichtlich 11/2020	a) Luspatercept + BSC b) Placebo + BSC
COMMANDS	nein	ja	laufend	01/2019 – voraussichtlich 04/2021	a) Luspatercept b) Epoetin-alfa
BSC = Best-Supportive-Care					

*Geben Sie an, welchen Stand die Information in Tabelle 4-7 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.*

Stand der Informationen: 18.05.2020

*Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-7 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-8: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
COMMANDS	Nur ein Teil der Studienpopulation entspricht E1; Laufende, derzeit rekrutierende Studie; Es liegen noch keine Ergebnisse vor

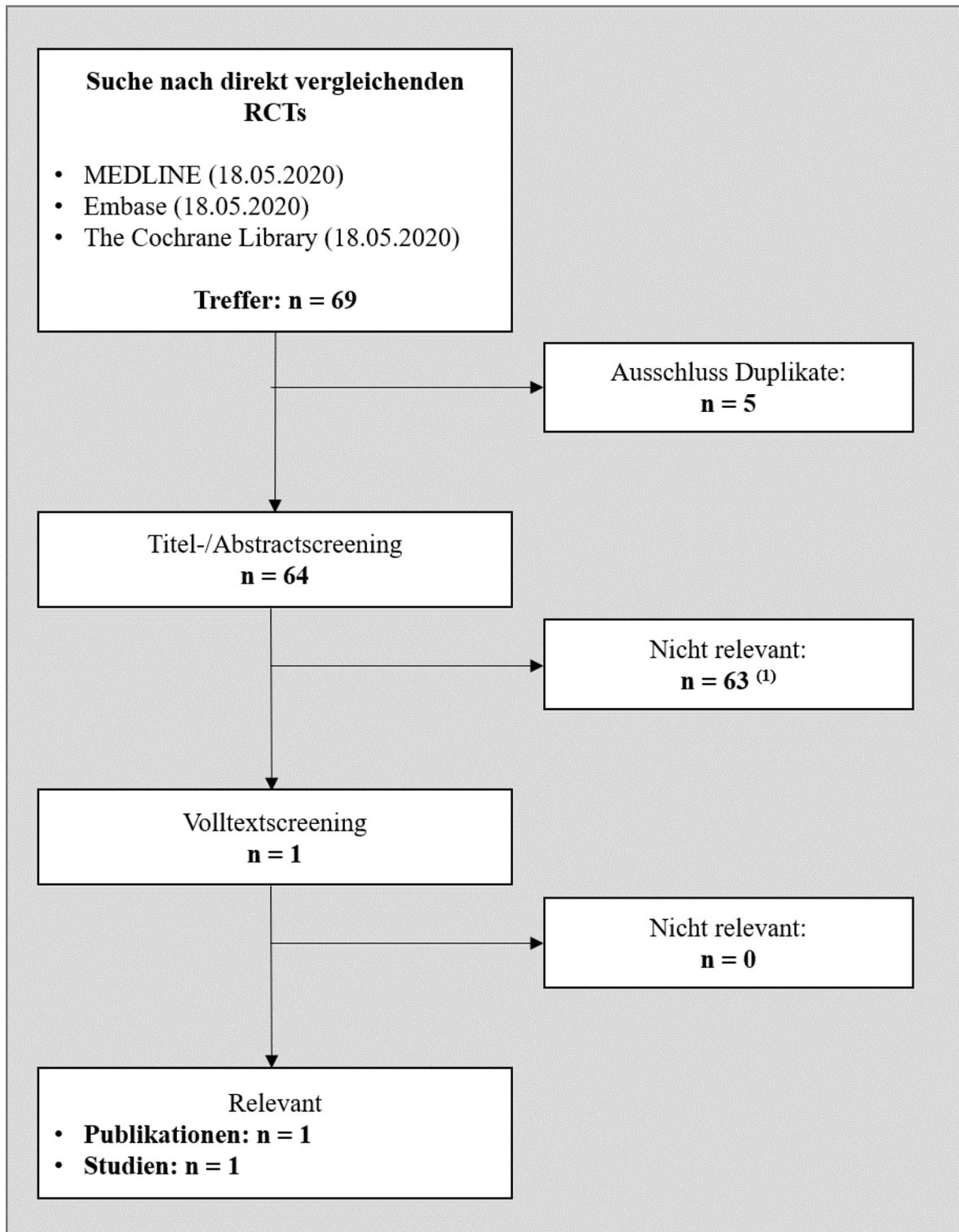
#### 4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

*Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.*

*[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]*

*Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.*

*Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.*



<sup>(1)</sup> Studienregistereinträge aus anderweitig durchsuchten Registern (d. h. clinicaltrials.gov, EUCTR oder WHO ICTRP), die im Rahmen der bibliografischen Literaturrecherche identifiziert wurden, wurden an dieser Stelle als nicht relevant ausgeschlossen

Abbildung 4-2: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Am 18.05.2020 wurde eine bibliografische Literaturrecherche in den Datenbanken MEDLINE, EMBASE und The Cochrane Library zur Identifizierung von relevanten RCTs mit dem zu bewertenden Arzneimittel Luspatercept durchgeführt. Zur Bewertung der Relevanz wurden die in Abschnitt 4.2.2 definierten Kriterien für den Ein- und Ausschluss von Studien herangezogen. Insgesamt ergab die Recherche 69 Treffer, von denen 64 nach Entfernung der Dubletten (n = 5) anhand der Titel und Abstracts gesichtet wurden. Publikationen, deren Relevanz anhand der Titel und/oder Abstracts nicht eindeutig erkennbar war, wurden anschließend anhand der Volltexte bewertet. Nicht relevante Publikationen wurden ausgeschlossen (n = 63) nach Titel-/Abstractscreening, n = 0 nach Volltextscreening), sodass eine relevante Publikation verblieb, welche sich auf eine relevante Einzelstudie bezieht.

#### **4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken**

*Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-7) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.*

*[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]*

*Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.*

Tabelle 4-9: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank und Angabe der Zitate <sup>a</sup> )	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Status (abgeschlossen/ abgebrochen/ laufend)
MEDALIST	<a href="https://clinicaltrials.gov/ct2/show/study/NCT02631070">clinicaltrials.gov:</a> NCT02631070 [Studienregistereintrag] [74] <a href="https://eudract.europa.eu/number/2015-003454-41">EUCTR:</a> 2015-003454-41 [Studienregistereintrag] [75] <a href="https://www.who.int/clinical-trials/ctcp/2015-003454-41">ICTRP:</a> 2015-003454-41 [Studienregistereintrag] [76], NCT02631070 [Studienregistereintrag] [77] <a href="https://pharmnet-bund.de/">pharmnet-bund.de</a> <sup>(1)</sup> : 2015-003454-41 [Studienregistereintrag] [78]	ja	ja [79]	laufend
COM-MANDS	<a href="https://clinicaltrials.gov/ct2/show/study/NCT03682536">clinicaltrials.gov:</a> NCT03682536 [Studienregistereintrag] [80] <a href="https://eudract.europa.eu/number/2017-003190-34">EUCTR:</a> 2017-003190-34 [Studienregistereintrag] [81] <a href="https://www.who.int/clinical-trials/ctcp/2017-003190-34">ICTRP:</a> 2017-003190-34 [Studienregistereintrag] [82, 83] <a href="https://pharmnet-bund.de/">pharmnet-bund.de</a> <sup>(1)</sup> : 2017-003190-34 [Studienregistereintrag] [84]	ja	nein	laufend
a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder -ergebnisse. <sup>(1)</sup> Suche im Portal für Klinische Prüfungen; Suche im Arzneimittelinformationssystem (AMIS) blieb ohne Treffer				

*Geben Sie an, welchen Stand die Information in Tabelle 4-9 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.*

Stand der Informationen: 18.05.2020

#### 4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-7) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-10: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	Relevante Quellen <sup>a</sup>	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Studie durch Suche in Studienregistern /Studienergebnisdatenbanken identifiziert (ja/nein)
Es wurde keine relevante Studie auf der Internetseite des G-BA identifiziert.	k. A.	k. A.	k. A.	k. A.
a: Quellen aus der Suche auf der Internetseite des G-BA k: A. = keine Angabe				

Geben Sie an, welchen Stand die Information in Tabelle 4-10 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Informationen: 18.05.2020

**4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel**

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2, 4.3.1.1.3 und 0) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich aller verfügbaren Quellen in Abschnitt 4.3.1.4 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Quellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-11: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienberichte (ja/nein [Zitat])	Register-einträge <sup>c</sup> (ja/nein [Zitat])	Publikation und sonstige Quellen <sup>d</sup> (ja/nein [Zitat])
<b>placebokontrolliert</b>						
MEDA-LIST	ja	ja	nein	ja [1]	ja [74-78]	ja [79]
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (Liste der eingeschlossenen Studien) aufgeführt werden.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>d: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.</p>						

### **4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel**

#### **4.3.1.2.1 Studiendesign und Studienpopulationen**

*Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Falls Teilpopulationen berücksichtigt werden, ist die Charakterisierung der Studienpopulation auch für diese Teilpopulation durchzuführen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Geben Sie bei den Datenschnitten auch den Anlass des Datenschnittes an. Fügen Sie für jede Studie eine neue Zeile ein.*

*Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.*

Tabelle 4-12: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach, verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
MEDALIST	RCT, doppelblind, parallel	Erwachsene, Anämie aufgrund von MDS mit Ringsideroblasten mit sehr niedrigem, niedrigem oder intermediärem Risiko, auf Transfusionen mit EK angewiesen, ESA bereits erhalten oder für ESA nicht geeignet	Luspatercept + BSC <sup>(1)</sup> (n = 153) Placebo + BSC <sup>(1)</sup> (n = 76)	<p><u>Screening:</u> fünf Wochen vor Randomisierung<sup>(2)</sup></p> <p><u>Primäre Behandlungsphase:</u> 24 Wochen</p> <p><u>Extensionsphase:</u> ab Woche 25 bis zum Therapieabbruch<sup>(3)</sup></p> <p><u>Posttreatment-Follow-up-Phase:</u> drei Jahre nach letzter Gabe der Studienmedikation</p> <p><u>Futility-Analyse:</u> 27.06.2017 (geplante Interims-Futility-Analyse)<sup>(4)</sup></p> <p><u>1. Datenschnitt:</u> 08.05.2018 (geplante Analyse)</p> <p><u>2. Datenschnitt:</u> 07.01.2019 (Nachbeobachtungsanalyse zur Unterstützung der Beantwortung einer Rückfrage der EMA)</p>	Europa (Belgien, Deutschland, Frankreich, Italien, Niederlande, Schweden, Spanien, Türkei, Vereinigtes Königreich), Nordamerika (Kanada, USA) 02/2016 – voraussichtlich 11/2020	<p><u>Primärer Endpunkt:</u> Transfusionsfreiheit</p> <p><u>Sekundäre Endpunkte:</u> Erythroides Ansprechen (HI-E), EK-Transfusionslast, Hb-Wert, sFerritin-Wert, Eisenchelationstherapie, Hospitalisierungen, gesundheitsbezogene Lebensqualität, unerwünschte Ereignisse, OS</p>

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie	Studiendesign <RCT, doppelblind/einfach, verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
				<p><u>3. Datenschnitt:</u> 01.07.2019 (Nachbeobachtungsanalyse zur Unterstützung der Beantwortung einer Rückfrage der EMA)</p> <p><u>Finaler Datenschnitt:</u> voraussichtlich 11/2020 (Nachbeobachtungsanalyse OS, Progression, Folgetherapien)</p>		
<p>(1) BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).</p> <p>(2) Eine historische Dokumentation der Transfusionsabhängigkeit der 16 Wochen vor Behandlungsbeginn (transfundierte EK-Einheiten, Prätransfusions-Hb-Wert) musste vorhanden sein.</p> <p>(3) Patienten, die die Kriterien für den Übergang in die Extensionsphase nach Vollendung der Visite zu Woche 25 erfüllen, verbleiben in doppelblinder Behandlung bis zum Therapieabbruch aufgrund nicht akzeptabler Toxizität, Progression der Erkrankung gemäß IWG-Kriterien (Veränderung des natürlichen Verlaufs der MDS), Rücknahme der Einverständniserklärung oder Erfüllens jeglichen anderen Abbruchkriteriums.</p> <p>(4) Geplante Interims-Analyse zur Feststellung der Sinnhaftigkeit der Studie (<i>futility analysis</i>). Datenschnitt bei ca. 105 Patienten, die die erste Behandlungsphase (Woche 1 – 24) beendet haben oder vor Vollendung dieser Phase die Studie abgebrochen haben (entspricht ca. 50 % der geplanten Ereignisse des primären Endpunkts). Auf eine Darstellung der Ergebnisse dieses Datenschnitts wird verzichtet, da eine vergleichende Analyse für diesen Datenschnitt nicht a priori geplant war, dieser Datenschnitt nicht zulassungsbegründend ist und die Ergebnisse des dargestellten finalen Datenschnitts vollumfänglicher sind und somit geeigneter sind, um eine Bewertung des medizinischen Nutzens und Zusatznutzens durchzuführen.</p> <p>BSC = Best-Supportive-Care; EK = Erythrozytenkonzentrat; EMA = <i>European Medicines Agency</i>; ESA = Erythropoese-stimulierende Agenzien (<i>erythropoiesis-stimulating agents</i>); Hb = Hämoglobin; HI-E = erythroides Ansprechen (<i>hematologic improvement-erythroid</i>); MDS = Myelodysplastische Syndrome; OS = Gesamtüberleben (<i>overall survival</i>); RCT = randomisierte kontrollierte Studie (<i>randomized controlled trial</i>); sFerritin = Serum-Ferritin; USA = Vereinigte Staaten von Amerika (<i>United States of America</i>)</p>						

Tabelle 4-13: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.
MEDALIST	1,0 – 1,75 mg/kg Körpergewicht Luspatercept s. c. alle drei Wochen	Volumenäquivalent Placebo zur Luspatercept-Gruppe s. c. alle drei Wochen	<u>Dosistitration</u> ab dem 1. Tag des 3. Behandlungszyklus unter Berücksichtigung definierter Kriterien (Transfusionslast, Dosis vorheriger Behandlungszyklen) auf 1,33 mg/kg Körpergewicht bzw. maximal 1,75 mg/kg Körpergewicht  In beiden Behandlungsarmen war eine supportive Behandlung (BSC) mit Antimykotika, Antibiotika, Virostatika und begleitenden Ernährungsmaßnahmen, sofern angezeigt, erlaubt <sup>(3)</sup>
<p><sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie.</p> <p><sup>(2)</sup> Während der Studie explizit ausgeschlossen waren zytotoxische, chemotherapeutische, gezielte Substanzen/ Therapien oder Prüfsubstanzen, Azacitidin, Decitabin oder andere HMAs, Lenalidomid, Thalidomid oder andere IMiDs, ESA oder andere erythroide hämatopoetische Wachstumsfaktoren, G- oder GM-CSF (außer zur Behandlung neutropenen Fiebers oder bei klinischer Indikation gemäß Zulassung), Hydroxyurea, Androgene (außer zur Behandlung des Hypogonadismus), orale Retinoide, Arsentrioxid und Interferon.</p> <p>BSC = Best-Supportive-Care; ESA = Erythropoese-stimulierende Agenzien; G-CSF = Granulozyten-Kolonie-stimulierender Faktor; GM-CSF = Granulozyten-Makrophagen-Kolonie-stimulierender Faktor; HMA = Hypomethylierende Agenzien; IMiD = Immunmodulierende Arzneimittel; s. c. = subkutan</p>			

Tabelle 4-14: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie MEDALIST (Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup> )	Luspatercept + BSC <sup>(1)</sup> (N = 153)	Placebo + BSC <sup>(1)</sup> (N = 76)
<b>Demographische Charakteristika</b>		
Alter [Jahre]		
Mittelwert (SD)	70,5 (8,68)	70,7 (10,88)
Median (Min; Max)	71,0 (40; 95)	72,0 (26; 91)
Altersgruppe [n (%)]		
≤ 64 Jahre	29 (19,0)	16 (21,1)
65 – 74 Jahre	72 (47,1)	29 (38,2)
≥ 75 Jahre	52 (34,0)	31 (40,8)
Geschlecht [n (%)]		
männlich	94 (61,4)	50 (65,8)
weiblich	59 (38,6)	26 (34,2)

<b>Studie MEDALIST (Luspatercept + BSC vs. Placebo + BSC<sup>(1)</sup>)</b>	<b>Luspatercept + BSC<sup>(1)</sup> (N = 153)</b>	<b>Placebo + BSC<sup>(1)</sup> (N = 76)</b>
<b>Ethnische Zugehörigkeit [n (%)]</b>		
schwarz oder afroamerikanisch	1 (0,7)	0 (0,0)
weiß	107 (69,9)	51 (67,1)
nicht erfasst	44 (28,8)	24 (31,6)
andere	1 (0,7)	1 (1,3)
<b>Größe [cm]</b>	<b>(n = 152)</b>	<b>(n = 75)</b>
Mittelwert (SD)	168,9 (10,29)	168,8 (8,85)
Median (Min; Max)	169,0 (142; 192)	169,0 (148; 188)
<b>Gewicht [kg]</b>		
Mittelwert (SD)	76,2 (15,07)	77,4 (15,78)
Median (Min; Max)	76,0 (46; 124)	75,0 (51; 153)
<b>Gewicht-Gruppe [n (%)]</b>		
< 70 kg	53 (34,6)	24 (31,6)
70 – < 85 kg	56 (36,6)	32 (42,1)
85 – < 100 kg	32 (20,9)	14 (18,4)
≥ 100 kg	12 (7,8)	6 (7,9)
<b>BMI [kg/m<sup>2</sup>]</b>	<b>(n = 152)</b>	<b>(n = 75)</b>
Mittelwert (SD)	26,6 (4,19)	27,0 (4,58)
Median (Min; Max)	26,2 (17; 40)	27,1 (20; 48)
<b>BMI-Gruppe [n (%)]</b>		
< 20 kg/m <sup>2</sup>	6 (3,9)	3 (3,9)
20 – < 25 kg/m <sup>2</sup>	49 (32,0)	24 (31,6)
25 – < 30 kg/m <sup>2</sup>	68 (44,4)	29 (38,2)
≥ 30 kg/m <sup>2</sup>	29 (19,0)	19 (25,0)
fehlend	1 (0,7)	1 (1,3)
<b><i>Krankheitsspezifische Charakteristika</i></b>		
<b>Zeit seit ursprünglicher MDS-Diagnose<sup>(2)</sup> [Monate]</b>		
Mittelwert (SD)	57,8 (56,59)	52,7 (42,29)
Median (Min; Max)	44,0 (3; 421)	36,1 (4; 193)
<b>Zeit seit ursprünglicher MDS-Diagnose<sup>(2)</sup> – Gruppen [n (%)]</b>		
≤ 2 Jahre	40 (26,1)	19 (25,0)
> 2 – 5 Jahre	62 (40,5)	34 (44,7)
> 5 Jahre	51 (33,3)	23 (30,3)
<b>Ringsideroblasten-Status [n (%)]</b>		
≥ 15 % der Gesamtzahl erythroider Zellen im KM	153 (100,0)	76 (100,0)

<b>Studie MEDALIST (Luspatercept + BSC vs. Placebo + BSC<sup>(1)</sup>)</b>	<b>Luspatercept + BSC<sup>(1)</sup> (N = 153)</b>	<b>Placebo + BSC<sup>(1)</sup> (N = 76)</b>
<b>WHO-Klassifikation (2008) der MDS [n (%)]</b>		
MDS RARS	7 (4,6)	2 (2,6)
MDS RCMD <sup>(3)</sup>	145 (94,8)	74 (97,4)
andere <sup>(4)</sup>	1 (0,7)	0 (0,0)
<b>IPSS-R Klassifizierung [n (%)]</b>		
sehr niedrig, niedrig	127 (83,0)	63 (82,9)
sehr niedrig	18 (11,8)	6 (7,9)
niedrig	109 (71,2)	57 (75,0)
intermediär	25 (16,3)	13 (17,1)
hoch	1 (0,7)	0 (0,0)
<b>Serum-EPO<sup>(5)</sup> [U/l]</b>		
Mittelwert (SD)	279,6 (361,33)	284,5 (433,84)
Median (Min; Max)	156,9 (12; 2454)	130,8 (29; 2760)
<b>Serum-Erythropoetin-Gruppe [n (%)]</b>		
< 100 U/l	51 (33,3)	31 (40,8)
100 – < 200 U/l	37 (24,2)	19 (25,0)
200 – 500 U/l	43 (28,1)	15 (19,7)
> 500 U/l	21 (13,7)	11 (14,5)
fehlend	1 (0,7)	0 (0,0)
<b>SF3B1-Mutationsstatus [n (%)]</b>		
mutiert	141 (92,2)	65 (85,5)
nicht mutiert	12 (7,8)	10 (13,2)
fehlend	0 (0,0)	1 (1,3)
<b>ECOG-Performance-Status [n (%)]</b>		
0	54 (35,3)	33 (43,4)
1	91 (59,5)	32 (42,1)
2	8 (5,2)	11 (14,5)
<b>Transfusionslast zur Baseline</b>		
<b>Transfundierte EK-Einheiten/ letzte acht Wochen<sup>(6)</sup></b>		
Mittelwert (SD)	5,9 (2,97)	6,2 (2,99)
Median (Min; Max)	6,0 (2; 16)	6,0 (0; 16)
<b>Transfundierte EK-Einheiten/ letzte acht Wochen<sup>(6)</sup> – Gruppen [n (%)]</b>		
≥ 6 Einheiten	78 (51,0)	46 (60,5)
< 6 Einheiten	75 (49,0)	30 (39,5)
≥ 4 – < 6 Einheiten	47 (30,7)	19 (25,0)
< 4 Einheiten	28 (18,3)	11 (14,5)

<b>Studie MEDALIST (Luspatercept + BSC vs. Placebo + BSC<sup>(1)</sup>)</b>	<b>Luspatercept + BSC<sup>(1)</sup> (N = 153)</b>	<b>Placebo + BSC<sup>(1)</sup> (N = 76)</b>
Transfundierte EK-Einheiten/ acht Wochen der letzten 16 Wochen <sup>(7)</sup>		
Mittelwert (SD)	5,5 (2,76)	5,8 (2,95)
Median (Min; Max)	5,0 (1; 15)	5,0 (2; 20)
Transfundierte EK-Einheiten/ acht Wochen der letzten 16 Wochen <sup>(7)</sup> – Gruppen [n (%)]		
≥ 6 Einheiten	66 (43,1)	33 (43,3)
< 6 Einheiten	87 (56,9)	43 (56,6)
≥ 4 – < 6 Einheiten	41 (26,8)	23 (30,3)
< 4 Einheiten	46 (30,1)	20 (26,3)
Hämoglobin <sup>(8)</sup> [g/dl]	(n = 153)	(n = 76)
Mittelwert (SD)	7,7 (0,84)	7,6 (0,77)
Median (Min; Max)	7,6 (6; 10)	7,6 (5; 9)
Vorherige ESA-Therapie [n (%)]		
ja	148 (96,7)	70 (92,1)
nein	5 (3,3)	6 (7,9)
Gründe für den Therapieabbruch der vorherigen ESA-Therapie <sup>(9)</sup> [n (%)]		
refraktär <sup>(10)</sup>	144 (97,3)	69 (98,6)
intolerant <sup>(11)</sup>	4 (2,7)	1 (1,4)
Zeit seit Beendigung der vorherigen ESA-Therapie bis Studienbeginn <sup>(12)</sup> [Monate]	(n = 148) <sup>1</sup>	(n = 70) <sup>1</sup>
Mittelwert (SD)	14,79 (28,824)	11,18 (13,553)
Median (Min; Max)	5,26 (0,9; 257,9)	5,13 (0,2; 64,9)
Zeit seit Beendigung der vorherigen ESA-Therapie bis Studienbeginn <sup>(12)</sup> – Gruppen [n (%)] <sup>(9)</sup>		
< 6 Monate	82 (55,4)	37 (52,9)
6 – 12 Monate	21 (14,2)	13 (18,6)
> 12 – 24 Monate	19 (12,8)	7 (10,0)
> 24 Monate	26 (17,6)	13 (18,6)
Längste Dauer einer vorherigen ESA-Therapie <sup>(9)</sup> [Monate]	(n = 148) <sup>(13)</sup>	(n = 70) <sup>(13)</sup>
Mittelwert (SD)	17,83 (22,415)	19,51 (20,202)
Median (Min; Max)	10,48 (1,2; 143,2)	13,17 (1,4; 90,9)

<b>Studie MEDALIST (Luspatercept + BSC vs. Placebo + BSC<sup>(1)</sup>)</b>	<b>Luspatercept + BSC<sup>(1)</sup> (N = 153)</b>	<b>Placebo + BSC<sup>(1)</sup> (N = 76)</b>
Längste Dauer einer vorherigen ESA-Therapie <sup>(9)</sup> – Gruppen [n (%)]		
< 6 Monate	48 (32,4)	20 (28,6)
6 – 12 Monate	34 (23,0)	9 (12,9)
> 12 – 24 Monate	35 (23,6)	21 (30,0)
> 24 Monate	31 (20,9)	20 (28,6)
Vorherige Therapie mit Eisenchelatoren [n (%)]		
ja	71 (46,4)	40 (52,6)
nein	82 (53,6)	36 (47,4)
Vorherige Therapie mit G-CSF/GM-CSF <sup>(14)</sup> [n (%)]		
ja	51 (33,3)	22 (28,9)
nein	102 (66,7)	54 (71,1)
sFerritin-Spiegel [µg/l]		
Mittelwert (SD)	1.348,0 (971,24)	1.503,8 (1.242,94)
Median (Min; Max)	1.089,2 (64; 5.968)	1.122,1 (165; 5.849)
Historischer sFerritin-Spiegel über 16 Wochen [µg/l]	(n = 101)	(n = 59)
Mittelwert (SD)	1.418,8 (988,17)	1.689,9 (1.317,37)
Median (Min; Max)	1.149,7 (84; 6083)	1.444,3 (144; 7229)
Nierenfunktion: Kreatinin- Clearance [ml/min]		
< 40	7 (4,6)	1 (1,3)
40 – < 60	33 (21,6)	12 (15,8)
≥ 60	113 (73,9)	62 (81,6)
fehlend	0 (0,0)	1 (1,3)
Thrombozytenzahl <sup>(15)</sup> [10 <sup>9</sup> /l]		
Mittelwert (SD)	259,3 (122,58)	251,7 (123,93)
Median (Min; Max)	235,0 (59; 892)	222,5 (60; 689)
Absolute Neutrophilenzahl (ANC) [10 <sup>9</sup> /l]		
Mittelwert (SD)	2,8 (2,06)	2,7 (2,03)
Median (Min; Max)	2,3 (0; 16)	2,2 (1; 12)
Absolute Neutrophilenzahl (ANC) – Gruppen [10 <sup>9</sup> /l]		
< 0,5	1 (0,7)	0 (0,0)
0,5 – < 1,0	14 (9,2)	10 (13,2)
≥ 1,0	138 (90,2)	66 (86,8)

<b>Studie MEDALIST (Luspatercept + BSC vs. Placebo + BSC<sup>(1)</sup>)</b>	<b>Luspatercept + BSC<sup>(1)</sup> (N = 153)</b>	<b>Placebo + BSC<sup>(1)</sup> (N = 76)</b>
Kardiovaskuläre Komorbiditäten (MACE) [n (%)]		
ja	47 (30,7)	23 (30,3)
nein	105 (68,6)	53 (69,7)
Komorbiditätsscore (Charlson Comorbidity Index)	(n = 152)	(n = 76)
Mittelwert (SD)	2,1 (1,62)	1,9 (1,58)
Median (Min; Max)	2,0 (0; 8)	2,0 (0; 6)
<p>Population: ITT-Population</p> <p><sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).</p> <p><sup>(2)</sup> Zeit seit ursprünglicher MDS-Diagnose war definiert als die Anzahl Jahre seit ursprünglicher Diagnose bis zum Datum der Einverständniserklärung.</p> <p><sup>(3)</sup> Alle Patienten waren als RCMD-RS klassifiziert, da das Vorliegen von Ringsideroblasten Einschlusskriterium war.</p> <p><sup>(4)</sup> Lokal diagnostizierte MDS-RS mit multilineärer Dysplasie.</p> <p><sup>(5)</sup> Baseline-EPO war definiert als der höchste EPO-Spiegel innerhalb der 35 Tage vor Gabe der ersten Dosis der Studienmedikation.</p> <p><sup>(6)</sup> Anzahl an EK-Einheiten in den letzten acht Wochen vor dem Datum der ersten Dosis der Studienmedikation.</p> <p><sup>(7)</sup> Anzahl an EK-Einheiten in den neun bis 16 Wochen (Acht-Wochen-Intervall) vor dem Datum der ersten Dosis der Studienmedikation.</p> <p><sup>(8)</sup> Hämoglobin zur Baseline war definiert als der letzte gemessene Wert vor oder am Tag der ersten Dosis der Studienmedikation. Nach Anwendung der 14/3-Tage-Regel, wurde der Baseline-Hämoglobin-Wert festgelegt als der niedrigste Wert, der innerhalb von 35 Tagen vor oder am Tag der ersten Dosis der Studienmedikation vom Zentrallabor gemessen wurde. Sofern kein qualifizierter Wert des Zentrallabors vorlag, wurde der aus demselben Zeitraum stammende Hämoglobin-Wert des lokalen Labors herangezogen. Falls weder ein Wert aus dem Zentrallabor noch aus einem lokalen Labor vorlag, wurde der niedrigste Wert vor Transfusion basierend auf den Transfusionsdokumentationen der 35 Tage vor Gabe der ersten Dosis der Studienmedikation benutzt.</p> <p><sup>(9)</sup> Prozentsatz berechnet bezogen auf die Anzahl Patienten mit vorheriger ESA-Therapie.</p> <p><sup>(10)</sup> Definiert als dokumentiertes Nicht-Ansprechen oder nicht anhaltendes Ansprechen auf ein vorheriges ESA-haltiges Therapieregime.</p> <p><sup>(11)</sup> Definiert als dokumentierter Therapieabbruch eines ESA-haltigen Therapieregimes zu jeglichem Zeitpunkt nach Therapiebeginn aufgrund einer Unverträglichkeit oder einem unerwünschten Ereignis.</p> <p><sup>(12)</sup> Zeit seit Beendigung der vorherigen ESA-Therapie bis Studienbeginn war definiert als die Anzahl Monate seit dem Datum des Therapieabbruchs der vorherigen ESA-Therapie bis zum Datum des ersten Tages des ersten Zyklus. Falls das Datum des ersten Tags des ersten Zyklus fehlte, wurde das Datum der Randomisierung herangezogen.</p> <p><sup>(13)</sup> Anzahl Patienten mit einer vorherigen ESA-Therapie.</p> <p><sup>(14)</sup> Jegliches Arzneimittel mit ATC-Code L oder L03 für G-CSF-/GM-CSF-Gebrauch.</p> <p><sup>(15)</sup> Thrombozytenzahl zur Baseline war definiert als der kleinste Wert innerhalb von 35 Tagen der ersten Dosis der Studienmedikation.</p> <p><b>14/3-Tage-Regel:</b> Es werden nur Hämoglobin-Werte, die mindestens 14 Tage nach einer Transfusion gemessen wurden, berücksichtigt, außer eine weitere Transfusion findet innerhalb von 3 Tagen nach der Hämoglobin-Messung statt. In diesem Fall ist auch dieser Hämoglobin-Wert zu berücksichtigen (trotz Messung innerhalb von &lt; 14 Tagen der ersten Transfusion).</p> <p>ANC = absolute Neutrophilenzahl (<i>absolute neutrophil count</i>); ATC = <i>Anatomical Therapeutic Chemical</i>; BMI = Body Mass Index; BSC = Best-Supportive-Care; EK = Erythrozytenkonzentrat; EPO = Erythropoetin; ESA = Erythropoese-stimulierende Agenzien (<i>erythropoiesis stimulating agent</i>); G-CSF = Granulozyten-Kolonie-stimulierender Faktor (<i>granulocyte colony-stimulating factor</i>); GM-CSF = Granulozyten-Makrophagen-Kolonie-stimulierender Faktor (<i>granulocyte-macrophage colony-stimulating factor</i>); IPSS-R = <i>International Prognostic Scoring System-revised</i>; ITT = Intention-to-Treat; KM = Knochenmark; MACE = <i>Major Adverse Cardiac Events</i>; Max = Maximum; MDS = Myelodysplastische Syndrome; Min = Minimum; SD = Standardabweichung (<i>standard deviation</i>); sFerritin = Serum-Ferritin; WHO = World Health Organization</p> <p>Quelle: [1]</p>		

*Beschreiben Sie die Studien zusammenfassend. In der Beschreibung der Studien sollten Informationen zur Behandlungsdauer sowie zu geplanter und tatsächlicher Beobachtungsdauer enthalten sein. Sofern sich die Beobachtungsdauer zwischen den relevanten Endpunkten unterscheidet, sind diese unterschiedlichen Beobachtungsdauern endpunktbezogen anzugeben. Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

*Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.*

Die Studie MEDALIST ist eine doppelblinde, randomisierte Phase-III-Studie, die die Wirksamkeit und Sicherheit von Luspatercept in erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, untersucht.

### **Patientenpopulation**

Die in MEDALIST eingeschlossenen Patienten wurden anhand generischer Merkmale, wie Alter, Geschlecht, ethnischer Zugehörigkeit, sowie krankheitsspezifischer Merkmale charakterisiert. Relevante Unterschiede zwischen den Behandlungsarmen konnten für keine der untersuchten Kenngrößen identifiziert werden. Wichtige Krankheitscharakteristika, die unter anderem sicherstellen, dass die hier vorgelegten Ergebnisse sich auf das vom Anwendungsgebiet von Luspatercept umfasste Patientenkollektiv beziehen, beinhalten insbesondere die MDS-Diagnose und Risikoklassifizierung, die Transfusionslast zur Baseline, sowie die Vortherapie mit ESA. Weitere untersuchte Merkmale dienen vornehmlich der Abschätzung der Schwere der Erkrankung und des Gesundheitsstatus des Patienten, wie beispielsweise die Kenngrößen zur Bestimmung weiterer Zytopenien, insbesondere der Neutro- und Thrombozytopenien, sowie anderer Komorbiditäten und dem Vorliegen einer Eisenüberladung. Insgesamt konnten keine systematischen Unterschiede zwischen den beiden Behandlungsarmen identifiziert werden, die Studienpopulation kann damit als ausreichend homogen angesehen werden (siehe Tabelle 4-14).

Entsprechend dem zugelassenen Anwendungsgebiet von Luspatercept ist die Patientenpopulation dieser Studie durch eine Reihe von Faktoren, die das Krankheitsbild, die Prognose als auch die Vortherapie der Patienten umfassen, anhand der Ein- und Ausschlusskriterien eingegrenzt. Dazu zählen die folgenden Charakteristika:

- MDS mit Ringsideroblasten gemäß WHO-Klassifikation
- MDS mit sehr niedrigem, niedrigem oder intermediärem Risiko gemäß IPSS-R
- Transfusionsabhängigkeit
- Vortherapie mit ESA mit nicht zufriedenstellendem Ansprechen oder fehlende Eignung für eine ESA-Therapie

### ***MDS mit Ringsideroblasten gemäß WHO-Klassifikation***

Gemäß den aktuellen WHO-Kriterien 2016 zur Klassifikation der MDS werden Patienten mit Ringsideroblasten durch das Vorliegen einer der beiden folgenden Eigenschaften charakterisiert [2-4, 85]:

- Anteil der Ringsideroblasten an der Gesamtzahl erythroider Zellen im Knochenmark  $\geq 15\%$ ;  $< 5\%$  Blasten im Knochenmark;  $< 1\%$  Blasten im peripheren Blut; keine Auerstäbchen; Ausschluss einer isolierten Del(5q)-Anomalie
- oder
- Anteil der Ringsideroblasten an der Gesamtzahl erythroider Zellen im Knochenmark  $\geq 5\%$ , wenn gleichzeitig eine SF3B1-Genmutation vorliegt;  $< 5\%$  Blasten im Knochenmark;  $< 1\%$  Blasten im peripheren Blut; keine Auerstäbchen; Ausschluss einer isolierten Del(5q)-Anomalie.

In MEDALIST wurden entsprechend der aktuellen Kriterien der WHO (2016) nur solche MDS-Patienten eingeschlossen, deren Anteil Ringsideroblasten an erythroiden Vorläuferzellen im Knochenmark  $\geq 15\%$  oder, bei Patienten mit einer SF3B1-Mutation,  $\geq 5\%$  war. Das Kriterium ist somit erfüllt.

***MDS mit sehr niedrigem, niedrigem oder intermediärem Risiko gemäß IPSS-R***

Neben der Klassifizierung der zugrundeliegenden Zytopenie(n) und vorliegenden Mutationen hat die Einstufung der Patienten in prognostische Risikogruppen eine besondere Bedeutung, da das Gesamtüberleben und das Risiko einer Progression in eine AML dadurch abgeschätzt werden können. Dafür stehen sowohl das IPSS (*International Prognostic Scoring System*) als auch die aktuellste, überarbeitete Version, das IPSS-R (*IPSS-revised*), als etablierte und validierte Systeme zur Verfügung [2, 37]. Das IPSS-R ist eine Weiterentwicklung des IPSS und zeichnet sich im Vergleich zu diesem durch eine erweiterte Anzahl zytogenetischer Kategorien, einer differenzierteren Gewichtung des Anteils an Knochenmarksblasten und einer detaillierteren Berücksichtigung von Zytopenien aus [2, 37, 86]. Aufgrund der exakteren Risikostratifizierung ist das IPSS-R dem IPSS vorzuziehen [4].

Im IPSS-R werden Patienten mit MDS in die fünf Risikogruppen sehr niedrig, niedrig, intermediär, hoch und sehr hoch eingeteilt. Für Therapieempfehlungen erfolgt in aller Regel eine Zusammenfassung der Patienten mit sehr niedrigem, niedrigem oder intermediärem Risiko gemäß IPSS-R als Niedrigrisiko-MDS-Patienten (entsprechend der Patienten mit niedrigem oder intermediär-1 Risiko nach IPSS) sowie der Patienten mit hohem und sehr hohem Risiko als Hochrisiko-MDS-Patienten (entsprechend der Patienten mit intermediär-2 und hohem Risiko gemäß IPSS) [2-4].

Der Einschluss von Patienten in MEDALIST mit sehr niedrigem, niedrigem oder intermediärem IPSS-R-Risiko entspricht damit der Zulassungspopulation.

***Transfusionsabhängigkeit***

Eine symptomatische Anämie aufgrund von Niedrigrisiko-MDS wird gemäß Leitlinienempfehlungen mit einer supportiven, bedarfsgerechten Transfusionstherapie mit EK behandelt (vgl. auch Modul 3 B) [2, 4]. Infolge der progressiven Natur der MDS steigt die Transfusionslast der Patienten im Verlauf der Erkrankung stetig an. Auswertungen des EUMDS (*European LeukaemiaNet MDS*)-Registers zum Transfusionsbedarf bei Niedrigrisiko-MDS-Patienten zeigen, dass auch Patienten, die bei initialer Diagnosestellung nicht transfusionsabhängig sind, innerhalb eines Jahres transfusionsabhängig werden [87].

Entsprechend dem geplanten Anwendungsgebiet sollen nur solche Patienten mit Luspatercept behandelt werden, die transfusionsabhängig sind, d. h. regelmäßig mit EK-Transfusionen behandelt werden [5]. Diesem Umstand wurde in MEDALIST Rechnung getragen, indem nur solche Patienten in die Studie eingeschlossen wurden, die über einen Zeitraum von mindestens 16 Wochen vor Behandlungsbeginn regelmäßig EK-Transfusionen benötigten. Eingeschlossene Patienten durften in diesen 16 Wochen kein transfusionsfreies Intervall von  $\geq 56$  Tagen aufweisen. Einmalig und somit nicht regelmäßig transfundierte Patienten wurden aus der Studie ausgeschlossen, sodass sichergestellt ist, dass nur regelmäßig und dauerhaft transfundierte und somit transfusionsabhängige Patienten in die Studie aufgenommen wurden. Angaben zur Transfusionshistorie finden sich in Tabelle 4-14.

***Vorthherapie mit ESA mit nicht zufriedenstellendem Ansprechen oder fehlende Eignung für eine ESA-Therapie***

Gemäß Anwendungsgebiet kommen nur solche Patienten für eine Therapie mit Luspatercept in Frage, die auf eine EPO-basierte Therapie, d. h. Therapie mit ESA, nicht zufriedenstellend angesprochen haben oder für eine solche Therapie nicht geeignet sind.

Der erste Teil dieser Bedingung – bereits mit ESA behandelte Patienten mit nicht zufriedenstellendem Ansprechen – definiert Patienten mit bereits abgeschlossener ESA-Behandlung. Dementsprechend waren diese Patienten im Vorfeld der Luspatercept-Behandlung für eine ESA-Therapie gemäß Nordic-Score und ESA-Zulassung (Serum-EPO < 200 U/l) geeignet und wurden auch damit behandelt, mussten diese Behandlung aber aufgrund eines unzureichenden Ansprechens, eines Verlusts des Ansprechens oder einer Unverträglichkeit wieder abbrechen (nicht zufriedenstellendes Ansprechen auf ESA). Initial sprechen rund 30 – 50 % der Patienten, die für eine ESA-Therapie geeignet sind, auf diese Behandlung an, ca. 70 % dieser Patienten werden innerhalb von ein bis zwei Jahren jedoch therapierefraktär, sodass eine Weiterbehandlung mit ESA nicht indiziert ist [88, 89].

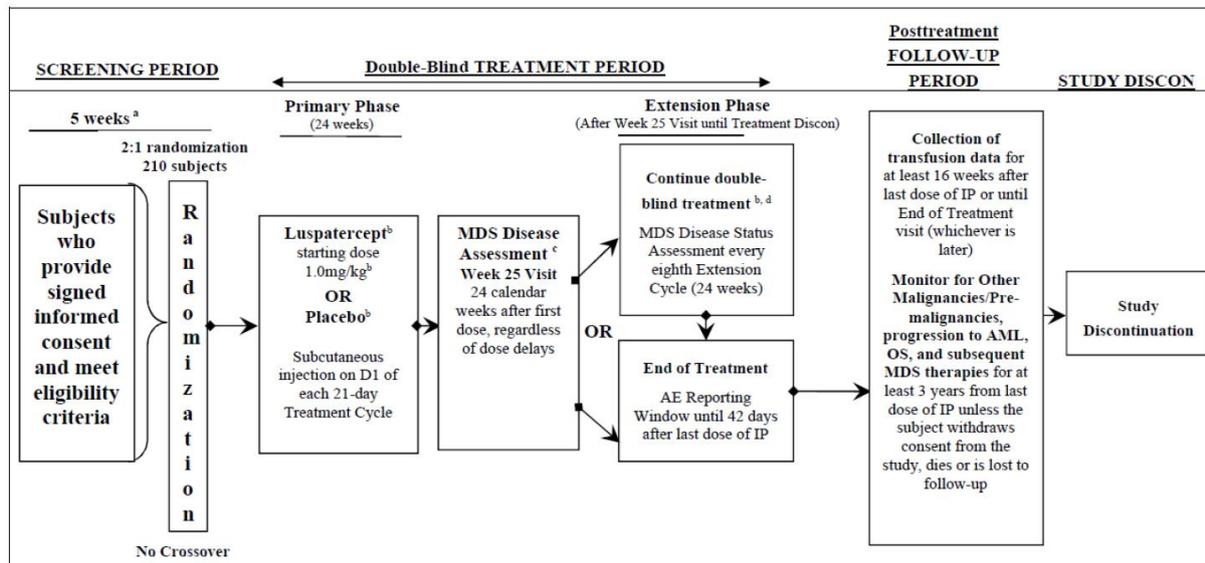
Zudem können auch solche Patienten, die nicht für eine ESA-Therapie geeignet sind – entsprechend dem zweiten Teil der Bedingung – mit Luspatercept behandelt werden. Mögliche Gründe hierfür sind bekannte Unverträglichkeiten, Kontraindikationen, fehlende Eignung gemäß Zulassung (Serum-EPO  $\geq$  200 U/l) oder auch eine geringe Ansprechprognose gemäß Nordic-Score [2, 90].

In MEDALIST wurden nur solche Patienten eingeschlossen, deren Therapieversagen einer ESA-Behandlung dokumentiert war (nicht zufriedenstellendes Ansprechen auf ESA) oder die für eine ESA-Therapie nicht geeignet sind.

Die Ein- und Ausschlusskriterien in MEDALIST stellen somit eine adäquate Eingrenzung des von der Zulassung umfassten Patientenkollektivs hinsichtlich der ESA-Vorbehandlung und -Eignung sicher.

## Studiendesign

Die Studie MEDALIST wurde an mehreren Zentren in nordamerikanischen und elf europäischen Ländern, darunter auch Deutschland, durchgeführt. Die Studie war unterteilt in eine Screening- und eine Behandlungsphase, an die sich entweder eine Extensions- und/oder *Post-treatment-Follow-up-Phase* anschloss (siehe Abbildung 4-3).



<sup>a</sup> Historische Dokumentation der Transfusionsabhängigkeit sollte für mindestens 16 Wochen vor Randomisierung verfügbar sein (Anzahl transfundierte EK-Einheiten, Prätransfusions-Hb).

<sup>b</sup> Dosistitration bis maximal 1,75 mg/kg war erlaubt.

<sup>c</sup> Patienten, die nach Ermessen des Prüfarztes zu Woche 25 von der Behandlung profitiert haben, durften im Rahmen der doppelblinden Extensionsphase weiter behandelt werden bis Kriterien zum Therapieabbruch erfüllt wurden.

<sup>d</sup> Eine Bewertung der Erkrankung entsprechend der Kriterien zur Woche 25-Visite musste zu Tag 1 von jedem 8. Behandlungszyklus durchgeführt werden.

AE = unerwünschtes Ereignis (*adverse event*); AML = akute myeloische Leukämie; D1 = Tag 1 (*day 1*); EK = Erythrozytenkonzentrat; Hb = Hämoglobin; IP = Prüfsubstanz (*investigational product*); MDS = Myelodysplastische Syndrome; OS = Gesamtüberleben (*overall survival*)

Abbildung 4-3: Übersicht des Studiendesigns der MEDALIST-Studie [1]

### Screening-Phase

Während einer initialen Screening-Phase wurde die Eignung der Patienten für die Studienteilnahme geprüft. Diese Prüfung umfasste unter anderem die Diagnose und Klassifizierung der MDS gemäß Einschlusskriterien sowie die Erhebung der Transfusionslast in den fünf Wochen vor Randomisierung. Darüber hinaus musste eine Transfusionshistorie über einen Zeitraum von mindestens 16 Wochen unmittelbar vor Randomisierung vorliegen. Transfusionslast und -historie mussten die Art der Transfusion (z. B. EK, Thrombozytenkonzentrat), Anzahl an verabreichten Einheiten und das Datum der Transfusionen enthalten. Darüber hinaus sollte der Hb-Wert vor Verabreichung einer EK-Transfusion dokumentiert werden (Prätransfusions-Hb).

***Randomisierung und primäre Behandlungsphase***

Die Patienten wurden im Verhältnis 2:1 in den Luspatercept- (n = 153) oder den Placebo-Arm (n = 76) randomisiert. Die Randomisierung erfolgte stratifiziert nach Transfusionslast und IPSS-R-Score zur Baseline.

Innerhalb von drei Tagen nach Randomisierung wurde die erste Dosis der Studienmedikation verabreicht. Im Luspatercept-Arm wurde Luspatercept zu einer anfänglichen Dosierung von 1,0 mg/kg Körpergewicht alle drei Wochen als subkutane Injektion verabreicht, welche auf maximal 1,75 mg/kg Körpergewicht auftitriert werden durfte. Patienten im Placebo-Arm erhielten eine subkutane Placebo-Injektion (Natriumchlorid-Lösung, 0,9 %), die analog zur Luspatercept-Behandlung alle drei Wochen verabreicht wurde, um die Verblindung in der Studie aufrecht zu erhalten. In beiden Behandlungsarmen war im Rahmen der BSC eine regelmäßige, bedarfsgerechte Transfusionstherapie in Kombination mit Eisenchelatherapie erlaubt, welche dem Behandlungsstandard in der vorliegenden Indikation entspricht (nähere Angaben zur Umsetzung von BSC in MEDALIST siehe Tabelle 4-13).

Die Gabe der Studienmedikation sollte mindestens über den Zeitraum der ersten 24 Wochen erfolgen, es sei denn, es traten nicht akzeptable Toxizitäten auf, der Patient zog die Einverständniserklärung zurück oder ein anderes Kriterium für den Studienabbruch war erfüllt.

***Extensionsphase***

Zu Woche 25 wurde durch den Prüfarzt beurteilt, ob der Patient innerhalb der ersten 24 Wochen von der Behandlung profitiert hatte. Die Beurteilung erfolgte anhand der folgenden Kriterien:

- Nachweis eines klinischen Nutzens (z. B. Reduktion der benötigten EK-Transfusionen oder Erhöhung des Hb-Werts im Vergleich zur Baseline) und
- fehlende Progression der Erkrankung gemäß IWG-Kriterien [38].

Patienten, die diese beiden Kriterien nach Ermessen des Prüfarztes erfüllten, durften weiterhin im doppelblinden Studiendesign behandelt werden (Extensionsphase), es sei denn, es traten nicht akzeptable Toxizitäten oder eine Progression der Erkrankung gemäß IWG-Kriterien auf [38], die Einverständniserklärung wurde widerrufen oder ein jegliches anderes Kriterium für den Studienabbruch mit der Studienmedikation trat ein. Die Beurteilung des Therapieerfolgs gemäß den Kriterien der Woche-25-Visite wurde alle acht Behandlungszyklen bis zum Behandlungsende wiederholt. Wirksamkeit und Sicherheit wurden weiterhin zu jedem Behandlungszyklus erhoben, Begleitmedikationen entsprechend der Vorgaben während der Behandlungsphase in Woche eins bis 24 waren weiterhin erlaubt – darunter auch die bedarfsgerechte Therapie mit EK-Transfusionen und Eisenchelatoren.

Patienten, die die Kriterien zur Beurteilung des Therapieerfolgs zu Woche 25 nicht erfüllten beendeten die Studienmedikation und wurden im Rahmen des *Posttreatment-Follow-ups* weiter beobachtet. Ebenso wurden Patienten, die während der Extensionsphase die Therapie abbrachen, im Rahmen der *Posttreatment-Follow-up*-Phase weiter beobachtet.

Im Luspatercept-Arm war der Anteil an Patienten, die nach Ermessen des Prüfarztes von der Therapie profitierten, deutlich größer als im Placebo-Arm. So wurden 78 Patienten (51,0 %) im Luspatercept-Arm über Woche 25 hinaus bis Woche 48 behandelt, während dieser Anteil im Placebo-Arm mit nur 12 Patienten (15,8 %) deutlich geringer ausfiel. Dadurch beendete ein größerer Anteil der Patienten im Placebo-Arm die doppelblinde Behandlung vorzeitig, was über den Zeitraum von Woche eins bis 48 betrachtet zu deutlichen Unterschieden in den tatsächlichen Behandlungszeiten zwischen den Therapiearmen führt (siehe Tabelle 4-15). Bis zum Datenschnitt vom 07.01.2019 hatten alle Patienten im Placebo-Arm ihre Behandlung abgebrochen. Aufgrund der daraus resultierenden Unsicherheit hinsichtlich der Interpretierbarkeit der Daten werden die Ergebnisse über den Zeitraum von Woche eins bis 48 ergänzend dargestellt. Dies gilt – mit Ausnahme des Endpunkts Mortalität und der Zeit bis zur Progression in eine AML – für alle dargestellten Endpunkte.

Tabelle 4-15: Behandlungsdauer der Studienpopulation in MEDALIST

<b>Studie MEDALIST</b> <b>(Luspatercept + BSC vs.</b> <b>Placebo + BSC<sup>(1)</sup>)</b>	<b>Luspatercept + BSC<sup>(1)</sup></b> <b>(N = 153)</b>	<b>Placebo + BSC<sup>(1)</sup></b> <b>(N = 76)</b>
<b>Studienzeitraum Woche 1–24</b>		
Anteil Patienten, die Behandlungszeitraum vollständig abgeschlossen haben [n (%)]	128 (83,7)	68 (89,5)
Behandlungsdauer [Wochen]		
Durchschnitt (SD)	22,8 (3,68)	23,0 (3,30)
Median (Min; Max)	24,1 (6; 24)	24,0 (7; 24)
<b>Studienzeitraum Woche 1–48</b>		
Anteil Patienten, die Behandlungszeitraum vollständig abgeschlossen haben [n (%)]	78 (51,0)	12 (15,8)
Behandlungsdauer [Wochen]		
Durchschnitt (SD)	37,2 (13,47)	28,6 (10,44)
Median (Min; Max)	48,1 (6; 48)	24,0 (7; 48)
Datenschnitt: 08.05.2018; Population: ITT-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (s. auch Tabelle 4-13).		
BSC = Best-Supportive-Care; ITT = Intention-to-Treat; Max = Maximum; Min = Minimum; SD = Standardabweichung Quelle: [1, 18]		

### ***Posttreatment-Follow-up-Phase***

Patienten, die die Behandlung abbrachen, wurden im Rahmen der *End-of-treatment* (EOT)-Visite untersucht und der Grund des Studienabbruchs wurde dokumentiert. Unerwünschte Ereignisse wurden bis 42 Tage nach Therapieabbruch erhoben. Darüber hinaus wurde die Transfusionslast für weitere 16 Wochen nach letzter Dosis der Studienmedikation oder EOT-Visite, je nachdem was später eintrat, erhoben.

Die Progression in eine AML und in andere maligne oder prämaligne Erkrankungen, die Folge-therapien der MDS sowie das Überleben wurden für Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben, nach Erhalt der letzten Dosis Studienmedikation für weitere drei Jahre erhoben. Davon ausgenommen sind Patienten, die die Einverständniserklärung zurückgenommen haben, gestorben sind oder während der Verlaufskontrolle verloren gegangen sind.

### ***Datenschnitte***

Die geplante Interims-Futility-Analyse (Datenschnitt vom 27.06.2017) wurde lediglich zur Beurteilung der Sinnhaftigkeit der Studie durchgeführt; eine Bewertung der Wirksamkeit war zu diesem Zeitpunkt explizit ausgeschlossen, sodass auch auf eine Darstellung im Dossier verzichtet wird.

Für die primäre Analyse wurde die Datenbank zu dem Zeitpunkt geschlossen, zu dem der letzte Patient 48 Wochen behandelt worden war oder die Behandlung vorzeitig abgebrochen hatte (Datenschnitt vom 08. Mai 2018). Weitere Analysen vor Studienende waren nicht a priori geplant. Im Rahmen des Zulassungsprozesses von Luspatercept wurden zwei weitere Datenschnitte zur Beantwortung von Fragen der EMA im Rahmen des Zulassungsprozesses durchgeführt (Datenschnitte vom 07.01.2019 und 01.07.2019).

Für die Darstellung im Dossier wird vorrangig die Auswertung nach den ersten 24 Wochen Behandlungsphase herangezogen, da in diesem Studienzeitraum eine doppelblinde Behandlung gemäß Zuteilung in beiden Behandlungsarmen gleichermaßen sichergestellt ist. Für diese Auswertungen lagen zum Zeitpunkt des Datenschnitts im Mai 2018 alle Erhebungen vollständig vor.

Eine Weiterbehandlung entsprechend der ursprünglichen Zuteilung nach Woche 24 war möglich, sofern die Patienten laut Einschätzung des Prüfarztes von der Behandlung profitierten. Diese Einschätzung wurde im weiteren Verlauf alle 24 Wochen überprüft. Durch diese Selektion von Patienten sind die Ergebnisse ab Woche 25 potentiell verzerrt und die Darstellung der Auswertungen über Woche 24 hinaus erfolgt aufgrund der Unsicherheit hinsichtlich der Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten lediglich ergänzend. Bereits zum Datenschnitt vom 08.05.2018 hatten 92,1 % (70/76), zum Datenschnitt vom 07.01.2019 sogar 100 % der Patienten unter Placebo (76/76) die Behandlung abgebrochen. Unter Luspatercept waren es deutlich weniger Patienten mit Behandlungsabbruch (54,2 % (83/153) am 08.05.2018; 68,6 % (105/153) am 07.01.2019). Die Ergebnisse der zusätzlichen, von der EMA

geforderten Datenschnitte vom 07.01.2019 und 01.07.2019 werden daher im vorliegenden Dossier nicht separat dargestellt. Bei Endpunkten, die als Ereigniszeitanalysen ausgewertet wurden, werden die Ereignisse bis zum letzten Datenschnitt (01.07.2019) berücksichtigt.

Die Analyse der Ergebnisse der langfristigen Verlaufskontrolle (*Posttreatment-Follow-up* bis zum finalen Datenschnitt) bezüglich des Gesamtüberlebens, des Progresses der Erkrankung und Folgetherapien der MDS steht zum Zeitpunkt der Einreichung des vorliegenden Dossiers noch nicht zur Verfügung.

### **Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext**

Insgesamt ist von einer guten Übertragbarkeit der Ergebnisse der Studie MEDALIST auf den deutschen Versorgungskontext auszugehen:

- Die Baseline-Charakteristika der eingeschlossenen Patienten spiegeln die Charakteristika der in Deutschland lebenden Patienten mit MDS wider, welche im Rahmen des MDS-Registers in Düsseldorf gut beschrieben sind [91-93]:
  - Medianes Erkrankungsalter 70 – 72 Jahre und
  - leicht erhöhtes Auftreten der Erkrankung bei Männern [91-93].
- Durch den Einschluss von Patienten in den USA und Europa – darunter auch 5 Zentren in Deutschland – wurden vornehmlich weiße Patienten eingeschlossen. Dies entspricht der Bevölkerungszusammensetzung in Deutschland.
- Die beteiligten Studienzentren liegen vornehmlich in westeuropäischen und nordamerikanischen Ländern, sodass von einem ähnlichen Versorgungsstandard wie in Deutschland auszugehen ist. Systematische Unterschiede hinsichtlich der Behandlung von MDS-Patienten zwischen den beteiligten Ländern sind nicht bekannt.

#### 4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-16: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
MEDALIST	ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Die Studie MEDALIST wurde randomisiert und doppelblind durchgeführt. Um die adäquate Gruppenzuteilung und Wahrung der Verblindung der Studienteilnehmer, behandelnden Personen sowie der Firma Celgene zu gewähren, wurde die Randomisierung, Arzneimittel-ausgabe, Dosistitration/-reduktion und Beendigung der Studienmedikation mittels eines IRT (*Interactive Response Technology*)-Systems gestützt. Ausgenommen davon waren lediglich ausgewählte Einzelpersonen (z. B. der Apotheker des Studienzentrums), die eine adäquate Herstellung der Placebo-Gaben sicherstellten.

Die Erhebung von präspezifizierten Endpunkten sowie die Darstellung derer Ergebnisse erfolgte wie vorab definiert. Der Patientenfluss ist nachvollziehbar dargestellt und es wurden keine außergewöhnlichen, das Verzerrungspotential der Studie maßgeblich beeinflussenden Unregelmäßigkeiten berichtet. Dementsprechend liegen weder Hinweise auf eine ergebnis-abhängige Berichterstattung noch sonstige Aspekte, die die Ergebnissicherheit der Studie in Frage stellen würden, vor.

Angaben zum Verzerrungspotential auf Endpunktebene finden sich in Abschnitt 4.3.1.3.1.

Insgesamt ist das Verzerrungspotential der Studie MEDALIST niedrig und die Studie somit geeignet, um Ergebnisse mit hoher Aussagesicherheit zu liefern.

#### 4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-17: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität – Gesamtüberleben	Morbidität				Lebensqualität		Sicherheit <sup>(1)</sup>
		Transusionsfreiheit	Reduktion der Transusionslast	Symptomatik (ERTC QLQ-C30)	Hospitalisierungen	gesundheitsbezogen (EORTC QLQ-C30)	krankheitsspezifisch (QoL-E)	
MEDALIST	ja	ja	ja	ja	ja	ja	ja	ja

<sup>(1)</sup> Dargestellt als jegliche UE, schwere UE (CTCAE ≥ 3), SUE und Therapieabbrüche aufgrund von UE

CTCAE = *Common Terminology Criteria for Adverse Events*; EORTC QLQ-C30 = *European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core-Module*; QoL-E = *Quality of Life-E*; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

#### 4.3.1.3.1 Endpunkte – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind inkl. Angaben zur Häufigkeit von und zum Umgang mit nicht oder nicht vollständig beobachteten Patienten (bei Verlaufsbeobachtungen pro Messzeitpunkt)
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler
- Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde.

Unterschiedliche Beobachtungszeiten zwischen den Behandlungsgruppen sollen durch adäquate Analysen (z.B. Überlebenszeitanalysen) adressiert werden, und zwar für alle Endpunkte (einschließlich UE nach den nachfolgend genannten Kriterien), für die eine solche Analyse aufgrund deutlich unterschiedlicher Beobachtungszeiten erforderlich ist.

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden. Dabei ist für jeden Endpunkt, für den eine solche Analyse durchgeführt wird, eine separate Kaplan-Meier-Kurve darzustellen.

Zu mit Skalen erhobenen patientenberichteten Endpunkten (z.B. zur gesundheitsbezogenen Lebensqualität oder zu Symptomen) sind immer auch die Werte im Studienverlauf anzugeben, auch als grafische Darstellung, sowie eine Auswertung, die die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt (z.B. als Symptomlast über die Zeit, geschätzt mittels MMRM-Analyse [falls aufgrund der Datenlage geeignet]).

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen:

1. Gesamtrate UE,
2. Gesamtrate schwerwiegender UE (SUE),

3. Gesamtrate der Abbrüche wegen UE,

4. Gesamtraten von UE differenziert nach Schweregrad, sofern dies in der/den relevante/n Studie/n erhoben wurde (z.B. gemäß CTCAE und/oder einer anderen etablierten bzw. validierten indikationsspezifischen Klassifikation) einschließlich einer Abgrenzung schwerer und nicht schwerer UE,

5. zu den unter 1, 2 und 4 genannten Kategorien (UE ohne weitere Differenzierung, SUE, UE differenziert nach Schweregrad) soll zusätzlich zu den Gesamtraten die Darstellung nach Organsystemen und Einzelereignissen (als System Organ Class [SOCs] und Preferred Terms [PT] nach MedDRA) jeweils nach folgenden Kriterien erfolgen:

- UE (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10% der Patienten in einem Studienarm aufgetreten sind

- Schwere UE (z.B. CTCAE-Grad  $\geq 3$ ) und SUE: Ereignisse, die bei mindestens 5% der Patienten in einem Studienarm aufgetreten sind

- zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patienten UND bei mindestens 1 % der Patienten in einem Studienarm aufgetreten sind.

6. A priori definierte UE von besonderem Interesse [AESI]) sowie prädefinierte SOC-übergreifende UE-Auswertungen (z.B. als Standardised MedDRA Queries, SMQs) sollen unabhängig von der Ereignisrate dargestellt werden und zwar differenziert nach Schweregrad (dargestellt als Gesamtrate und differenziert nach Schweregrad, nicht schwer, schwer, schwerwiegend).

7. zu Kategorie 3: Die Abbruchgründe auf SOC/PT-Ebene müssen vollständig, jedoch nur deskriptiv dargestellt werden.

Sofern bei der Erhebung unerwünschter Ereignisse erkrankungsbezogenen Ereignisse (z. B. Progression, Exazerbation) berücksichtigt werden (diese Ereignisse also in die UE-Erhebung eingehen), sollen für die Gesamtraten (UE, schwere UE und SUE) zusätzliche UE-Analysen durchgeführt werden, bei denen diese Ereignisse unberücksichtigt bleiben. Alle Auswertungen zu UE können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine Darstellung ausschließlich in Modul 5 ist nicht ausreichend. Davon unbenommen sind die Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE), sowie die für die Gesamtaussage zum Zusatznutzen herangezogenen Ergebnisse im vorliegenden Abschnitt darzustellen.

Auswertungen zu den im Abschnitt 4.3.1.2.1 aufgeführten Datenschnitten sollen vollständig, d.h. für alle erhobenen relevanten Endpunkte, durchgeführt und vorgelegt werden. Das gilt auch dann, wenn ein Datenschnitt ursprünglich nur zur Auswertung einzelner Endpunkte geplant war. Auf die Darstellung der Ergebnisse einzelner Endpunkte eines Datenschnitts bzw. eines gesamten Datenschnitts kann verzichtet werden, wenn hierdurch kein wesentlicher Informationsgewinn gegenüber einem anderen Datenschnitt zu erwarten ist (z. B. wenn die Nachbeobachtung zu einem Endpunkt bereits zum vorhergehenden Datenschnitt nahezu vollständig war oder ein Datenschnitt in unmittelbarer zeitlicher Nähe zu einem anderen Datenschnitt liegt).

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

#### 4.3.1.3.1.1 Mortalität – Gesamtüberleben

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-18: Operationalisierung des Endpunkts Mortalität – Gesamtüberleben

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Mortalität, dargestellt als Gesamtüberleben, war definiert als Zeit zwischen der Randomisierung und dem Todeszeitpunkt. Verstorbene Patienten, unabhängig von der Todesursache, wurden als Patienten mit Ereignis berücksichtigt. Patienten, die zum Zeitpunkt der Analyse noch lebten, die Studie abgebrochen haben oder während der Verlaufsbeobachtung verloren gegangen sind, wurden zu dem letzten Zeitpunkt zensiert, an dem sie bekanntermaßen noch am Leben waren.</p> <p>Das Gesamtüberleben wurde über den gesamten Studienzeitraum (Behandlungs-, Extensions- und Posttreatment-Follow-up-Phase) ausgewertet. Überlebenskurven wurden mittels Kaplan-Meier-Schätzern berechnet. Kaplan-Meier-Schätzer des medianen Gesamtüberlebens und dazugehörige zweiseitige 95 %-KI, welche über eine Log-Log-Transformation nach Brookmeyer und Crowley berechnet wurden, wurden für jeden Behandlungsarm zusammengefasst. Unterschiede zwischen den beiden Behandlungsarmen wurden anhand eines Log-Rank-Tests, stratifiziert nach durchschnittlicher Transfusionslast zur Baseline und dem IPSS-R-Score zur Baseline, analysiert. Das HR inklusive zweiseitigem 95 %-KI für Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, wurde anhand eines stratifizierten Cox-Proportional-Hazards-Modells mit Behandlung als Kovariable geschätzt.</p> <p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Hauptanalyse:</b> Gesamtüberleben; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-21)</li> </ul>
BSC = Best-Supportive-Care; HR = Hazard Ratio; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-19: Bewertung des Verzerrungspotenzials für den Endpunkt Mortalität – Gesamtüberleben in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
MEDALIST	niedrig	ja	ja	ja	nein	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Mortalität – Gesamtüberleben wurde verblindet erhoben. Das ITT-Prinzip wurde adäquat umgesetzt. Anhaltspunkte für eine ergebnisabhängige Berichterstattung liegen nicht vor.

Der Eintritt des Todes jeglicher Ursache ist ein objektives Kriterium, dessen Feststellung frei von Subjektivität erfolgt. In der Studie MEDALIST sollten alle Patienten unabhängig von der Dauer der Behandlung im Rahmen der Studie hinsichtlich ihres Überlebens nachbeobachtet werden. Unterschiede in der Beobachtungszeit liegen nicht vor (siehe Tabelle 4-20). Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Mortalität – Gesamtüberleben und eine hohe Aussagesicherheit der Ergebnisse.

Tabelle 4-20: Beobachtungszeiten für den Endpunkt Mortalität – Gesamtüberleben aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

<b>Studie MEDALIST</b> <b>Beobachtungszeit in Monaten</b>	<b>Luspatercept + BSC<sup>(1)</sup></b> <b>N = 153</b>	<b>Placebo + BSC<sup>(1)</sup></b> <b>N =76</b>
Mittelwert (SD)	23,8 (9,14)	23,2 (9,22)
Median (Min; Max)	26,4 (2,8; 39,6)	26,1 (1,7; 35,6)
Datenschnitt: 01.07.2019; Population: ITT-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).		
BSC = Best-Supportive-Care; ITT = Intention-to-Treat; Max = Maximum; Min = Minimum; SD = Standardabweichung Quelle: [20]		

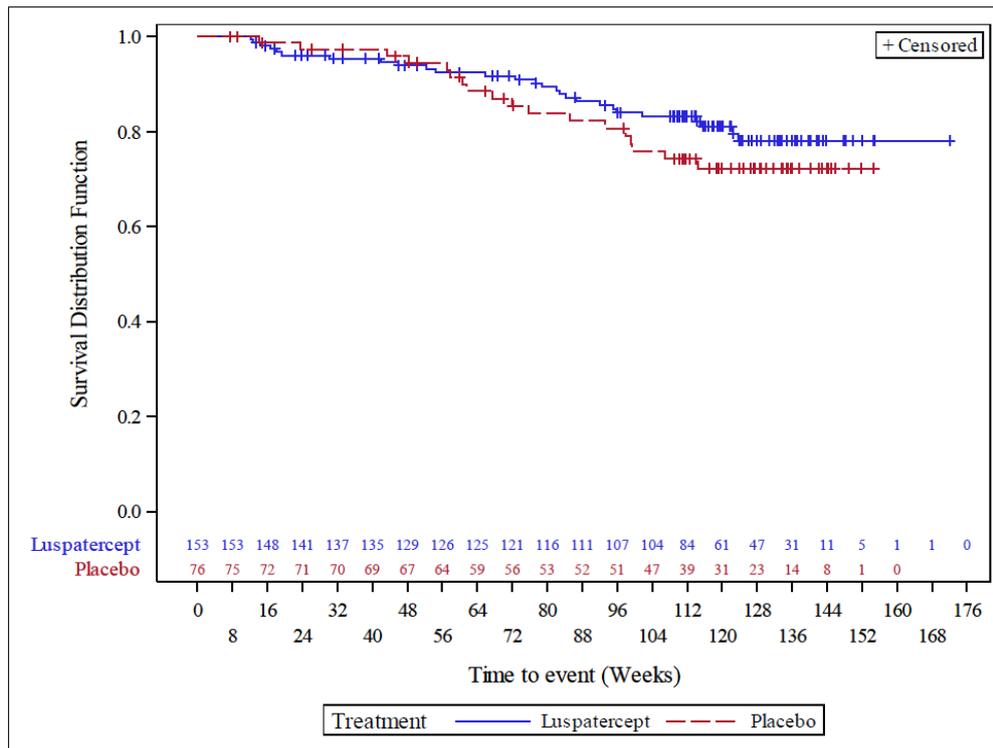
*Stellen Sie die Ergebnisse für den Endpunkt Mortalität – Gesamtüberleben für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

Tabelle 4-21: Ergebnisse für den Endpunkt Mortalität – Gesamtüberleben aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST Endpunkt	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Gesamtüberleben bis Datenschnitt vom 08.05.2018	12/153 (7,8)	N.E. [N.E.; N.E.]	9/76 (11,8)	N.E. [N.E.; N.E.]	0,76 [0,32; 1,83]	0,543
Gesamtüberleben bis Datenschnitt vom 07.01.2019	22/153 (14,4)	N.E. [N.E.; N.E.]	14/76 (18,4)	N.E. [N.E.; N.E.]	0,81 [0,41; 1,58]	0,532
Gesamtüberleben bis Datenschnitt vom 01.07.2019	27/153 (17,6)	N.E. [N.E.; N.E.]	18/76 (23,7)	N.E. [N.E.; N.E.]	0,75 [0,41; 1,38]	0,357

Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.  
<sup>(3)</sup> HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.  
<sup>(4)</sup> Der p-Wert des HR basiert auf einem stratifizierten Log-Rank-Test.

BSC = Best-Supportive-Care; HR = Hazard-Ratio; ITT = Intention-to-Treat; IPSS-R = *International Prognostic Scoring System-revised*; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar  
Quelle: [1, 20, 94]



Datenschnitt: 01.07.2019; Population: ITT-Population;

CI = Konfidenzintervall; ITT = Intention-to-Treat

Quelle: [20]

Abbildung 4-4: Kaplan-Meier-Kurve des Gesamtüberlebens aus MEDALIST; Analysezeitraum bis zum Datenschnitt vom 01.07.2019

Für den Endpunkt Mortalität – Gesamtüberleben zeigte sich zu den dargestellten Datenschnitten kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (Datenschnitt vom 01.07.2019: HR = 0,75 (95 %-KI: [0,41; 1,38]); p = 0,357). Für die Schätzung des medianen Gesamtüberlebens wurden nicht genügend Ereignisse erreicht. Im Luspatercept-Arm traten 27 (17,6 %) und im Placebo-Arm 18 Todesfälle (23,7 %) auf (siehe Tabelle 4-21 & Abbildung 4-4).

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

### **Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

### **Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Mortalität – Gesamtüberleben liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.1.2 Morbidität – Transfusionsfreiheit

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-22: Operationalisierung des Endpunkts Morbidität – Transfusionsfreiheit

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Als transfusionsfreier Zeitraum wird der Zeitraum zwischen zwei aufeinanderfolgenden EK-Transfusionen angesehen. Eine Transfusionsfreiheit war dann erreicht, wenn der Patient über einen entsprechenden Mindestzeitraum, d. h. acht, 12, 16 oder 24 Wochen keine EK-Transfusionen erhalten hatte.</p> <p>Unterschiede zwischen den Behandlungsarmen bei Auswertungen zum Anteil der Patienten mit einer Transfusionsfreiheit wurden mittels CMH-Tests ermittelt, stratifiziert nach durchschnittlicher Transfusionslast und dem IPSS-R-Score zur Baseline. Dargestellt werden die Ergebnisse als OR, stratifiziert berechnet mittels Normal-Approximation, RR und ARR, unstratifiziert berechnet mittels Normal-Approximation, jeweils inkl. zugehörigem 95 %-KI.</p> <p>Die Zeit bis zum Erreichen einer Transfusionsfreiheit war definiert als Zeit zwischen dem Datum der ersten Dosis Studienmedikation und dem Datum des Beginns einer ersten Transfusionsfreiheit über mindestens acht, 12, 16 bzw. 24 Wochen, unter Berücksichtigung der Behandlungs- und Extensionsphase bis zum Datenschnitt vom 01.07.2019. Patienten, die bis zum Datenschnitt vom 01.07.2019 keine Transfusionsfreiheit erreicht hatten, wurden zensiert.</p> <p>Die längste Dauer eines transfusionsfreien Zeitraums war definiert als längster zusammenhängender Zeitraum zwischen zwei aufeinanderfolgenden Transfusionszeitpunkten, in denen keine EK-Transfusion verabreicht wurde. Für Patienten, die bis zum Datenschnitt vom 01.07.2019 an weniger als zwei Transfusionszeitpunkten transfundiert wurden, wurde der Behandlungsbeginn mit Luspatercept als Startdatum eines transfusionsfreien Zeitraums herangezogen. Patienten, deren längster transfusionsfreier Zeitraum zum Zeitpunkt der Analyse anhielt, wurden zensiert.</p> <p>Die mediane Zeit bis zum Erreichen einer Transfusionsfreiheit bzw. die mediane Dauer des längsten transfusionsfreien Zeitraums wurde mittels Kaplan-Meier-Methode geschätzt und inkl. zugehörigem 95 %-KI, welche über eine Log-Log-Transformation nach Brookmeyer und Crowley berechnet wurden, dargestellt. Unterschiede zwischen den Behandlungsarmen wurden mittels stratifiziertem Log-Rank-Test analysiert. Das HR inklusive zweiseitigem 95 %-KI für Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, wurde anhand eines stratifizierten Cox-Proportional-Hazards-Modells mit Behandlung als Kovariable geschätzt.</p> <p>Auswertungen zur mittleren Änderung des Hb-Werts im Vergleich zur Baseline wurden mittels ANCOVA mit Behandlung und Stratifizierungsfaktoren als Kovariablen berechnet. Der Baseline-Hb-Wert war definiert als der mittlere Hb-Wert in den 16 Wochen vor Behandlungsbeginn nach Anwendung der 14/3-Tage-Regel<sup>(1)</sup>. Als Werte im Studienverlauf wurde das beste 16-Wochen-Intervall innerhalb des Analysezeitraums herangezogen. Nur Patienten mit Werten zu Studienbeginn und zum Auswertungszeitpunkt wurden in der Analyse berücksichtigt.</p>

Studie	Operationalisierung
	<p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Hauptanalyse:</b> Anteil der Patienten mit einer Transfusionsfreiheit; Woche 1 – 24 (siehe Tabelle 4-24)</li> <li>• <b>Zusatzanalyse 1:</b> Zeit bis zum Erreichen einer Transfusionsfreiheit; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-25)</li> <li>• <b>Zusatzanalyse 2:</b> Längste Dauer eines transfusionsfreien Zeitraums; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-26)</li> <li>• <b>Zusatzanalyse 3:</b> Swimmer-Plots der Transfusionszeitpunkte und transfusionsfreien Zeiträume; Analysezeitraum bis zum Datenschnitt vom 01.07.2019 (siehe Abbildung 4-10)</li> <li>• <b>Zusatzanalyse 4:</b> Mittlere Änderung des Hb-Werts im Vergleich zur Baseline; Woche 1 – 24, Woche 1 – 48 (siehe Tabelle 4-27 f.)</li> </ul>
	<p><sup>(1)</sup> <b>14/3-Tage-Regel:</b> Es werden nur Hämoglobin-Werte, die mindestens 14 Tage nach einer Transfusion gemessen wurden, berücksichtigt, außer eine weitere Transfusion findet innerhalb von 3 Tagen nach der Hämoglobin-Messung statt. In diesem Fall ist auch dieser Hämoglobin-Wert zu berücksichtigen (trotz Messung innerhalb von &lt; 14 Tagen der ersten Transfusion).</p> <p>ANCOVA = <i>analysis of covariance</i>; ARR = absolute Risikoreduktion; BSC = Best—Supportive-Care; CMH = Cochran-Mantel-Haenszel; EK = Erythrozytenkonzentrat; Hb = Hämoglobin; HR = Hazard Ratio; IPSS-R = <i>International Prognostic Scoring System-revised</i>; IWG = <i>International Working Group</i>; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-23: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Transfusionsfreiheit in MEDALIST

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
MEDALIST	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Morbidität – Transfusionsfreiheit wurde verblindet erhoben. Das ITT-Prinzip wurde adäquat umgesetzt. Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Morbidität – Transfusionsfreiheit und eine hohe Aussagesicherheit der Ergebnisse.

Stellen Sie die Ergebnisse für den Endpunkt Morbidität - Transfusionsfreiheit für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

### Hauptanalyse: Anteil der Patienten mit einer Transfusionsfreiheit

Tabelle 4-24: Ergebnisse für den Endpunkt Morbidität – Transfusionsfreiheit aus MEDALIST – Anteil der Patienten mit einer Transfusionsfreiheit; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(3)</sup>	ARR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Endpunkt	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)				
Transfusionsfrei- heit ≥ 8 Wochen	58/153 (37,9)	10/76 (13,2)	5,06 [2,28; 11,26]	2,88 [1,56; 5,31]	24,75 [13,94; 35,56]	< 0,001
Transfusionsfrei- heit ≥ 12 Wochen	43/153 (28,1)	6/76 (7,9)	5,07 [2,00; 12,84]	3,56 [1,59; 7,99]	20,21 [10,86; 29,56]	< 0,001
Transfusionsfrei- heit ≥ 16 Wochen	29/153 (19,0)	3/76 (3,9)	5,96 [1,75; 20,34]	4,80 [1,51; 15,26]	15,01 [7,41; 22,61]	0,002
Transfusionsfrei- heit ≥ 24 Wochen	20/153 (13,1)	1/76 (1,3)	10,19 [1,47; 70,59]	9,93 [1,36; 72,64]	11,76 [5,83; 17,68]	0,003
Datenschnitt: 08.05.2018; Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittlicher Transfusionslast zur Baseline. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR berechnet mittels Normal-Approximation stratifiziert nach Stratifizierungsfaktoren. <sup>(3)</sup> RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(4)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.  ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; ITT = Intention-to-Treat; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko Quelle: [1, 18]						

Das Erreichen einer Transfusionsfreiheit stellt den im Anwendungsgebiet mit nicht kurativen Therapieoptionen größtmöglichen Behandlungserfolg und therapierelevanten Nutzen für den Patienten dar. Entsprechend war das Erreichen einer Transfusionsfreiheit über einen klinisch relevanten Mindestzeitraum von acht Wochen gemäß IWG-Kriterien aus dem Jahr 2006 auch primärer Endpunkt der Studie MEDALIST. Die im Folgenden näher beschriebenen Ergebnisse belegen den großen therapeutischen Nutzen von Luspatercept und die deutliche Überlegenheit von Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, hinsichtlich dieses Therapieziels.

In der primären Behandlungsphase der Studie (Woche 1 – 24) zeigte sich konsistent über alle betrachteten Zeiträume einer Transfusionsfreiheit ( $\geq 8$ ,  $\geq 12$ ,  $\geq 16$  und  $\geq 24$  Wochen) eine deutliche und statistisch signifikante Überlegenheit für Luspatercept im Vergleich zum Placebo-Arm ( $p < 0,005$  für alle betrachteten Intervalle mit Transfusionsfreiheit).

Für den Mindestzeitraum einer Transfusionsfreiheit von acht Wochen, welcher gemäß IWG-Kriterien (2006) einem Therapieansprechen entspricht [38, 53], erlangten 37,9 % der Patienten (58/153) im Luspatercept-Arm und 13,2 % der Patienten (10/76) im Placebo-Arm eine Transfusionsfreiheit. Dies entspricht einem RR von 2,88 (95 %-KI: [1,56; 5,31]) und einer absoluten Risikodifferenz (Luspatercept – Placebo) von 24,75 %-Punkten (95 %-KI: [13,94; 35,56]) zugunsten von Luspatercept (siehe Tabelle 4-24).

Auch für die Betrachtung der Transfusionsfreiheit über mindestens 12, 16 und 24 Wochen war der Anteil der Patienten mit einer Transfusionsfreiheit unter Luspatercept dreieinhalb- bis fast zehnmal höher als unter Placebo; der Effekt war über alle betrachteten Zeiträume statistisch signifikant.

Dabei ist zu betonen, dass Patienten, die während der primären Behandlungsphase (Woche 1 – 24) eine Transfusionsfreiheit über  $\geq 24$  Wochen erreichten, ab Behandlungsbeginn vollständig transfusionsfrei waren, d. h. mit Beginn der Luspatercept-Therapie keine EK-Transfusionen mehr benötigten. Dies erreichten 13,1 % der Patienten (20/153) unter Luspatercept und nur 1,3 % der Patienten (1/76) im Placebo-Arm (RR = 9,93 (95 %-KI: [1,36; 72,64])).

**Zusatzanalyse 1: Zeit bis zum Erreichen einer Transfusionsfreiheit**

Tabelle 4-25: Zusatzanalyse 1: Ergebnisse für den Endpunkt Morbidität – Transfusionsfreiheit aus MEDALIST – Zeit bis zum Erreichen einer Transfusionsfreiheit; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Transfusionsfrei- heit ≥ 8 Wochen	73/153 (47,7)	52,3 [21,9; N.E.]	12/76 (15,8)	N.E. [N.E.; N.E.]	3,92 [2,12; 7,25]	< 0,001
Transfusionsfrei- heit ≥ 12 Wochen	54/153 (35,3)	N.E. [78,1; N.E.]	9/76 (11,8)	N.E. [N.E.; N.E.]	3,56 [1,75; 7,23]	< 0,001
Transfusionsfrei- heit ≥ 16 Wochen	47/153 (30,7)	N.E. [78,1; N.E.]	6/76 (7,9)	N.E. [N.E.; N.E.]	4,25 [1,81; 9,98]	< 0,001
Transfusionsfrei- heit ≥ 24 Wochen	43/153 (28,1)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	5,78 [2,06; 16,17]	< 0,001

Datenschnitt: 01.07.2019; Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline

<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

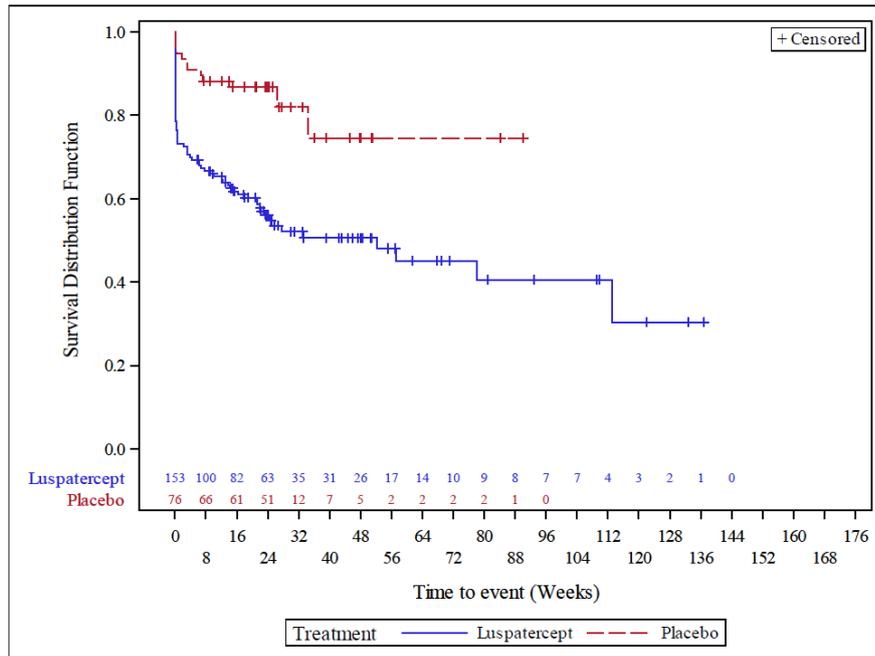
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.

<sup>(3)</sup> HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.

<sup>(4)</sup> Der p-Wert des HR basiert auf einem stratifizierten Log-Rank-Test.

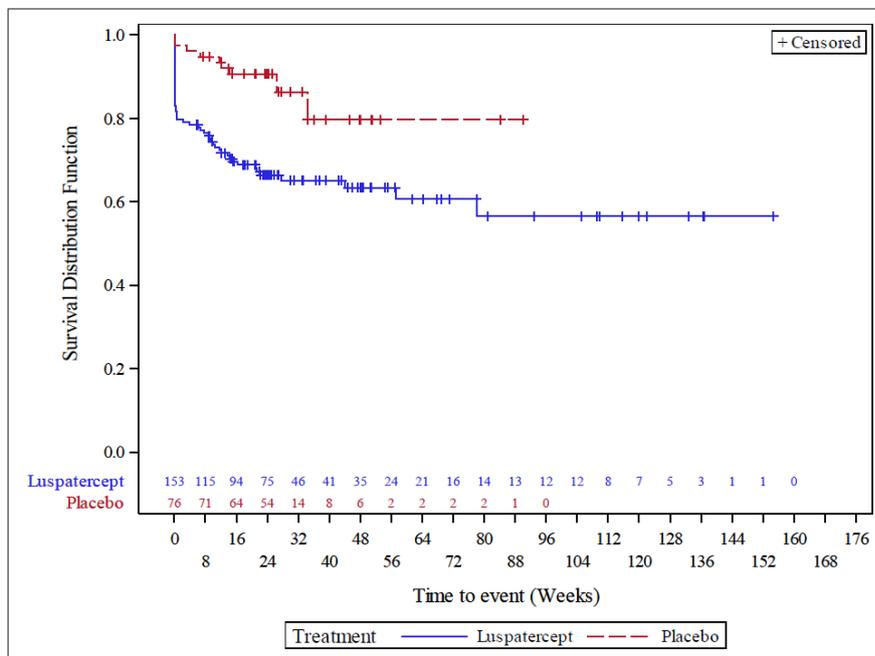
BSC = Best-Supportive-Care; HR = Hazard-Ratio; ITT = *Intention-to-Treat*; IPSS-R = *International Prognostic Scoring System-revised*; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar

Quelle: [20]



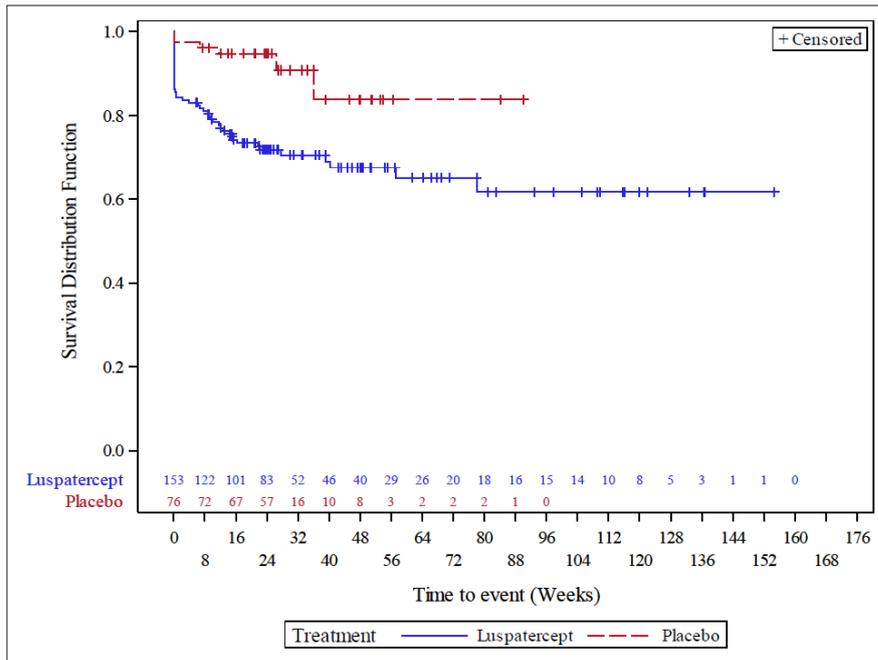
Datenschnitt: 01.07.2019; Population: ITT-Population; ITT = Intention-to-Treat  
 Quelle: [20]

Abbildung 4-5: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur Transfusionsfreiheit über  $\geq 8$  Wochen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



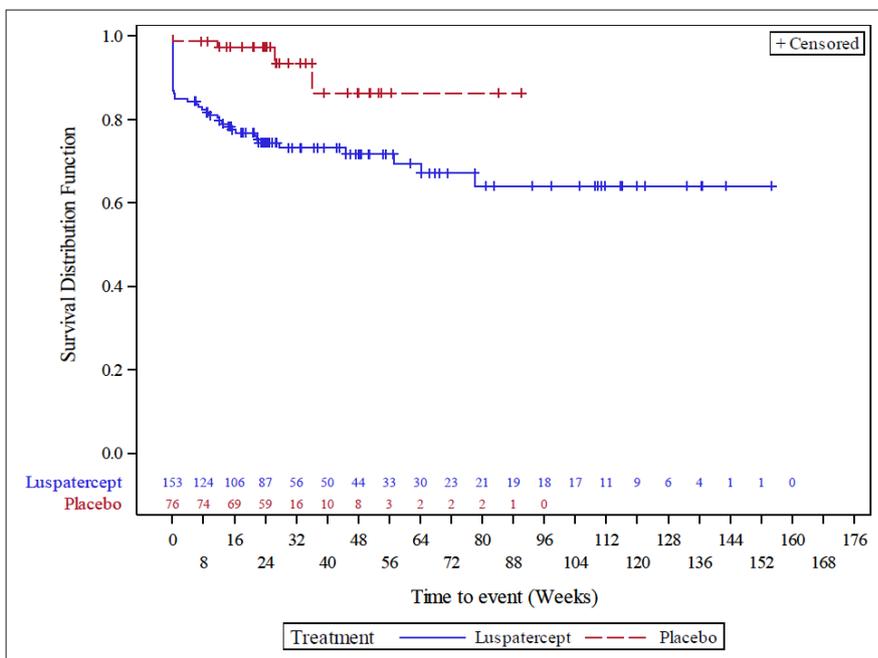
Datenschnitt: 01.07.2019; Population: ITT-Population; ITT = Intention-to-Treat  
 Quelle: [20]

Abbildung 4-6: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur Transfusionsfreiheit über  $\geq 12$  Wochen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019; Population: ITT-Population; ITT = Intention-to-Treat  
 Quelle: [20]

Abbildung 4-7: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur Transfusionsfreiheit über  $\geq 16$  Wochen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019; Population: ITT-Population; ITT = Intention-to-Treat  
 Quelle: [20]

Abbildung 4-8: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur Transfusionsfreiheit über  $\geq 24$  Wochen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

Auswertungen über den Studienzeitraum bis zum Datenschnitt vom 01.07.2019 werden aufgrund der Unsicherheit hinsichtlich der Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalysen dargestellt.

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit einer Transfusionsfreiheit in den ersten 24 Wochen der Behandlungsphase, zeigte sich auch in der Analyse der Zeit bis zur Transfusionsfreiheit ein deutlicher, statistisch signifikanter Vorteil für alle betrachteten Mindestzeiträume einer Transfusionsfreiheit ( $p < 0,001$ ). Der Anteil an Patienten, die eine Transfusionsfreiheit erreichten, war unter Placebo mit  $< 10\%$  (Transfusionsfreiheit über mindestens 16 und 24 Wochen) bzw.  $< 16\%$  (Transfusionsfreiheit über mindestens acht und 12 Wochen) deutlich geringer als im Luspatercept-Arm, in dem die Anteile an Patienten mit Ansprechen durchweg  $> 25\%$  lagen, hinsichtlich einer Transfusionsfreiheit über mindestens acht Wochen sogar  $> 45\%$  (siehe Tabelle 4-25). Besonders hervorzuheben ist auch die Schnelligkeit des Wirkungseintritts unter Luspatercept, die in den Kaplan-Meier-Kurven eindrucksvoll abgebildet wird (siehe Abbildung 4-5 ff.).

Im Verlauf der Studie erreichten die mit Luspatercept behandelten Patienten eine Transfusionsfreiheit über acht, 12, 16 und 24 Wochen somit nicht nur häufiger, sondern auch signifikant früher als unter Placebo.

### Zusatzanalyse 2: Längste Dauer eines transfusionsfreien Zeitraums

Tabelle 4-26: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Transfusionsfreiheit aus MEDALIST – Längste Dauer eines transfusionsfreien Zeitraums; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Beobachtungs- zeitraum	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>
Woche 1 bis zum Datenschnitt vom 01.07.2019	135/153 (88,2)	8,1 [6,0; 10,0]	65/76 (85,5)	4,0 [3,7; 4,6]	0,46 [0,33; 0,63]	$< 0,001$

Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline

<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

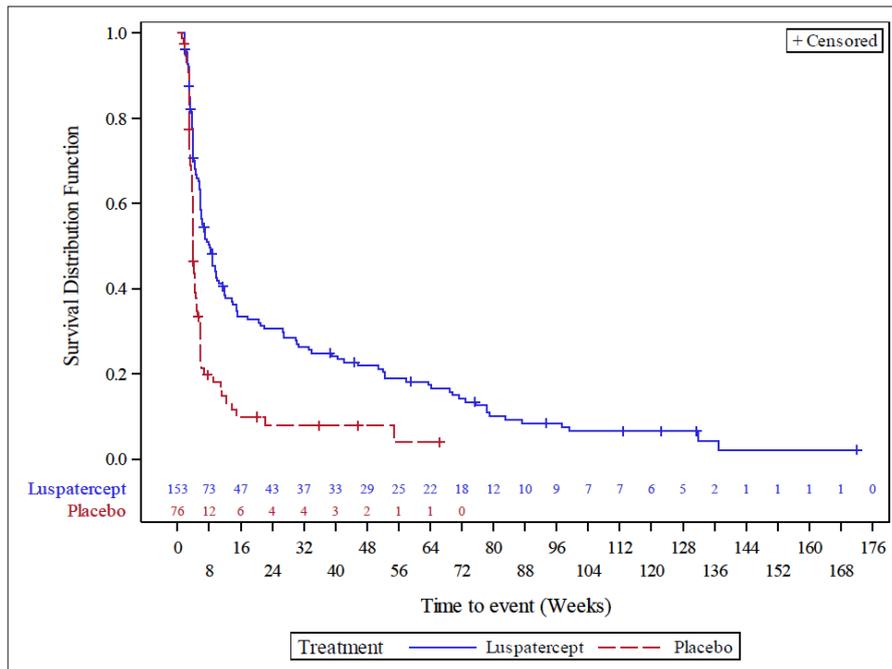
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.

<sup>(3)</sup> HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.

<sup>(4)</sup> Der p-Wert des HR basiert auf einem stratifizierten Log-Rank-Test.

BSC = Best-Supportive-Care; HR = Hazard-Ratio; ITT = Intention-to-Treat; IPSS-R = *International Prognostic Scoring System-revised*; KI = Konfidenzintervall

Quelle: [20]



Datenschnitt: 01.07.2019; + = Zensierung

Quelle: [20]

Abbildung 4-9: Zusatzanalyse 2: Kaplan-Meier-Kurve der längsten Dauer eines transfusionsfreien Zeitraums in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

Die längste Dauer eines transfusionsfreien Zeitraums wird mittels Kaplan-Meier-Methode geschätzt, dementsprechend wurde die Dauer ab dem ersten Tag eines zusammenhängenden Zeitraums ohne EK-Transfusionen bis das Ereignis einer erneuten EK-Transfusion eintritt gezählt. Neben der medianen Dauer eines transfusionsfreien Zeitraums kann mithilfe dieser Analyse somit auch das Risiko für Folgetransfusionen abgeschätzt werden.

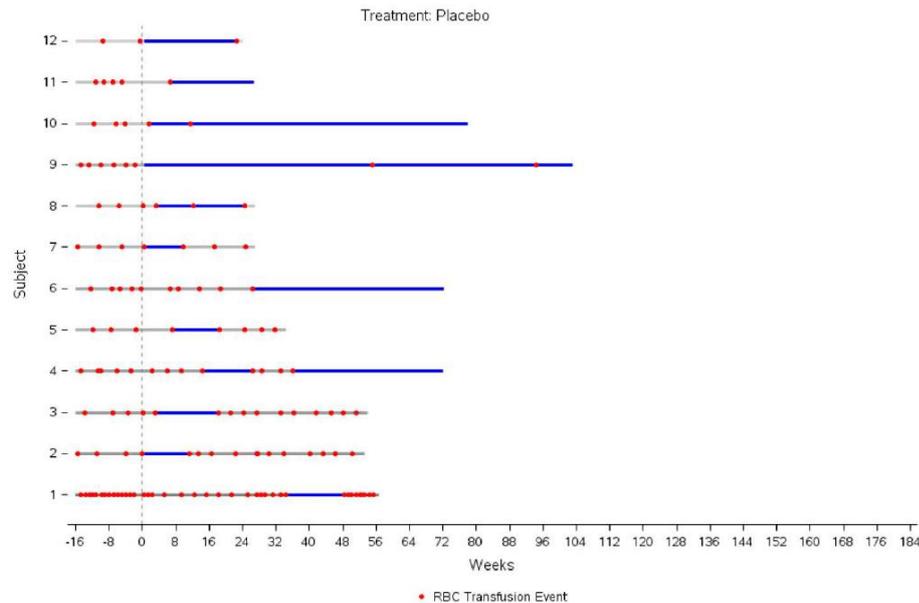
Die Ergebnisse zur längsten Dauer eines transfusionsfreien Zeitraums bis zum Datenschnitt vom 01.07.2019 belegen, dass für Luspatercept nicht nur hinsichtlich des Erreichens einer Transfusionsfreiheit über einen klinisch relevanten Zeitraum von mindestens acht, 12, 16 und 24 Wochen ein deutlicher Vorteil vorliegt, sondern auch hinsichtlich der generellen Verlängerung der Abstände zwischen aufeinanderfolgenden Transfusionszeitpunkten, ein deutlicher, statistisch signifikanter Vorteil gegenüber Placebo besteht ( $p < 0,001$ ).

Das Risiko, eine erneute EK-Transfusion zu erhalten, war für Patienten, die mit Luspatercept behandelt werden, gegenüber den Patienten im Placebo-Arm um fast 55 % reduziert (HR = 0,46; 95 %-KI: [0,33; 0,63]; siehe Tabelle 4-26 & Abbildung 4-9).

### Zusatzanalyse 3: Swimmer-Plots der Transfusionszeitpunkte und Intervalle einer Transfusionsfreiheit

Luspatercept  
+ BSC



**Placebo +  
BSC**

Schwimmbahn auf der y-Achse repräsentiert individuelle Patienten, die während der primären Behandlungsphase (Woche 1 – 24) eine Transfusionsfreiheit mindestens acht Wochen erreichten. blaue Schwimmbahn = Transfusionsfreiheit über mindestens acht Wochen; roter Punkt = Transfusionszeitpunkt; BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Abbildung 4-10: Zusatzanalyse 3: Swimmer-Plot der Transfusionszeitpunkte und transfusionsfreien Zeiträume über mindestens acht Wochen aus MEDALIST; Studienzeitraum bis zum Datenschnitt vom 01.07.2019

Bei Betrachtung der Swimmer-Plots der Transfusionszeitpunkte und transfusionsfreien Zeiträume wird der große Vorteil der Behandlung mit Luspatercept hinsichtlich des Erreichens einer Transfusionsfreiheit deutlich. Neben den bereits in den oben stehenden Analysen gezeigten statistisch signifikanten Effekten hinsichtlich des Anteils an Patienten mit einer Transfusionsfreiheit über mindestens acht, 12, 16 und 24 Wochen, der Zeit bis zum Erreichen einer Transfusionsfreiheit und der längsten Dauer transfusionsfreier Intervalle, lässt sich aus den hier gezeigten Schwimmbahnen ein weiterer Vorteil der Luspatercept-Therapie gegenüber Placebo ableiten: der Großteil der mit Luspatercept behandelte Patienten erreichte nicht nur einmal, sondern mehrfach eine Transfusionsfreiheit, sofern in der Zwischenzeit eine EK-Transfusion nötig wurde.

Insgesamt zeigte sich im Vergleich zum Placebo-Arm ein bemerkenswerter Vorteil bei mit Luspatercept behandelten Patienten hinsichtlich

1. des Anteils an Patienten, die eine Transfusionsfreiheit erreichen,
2. der Schnelligkeit, mit der Patienten eine Transfusionsfreiheit erreichten,
3. der Dauer der Transfusionsfreiheit und
4. der Anzahl an Intervallen mit einer Transfusionsfreiheit (siehe Abbildung 4-10).

**Zusatzanalyse 4: Mittlere Änderung des Hb-Werts**

Tabelle 4-27: Zusatzanalyse 4: Ergebnisse für den Endpunkt Morbidität – Transfusionsfreiheit aus MEDALIST – Mittlere Änderung des Hb-Werts

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studienbeginn	MW (SD) zum Auswertungszeitpunkt <sup>(3)</sup>	N <sup>(2)</sup>	MW (SD) zu Studienbeginn	MW (SD) zum Auswertungszeitpunkt <sup>(3)</sup>	LS-MWD [95 %-KI]	p-Wert <sup>(4)</sup>	SMD [95 %-KI]
<b>Woche 1 – 24 (Hb-Spiegel [g/dl]/16 Wochen)</b>								
152	7,68 (0,84)	9,07 (0,87)	75	7,64 (0,78)	8,51 (0,74)	0,54 [0,34; 0,74]	< 0,001	0,61 [0,33; 0,89]
<b>Woche 1 – 48 (Hb-Spiegel [g/dl]/16 Wochen)</b>								
152	7,68 (0,84)	9,31 (0,90)	75	7,64 (0,78)	8,75 (0,81)	0,54 [0,33; 0,76]	< 0,001	0,56 [0,28; 0,84]
<p>Datenschnitt: 08.05.2018; Population: ITT-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline.  Einheit: Hb-Spiegel [g/dl]/16 Wochen.  <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  <sup>(2)</sup> Anzahl Patienten mit nicht-fehlendem Wert zu Studienbeginn und Auswertungszeitpunkt.  <sup>(3)</sup> Bestes 16-Wochen-Intervall innerhalb des Analysezeitraums.  <sup>(4)</sup> Ergebnisse für die Veränderung zur Baseline aus ANCOVA-Modell mit Behandlung und Stratifizierungsfaktoren als Kovariablen.</p> <p>ANCOVA = <i>analysis of covariance</i>; BSC = Best-Supportive-Care; ITT = Intention-to-Treat; IPSS-R = <i>International Prognostic Scoring System-revised</i>; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate (<i>least square</i>); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SMD = standardisierte Mittelwertdifferenz nach Hedges' g  Quelle: [18]</p>								



**Zusammenfassung der Ergebnisse zum Endpunkt Transfusionsfreiheit**

Insgesamt zeigten sich für den Endpunkt Transfusionsfreiheit ausschließlich positive, über alle betrachteten Analysen hinweg konstante Effekte zugunsten von Luspatercept. Der bedeutende patientenrelevante Zusatznutzen von Luspatercept beruht primär auf dem deutlichen, statistisch signifikant höheren Anteil an Patienten, die eine Transfusionsfreiheit erreichten. Dieser Vorteil war über alle betrachteten Zeiträume einer Transfusionsfreiheit (acht, 12, 16 und 24 Wochen) im Ausmaß erheblich. Ergänzend dargestellte Analysen unterstreichen diesen bemerkenswerten, therapie relevanten Nutzen von Luspatercept hinsichtlich der folgenden Aspekte:

- Patienten unter Luspatercept erreichten nicht nur statistisch signifikant häufiger, sondern auch schneller eine Transfusionsfreiheit als mit Placebo behandelte Patienten.
- Unabhängig vom Erreichen einer Transfusionsfreiheit über einen klinisch relevanten Zeitraum von mindestens acht Wochen, bewirkte Luspatercept insgesamt eine statistisch signifikante Verlängerung der transfusionsfreien Intervalle. Das Risiko eine erneute EK-Transfusion zu benötigen war unter Luspatercept im Vergleich zu Placebo deutlich reduziert.

Der Großteil der Patienten, die nach Erreichen einer Transfusionsfreiheit erneute EK-Transfusionen benötigen, wurden unter Luspatercept wieder transfusionsfrei, d. h. sie sprachen mehrfach an. Patienten, die mit Placebo behandelt wurden, sprachen hingegen, sofern sie überhaupt einmal transfusionsfrei wurden, maximal zweimal an. Alle hier beobachteten patientenrelevanten Effekte können nur erreicht werden, wenn die Erythropoese in ausreichendem Maße funktionsfähig ist. Ergänzend dargestellte Ergebnisse zum Hb-Spiegel belegen die Luspatercept-induzierte Regulation der ineffektiven Erythropoese und damit verbundenen Verbesserung der Anämie, die die Notwendigkeit von EK-Transfusionen verhindert bzw. stark minimiert. Dieser Effekt trat sehr schnell, d. h. bereits während des ersten Behandlungszyklus ein, und hielt während der gesamten Studiendauer stabil an.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

### **Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

### **Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Morbidität – Transfusionsfreiheit liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.1.3 Morbidität – Reduktion der Transfusionslast

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-28: Operationalisierung des Endpunkts Morbidität – Reduktion der Transfusionslast

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Die Transfusionslast bzw. Transfusionshäufigkeit war definiert als Anzahl transfundierter EK-Einheiten bzw. Anzahl Transfusionszeitpunkte pro definiertem Zeitintervall.</p> <p>Die Transfusionslast/Transfusionshäufigkeit zur Baseline war definiert als Anzahl transfundierter EK-Einheiten/Anzahl Transfusionszeitpunkte in den acht bzw. 16 Wochen vor Behandlungsbeginn (Dosis 1 Tag 1).</p> <p>Im vorliegenden Anwendungsgebiet entspricht eine klinisch relevante Reduktion der Transfusionslast einem HI-E, das 2006 durch die IWG als etabliertes Ansprechkriterium zur Bewertung des Therapieerfolgs von Patienten mit MDS definiert wurde als:</p> <ul style="list-style-type: none"> <li>• für Patienten mit einer Transfusionslast von mindestens vier EK pro acht Wochen vor Behandlungsbeginn: Reduktion der Transfusionslast um mindestens vier EK-Einheiten pro acht Wochen im Vergleich zu dem Acht-Wochen-Zeitraum vor Behandlungsbeginn oder</li> <li>• für Patienten mit einer Transfusionslast von weniger als vier EK pro acht Wochen vor Behandlungsbeginn: Erhöhung des Hb-Werts um <math>\geq 1,5</math> g/dl bei gleichzeitiger Transfusionsfreiheit über mindestens acht Wochen im Vergleich zu dem Acht-Wochen-Zeitraum vor Behandlungsbeginn.</li> </ul> <p>Der Anteil an Patienten mit einem HI-E in den ersten 24 Wochen der Behandlungsphase wurde als Summe der Anteile der Patienten mit einer Transfusionslast <math>\geq 4</math> EK bzw. <math>&lt; 4</math> EK pro acht Wochen vor Behandlungsbeginn, die entsprechend ihrer jeweiligen Kriterien ein Ansprechen erreicht hatten, ausgewertet. Eine Darstellung der Patienten mit HI-E getrennt nach ihrer Transfusionslast zur Baseline erfolgt ergänzend.</p> <p>Unterschiede zwischen den Behandlungsarmen wurden mittels CMH-Tests, stratifiziert nach durchschnittlicher Transfusionslast und IPSS-R-Risikoklassifizierung zur Baseline, berechnet. Dargestellt werden die Ergebnisse als OR, stratifiziert berechnet mittels Normal-Approximation, RR und ARR, unstratifiziert berechnet mittels Normal-Approximation, inkl. jeweiligem 95 %-KI.</p> <p>Die Zeit bis zum Erreichen eines HI-E war definiert als Zeit zwischen dem Datum der ersten Dosis Studienmedikation und dem Datum des Beginns des ersten Intervalls mit einem HI-E bis zum Datenschnitt vom 01.07.2019. Patienten, die bis zum Datenschnitt vom 01.07.2019 kein HI-E erreicht hatten, wurden zensiert.</p> <p>Die mediane Zeit bis zum Erreichen eines HI-E wurde mittels Kaplan-Meier-Methode geschätzt und inkl. zugehörigem 95 %-KI, welche über eine Log-Log-Transformation nach Brookmeyer und Crowley berechnet wurden, dargestellt. Unterschiede zwischen den Behandlungsarmen wurden mittels stratifiziertem Log-Rank-Test analysiert. Das HR inklusive zweiseitigem 95 %-KI für Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, wurde anhand eines stratifizierten Cox-Proportional-Hazards-Modells mit Behandlung als Kovariable geschätzt.</p> <p>Auswertungen zur mittleren Änderung der Transfusionslast und der Transfusionshäufigkeit im Vergleich zur Baseline wurden mittels ANCOVA mit Behandlung und Stratifizierungsfaktoren als Kovariablen berechnet. Als Werte im Studienverlauf wurde das beste 16-Wochen-Intervall innerhalb des Analysezeitraums herangezogen. Nur Patienten mit Werten zu Studienbeginn und zum Auswertungszeitpunkt wurden in der Analyse berücksichtigt.</p>

Studie	Operationalisierung
	<p>Wenn für einen einzelnen Patienten mehrere Transfusionen an einem Datum gegeben werden, wurden diese als ein Ereignis (Transfusionszeitpunkt) gezählt. Für Patienten, die an zwei aufeinanderfolgenden Tagen mehrere Transfusionen erhielten, wurden diese als ein Ereignis (Transfusionszeitpunkt) gezählt. Wenn für einen einzelnen Patienten an drei aufeinanderfolgenden Tagen Transfusionen gegeben wurden, wurden diese als zwei Ereignisse (Transfusionszeitpunkte) gezählt.</p> <p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Hauptanalyse:</b> Anteil der Patienten mit einem HI-E; Woche 1 – 24 (siehe Tabelle 4-30)</li> <li>• <b>Zusatzanalyse 1:</b> Zeit bis zum Erreichen eines HI-E; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-31)</li> <li>• <b>Zusatzanalyse 2:</b> Mittlere Änderung der Transfusionslast im Vergleich zur Baseline; Woche 1 – 24 und Woche 1 – 48 (siehe Tabelle 4-32)</li> <li>• <b>Zusatzanalyse 3:</b> Mittlere Änderung der Transfusionshäufigkeit im Vergleich zur Baseline; Woche 1 – 24 und Woche 1 – 48 (siehe Tabelle 4-33)</li> </ul>
<p>ANCOVA = <i>analysis of covariance</i>; ARR = absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran-Mantel-Haenszel; EK = Erythrozytenkonzentrat; HI-E = erythroides Ansprechen (<i>hematologic improvement-erythroid</i>); HR = Hazard Ratio; IPSS-R = <i>International Prognostic Scoring System-revised</i>; IWG = International Working Group; KI = Konfidenzintervall; MDS = Myelodysplastische Syndrome; OR = Odds Ratio; RR = Relatives Risiko</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-29: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Reduktion der Transfusionslast in MEDALIST

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
MEDALIST	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Morbidität – Reduktion der Transfusionslast wurde verblindet erhoben. Das ITT-Prinzip wurde adäquat umgesetzt. Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Morbidität – Reduktion der Transfusionslast und eine hohe Aussagesicherheit der Ergebnisse.

*Stellen Sie die Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

### Hauptanalyse: Anteil der Patienten mit einem HI-E

Tabelle 4-30: Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST – Anteil der Patienten mit einem HI-E; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(4)</sup>	ARR [95 %-KI] <sup>(4)</sup>	p-Wert <sup>(3)</sup>
Endpunkt	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)				
HI-E (insgesamt)	81/153 (52,9)	9/76 (11,8)	8,59 [3,94; 18,70]	4,47 [2,38; 8,41]	41,10 [30,36; 51,84]	<0,001
HI-E I <sup>(5)</sup>	52/107 (48,6)	8/56 (14,3)	5,97 [2,52; 14,14]	3,40 [1,74; 6,65]	34,31 [21,13; 47,49]	<0,001
HI-E II <sup>(6)</sup>	29/46 (63,0)	1/20 (5,0)	23,25 [3,11; 173,59]	12,61 [1,84; 86,27]	58,04 [41,14; 74,95]	<0,001

Datenschnitt: 08.05.2018; Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittlicher Transfusionslast zur Baseline.

<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

<sup>(2)</sup> OR berechnet mittels Normal-Approximation stratifiziert nach Stratifizierungsfaktoren.

<sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.

<sup>(4)</sup> RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet.

<sup>(5)</sup> Anteil der Patienten mit hoher Transfusionslast zur Baseline ( $\geq 4$  EK-Einheiten pro acht Wochen) mit einer Reduktion der Transfusionslast um  $\geq 4$  EK-Einheiten pro acht Wochen.

<sup>(6)</sup> Anteil der Patienten mit niedriger Transfusionslast zur Baseline ( $< 4$  EK-Einheiten pro acht Wochen) mit einer Erhöhung des Hb-Werts um  $\geq 1,5$  g/dl bei gleichzeitiger Transfusionsfreiheit über mindestens acht Wochen.

ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; EK = Erythrozytenkonzentrat; HI-E = erythroides Ansprechen (*hematologic improvement-erythroid*); ITT = Intention-to-Treat; IPSS-R = *International Prognostic Scoring System-revised*; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko

Quelle: [1, 18]

Bei Patienten mit Niedrigrisiko-MDS steht die Therapie der Zytopenie(n) – v. a. die Anämie infolge einer ineffektiven Erythropoese – im Fokus der Behandlung. Aus Sicht der IWG ist eine Verbesserung der Erythropoese, d.h. ein erythroides Ansprechen gemäß IWG-Kriterien 2006 (HI-E), dementsprechend ein essentielles Therapieziel, welches die effektive Behandlung der Anämie widerspiegelt. Luspatercept zeigte hinsichtlich dieses Therapieziels eine deutliche und statistisch signifikante Überlegenheit im Vergleich zu Placebo ( $p < 0,001$ ).

Insgesamt erreichten 52,9 % der Patienten (81/153) im Luspatercept-Arm und nur 11,8 % der Patienten (9/76) im Placebo-Arm ein HI-E. Dies entspricht einem RR von 4,47 (95 %-KI: [2,38; 8,41]) und einer absoluten Risikodifferenz (Luspatercept – Placebo) von 41,1 %-Punkten (95 %-KI: [30,36; 51,84]) zugunsten von Luspatercept (siehe Tabelle 4-30). Dieser Vorteil besteht konsistent sowohl für Patienten mit einer hohen ( $\geq 4$  EK-Einheiten pro acht Wochen) und geringen Transfusionslast ( $< 4$  EK-Einheiten pro acht Wochen) vor Behandlungsbeginn entsprechend der IWG-Kriterien (2006).

### Zusatzanalyse 1: Zeit bis zum Erreichen eines HI-E

Tabelle 4-31: Zusatzanalyse 1: Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST – Zeit bis zum Erreichen eines HI-E; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept vs. Placebo	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
HI-E (insgesamt)	93/153 (60,8)	13,1 [4,0; 27,1]	14/76 (18,4)	60,1 [39,4; N.E.]	4,37 [2,48; 7,70]	< 0,001
HI-E I <sup>(5)</sup>	59/107 (55,1)	21,9 [7,0; 33,1]	12/56 (21,4)	N.E. [29,1; N.E.]	3,32 [1,77; 6,22]	< 0,001
HI-E II <sup>(6)</sup>	34/46 (73,9)	2,7 [0,1; 24,6]	2/20 (10,0)	60,1 [N.E.; N.E.]	10,31 [2,46; 43,19]	0,001

Datenschnitt: 01.07.2019; Population: Intention-to-Treat Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline

<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.

<sup>(3)</sup> HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.

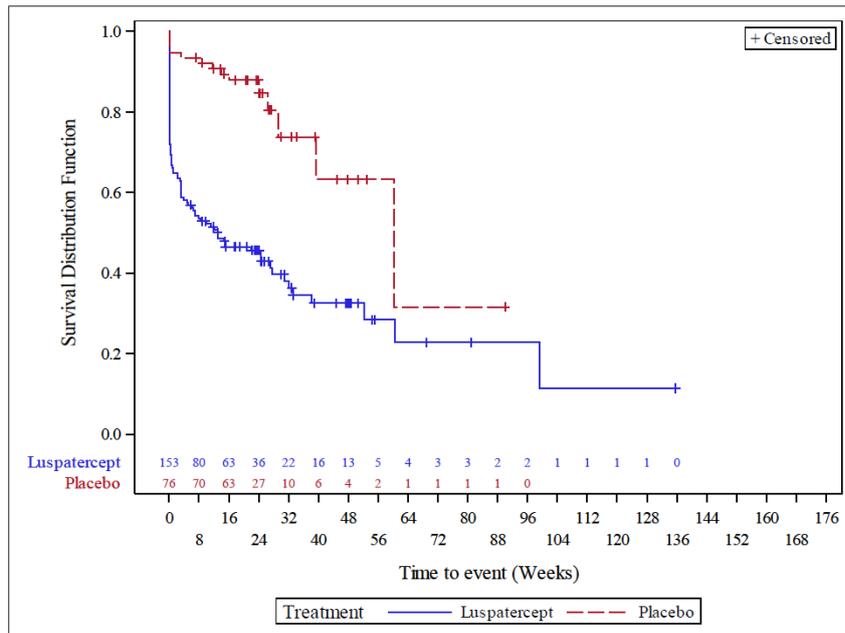
<sup>(4)</sup> Der p-Wert des HR basiert auf einem stratifizierten Log-Rank-Test.

<sup>(5)</sup> Anteil der Patienten mit hoher Transfusionslast zur Baseline ( $\geq 4$  EK-Einheiten pro acht Wochen) mit einer Reduktion der Transfusionslast um  $\geq 4$  EK-Einheiten pro acht Wochen.

<sup>(6)</sup> Anteil der Patienten mit niedriger Transfusionslast zur Baseline ( $< 4$  EK-Einheiten pro acht Wochen) mit einer Erhöhung des Hb-Werts um  $\geq 1,5$  g/dl bei gleichzeitiger Transfusionsfreiheit über mindestens acht Wochen.

BSC = Best-Supportive-Care; HR = Hazard-Ratio; EK = Erythrozytenkonzentrat; HI-E = erythroides Ansprechen (*hematologic improvement-erythroid*); ITT = Intention-to-Treat; IPSS-R = *International Prognostic Scoring System-revised*; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar

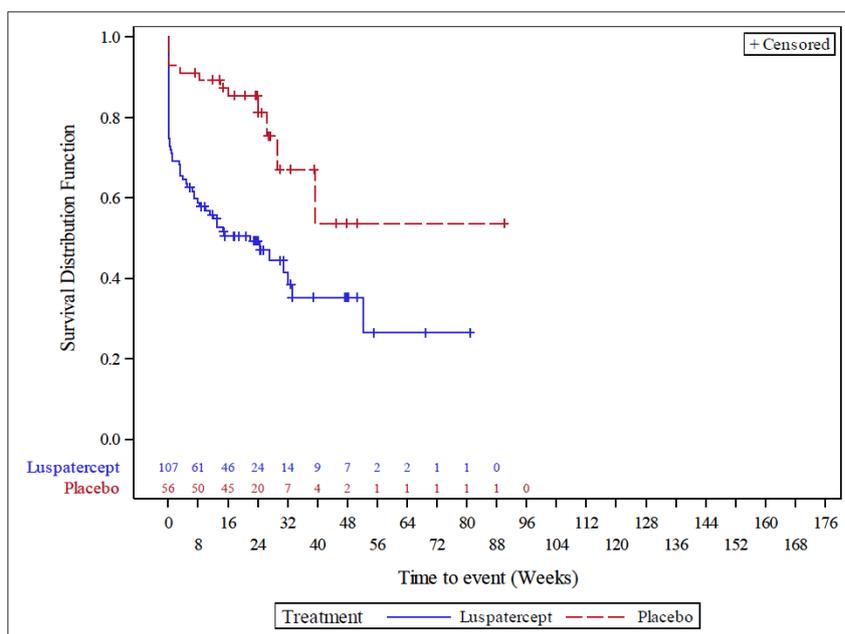
Quelle: [20]



Datenschnitt: 01.07.2019; + = Zensierung

Quelle: [20]

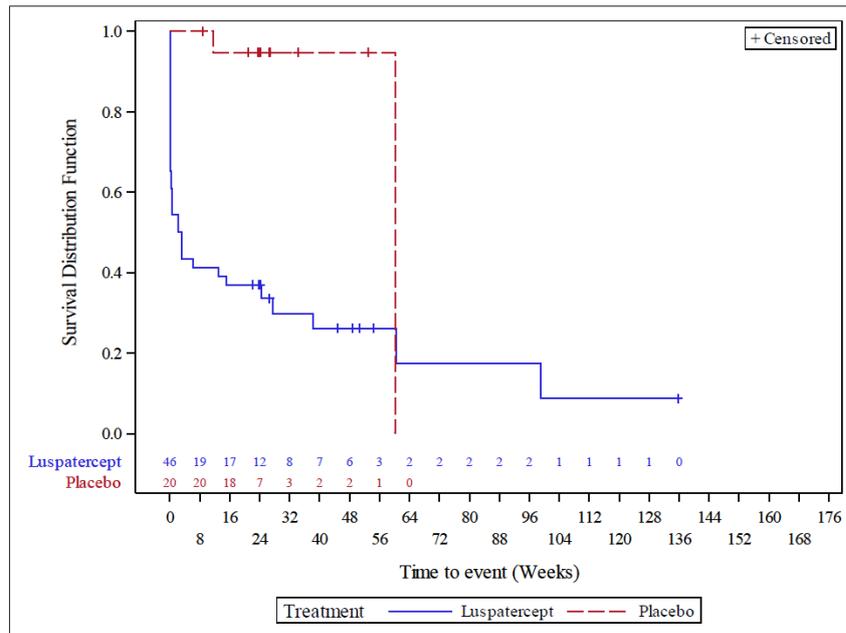
Abbildung 4-12: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum Erreichen eines HI-E (insgesamt) in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019; + = Zensierung

Quelle: [20]

Abbildung 4-13: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum Erreichen eines HI-E I für Patienten mit hoher Transfusionslast zur Baseline ( $\geq 4$  EK-Einheiten pro acht Wochen) in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019; + = Zensierung

Quelle: [20]

Abbildung 4-14: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum Erreichen eines HI-E II für Patienten mit niedriger Transfusionslast zur Baseline (< 4 EK-Einheiten pro acht Wochen) in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

Auswertungen über den Studienzeitraum bis zum Datenschnitt vom 01.07.2019 werden aufgrund der Unsicherheit hinsichtlich der Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalysen dargestellt.

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit einem HI-E in den ersten 24 Wochen der Behandlungsphase, zeigte sich auch in der Analyse der Zeit bis zum Erreichen eines HI-E ein deutlicher, statistisch signifikanter Vorteil für Luspatercept ( $p < 0,001$ ). Einerseits war die Wahrscheinlichkeit für das Erreichen eines HI-E unter Luspatercept 4,4-mal höher als unter Placebo (HR = 4,37; 95 %-KI: [2,48; 7,70]) und andererseits trat dieser Effekt unter Luspatercept auch deutlich schneller ein (mediane Zeit bis zum HI-E: Luspatercept: 13,1 Wochen (95 %-KI: [4,0; 27,1]; Placebo: 60,1 Wochen (95 %-KI: [39,4; N.E.]); siehe Tabelle 4-31 & Abbildung 4-12). Auch für Patienten mit einer Transfusionslast von mindestens vier EK-Einheiten pro acht Wochen und weniger als vier EK-Einheiten pro acht Wochen vor Behandlungsbeginn entsprechend der IWG-Kriterien (2006) bestand dieser große Vorteil konsistent zur Gesamtanalyse.

**Zusatzanalyse 2: Mittlere Änderung der Transfusionslast**

Tabelle 4-32: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST – Mittlere Änderung der Transfusionslast

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studienbeginn	MW (SD) zum Auswertungszeitpunkt <sup>(3)</sup>	N <sup>(2)</sup>	MW (SD) zu Studienbeginn	MW (SD) zum Auswertungszeitpunkt <sup>(3)</sup>	LS-MWD [95 %-KI]	p-Wert <sup>(4)</sup>	SMD [95 %-KI]
<b>Woche 1 – 24 (Anzahl EK-Einheiten/16 Wochen)</b>								
140	10,44 (5,16)	6,09 (5,47)	71	11,73 (6,03)	10,52 (5,80)	-3,53 [-4,70; -2,36]	<0,001	-0,71 [-1,00; -0,41]
<b>Woche 1 – 48 (Anzahl EK-Einheiten/16 Wochen)</b>								
140	10,44 (5,16)	5,29 (5,60)	71	11,73 (6,03)	9,86 (5,89)	-3,77 [-5,05; -2,49]	<0,001	-0,69 [-0,98; -0,40]
Datenschnitt: 08.05.2018; Population: ITT-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline. Einheit: EK-Einheiten/16 Wochen <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> Anzahl Patienten mit nicht-fehlerndem Wert zu Studienbeginn und Auswertungszeitpunkt. <sup>(3)</sup> Bestes 16-Wochen-Intervall innerhalb des Analysezeitraums. <sup>(4)</sup> Ergebnisse für die Veränderung zur Baseline aus ANCOVA-Modell mit Behandlung und Stratifizierungsfaktoren als Kovariablen.  ANCOVA = <i>analysis of covariance</i> ; BSC = Best-Supportive-Care; ITT = Intention-to-Treat; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SMD = standardisierte Mittelwertdifferenz nach Hedges' g Quelle: [18]								

Sowohl während der primären Behandlungsphase (Woche 1 – 24) als auch bis Woche 48 zeigte sich unter Luspatercept eine deutliche, statistisch signifikant höhere Reduktion der mittleren Transfusionslast als unter Placebo ( $p < 0,001$ ). Die Veränderung der SMD nach Hedges' g zeigt darüber hinaus einen klinisch relevanten Effekt an (95 %-KI der SMD liegt vollständig außerhalb der Irrelevanzschwelle von -0,2: SMD(Woche 1 – 24) = -0,71; 95 %-KI: [-1,00; -0,41]; SMD(Woche 1 – 48) = -0,69; 95 %-KI: [-0,98; -0,40]). Die Anzahl an transfundierten EK-Einheiten war unter Luspatercept um etwa 4 EK-Einheiten pro 16 Wochen niedriger als unter Placebo (LS-MWD(Woche 1 – 24): -3,53; 95 %-KI: [-4,70; -2,36]; LS-MWD(Woche 1 – 48): -3,77; 95 %-KI: [-5,05; -2,49]; siehe Tabelle 4-32). Für einen regelmäßig transfundierten Patienten mit einer durchschnittlichen Transfusionslast von 11 EK-Einheiten/16 Wochen entspricht diese Reduktion der Transfusionslast einer Einsparung der Menge benötigter EK-Einheiten um etwa ein Drittel.

**Zusatzanalyse 3: Mittlere Änderung der Transfusionshäufigkeit**

Tabelle 4-33: Zusatzanalyse 3: Ergebnisse für den Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST – Mittlere Änderung der Transfusionshäufigkeit

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studienbeginn	MW (SD) zum Auswertungszeitpunkt <sup>(3)</sup>	N <sup>(2)</sup>	MW (SD) zu Studienbeginn	MW (SD) zum Auswertungszeitpunkt <sup>(3)</sup>	LS-MWD [95 %-KI]	p-Wert <sup>(4)</sup>	SMD [95 %-KI]
<b>Woche 1 – 24 (Anzahl Transfusionszeitpunkte/16 Wochen)</b>								
140	5,70 (2,73)	3,46 (3,10)	71	6,32 (3,22)	5,66 (3,13)	-1,78 [-2,46; -1,11]	<0,001	-0,62 [-0,91; -0,32]
<b>Woche 1 – 48 (Anzahl Transfusionszeitpunkte/16 Wochen)</b>								
140	5,70 (2,73)	3,00 (3,10)	71	6,32 (3,22)	5,31 (3,10)	-1,95 [-2,67; -1,24]	<0,001	-0,64 [-0,93; -0,34]
Datenschnitt: 08.05.2018; Population: ITT-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline. Einheit: Transfusionszeitpunkte/16 Wochen <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> Anzahl Patienten mit nicht-fehlerndem Wert zu Studienbeginn und Auswertungszeitpunkt. <sup>(3)</sup> Bestes 16-Wochen-Intervall innerhalb des Analysezeitraums. <sup>(4)</sup> Ergebnisse für die Veränderung zur Baseline aus ANCOVA-Modell mit Behandlung und Stratifizierungsfaktoren als Kovariablen.  ANCOVA = <i>analysis of covariance</i> ; BSC = Best-Supportive-Care; ITT = Intention-to-Treat; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SMD = standardisierte Mittelwertdifferenz nach Hedges' g Quelle: [18]								

Neben einer Reduktion der Transfusionslast, d. h. der Menge der transfundierten EK-Einheiten, zeigte sich unter Luspatercept im Vergleich zu Placebo zudem eine deutliche, statistisch signifikante und auch klinisch relevante Reduktion der Transfusionshäufigkeit, also der Transfusionszeitpunkte pro 16 Wochen ( $p < 0,001$ ; siehe Tabelle 4-33). Bezogen auf einen transfusionsabhängigen Patienten, der beispielsweise alle drei Wochen eine EK-Transfusion erhalten hatte, entspricht die unter Luspatercept durchschnittlich erreichte Änderung der Transfusionshäufigkeit fast einer Halbierung der Transfusionshäufigkeit von etwa fünf Zeitpunkten pro 16 Wochen auf 3 Zeitpunkte pro 16 Wochen. Die Reduktion unter Placebo war hingegen mit einer Einsparung von knapp einem Transfusionszeitpunkt deutlich geringer.

Dies bedeutet, dass Patienten unter Luspatercept nicht nur die Menge an EK-Einheiten an sich deutlich und klinisch relevant reduzieren konnten, sondern dass zusätzlich auch die zeitliche Belastung durch eine Reduktion der Häufigkeit von Transfusionszeitpunkten gemildert werden konnte.

**Zusammenfassung der Ergebnisse zum Endpunkt Reduktion der Transfusionslast**

Die Reduktion der Transfusionslast wurde primär anhand der etablierten, klinisch relevanten Ansprechkriterien eines HI-E der IWG aus dem Jahr 2006 untersucht. Diese Kriterien berücksichtigen insbesondere, dass für Patienten mit hoher Transfusionslast bereits eine Reduktion der Transfusionslast einen patientenrelevanten Therapieerfolg darstellt. Patienten mit niedriger Transfusionslast wurden hingegen nur als Ansprecher bewertet, wenn eine effektive und anhaltende Verbesserung der Erythropoese, d. h. ein Anstieg des Hb-Spiegels um mindestens 1,5 g/dl über mindestens acht Wochen, mit einer Transfusionsfreiheit einher ging.

Der Anteil an Patienten, die ein HI-E erreichten, war unter Luspatercept mit 52,9 % gegenüber 11,8 % unter Placebo bedeutend höher, was einer fast fünffach höheren Wahrscheinlichkeit entspricht, dieses Therapieziel zu erreichen. Unabhängig von der Transfusionslast vor Behandlungsbeginn zeigte sich dieser deutliche Unterschied zugunsten von Luspatercept. Ergänzend dargestellte Analysen belegen zudem, dass mit Luspatercept behandelte Patienten statistisch signifikant schneller ein solches Ansprechen erreichten, so sprach ein Großteil der Patienten bereits innerhalb der ersten Wochen der Behandlung an. Weitere Auswertungen zur mittleren Reduktion der Transfusionslast und Transfusionshäufigkeit runden das Bild zu diesem Endpunkt weiter ab. Die deutlich größere Einsparung an notwendigen EK-Einheiten sowie an Transfusionszeitpunkten im Vergleich zu Placebo spiegeln die große Entlastung unter Luspatercept durch eine verminderte Transfusionstherapie wider.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

### **Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

### **Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Morbidität – Reduktion der Transfusionslast liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.1.4 Morbidität – Reduktion der Eisenüberladung

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-34: Operationalisierung des Endpunkts Morbidität – Reduktion der Eisenüberladung

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Die Reduktion der Eisenüberladung wurde gemessen anhand des klinisch relevanten Parameters sFerritin und dem Bedarf an Eisenchelatoren, gemessen als tägliche Dosis Eisenchelator.</p> <p>Der Baseline-sFerritin-Spiegel bzw. der Bedarf an Eisenchelatoren zur Baseline waren definiert als mittlerer sFerritin-Wert bzw. mittlere tägliche Dosis Eisenchelator in den 16 Wochen vor Behandlungsbeginn.</p> <p>Für Patienten, die keine Dosis Eisenchelator erhalten, wird der Bedarf auf 0 mg gesetzt.</p> <p>Mittlere Änderungen der Eisenparameter wurden mittels ANCOVA mit Behandlung und Stratifizierungsfaktoren (durchschnittliche Transfusionslast zur Baseline und IPSS-R-Score zur Baseline) als Kovariablen als mittlere Änderung in jeglichem, besten 16-Wochen-Intervall im Studienverlauf (sFerritin) bzw. dem 16-Wochen-Intervall von Woche 9 – 24 und 33 – 48 (Bedarf an Eisenchelatoren) berechnet. Nur Patienten mit Werten zu Studienbeginn und zum Auswertungszeitpunkt wurden in der Analyse berücksichtigt.</p> <p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Analyse 1:</b> Mittlere Änderung des sFerritin-Spiegels im Vergleich zur Baseline; Woche 1 – 24 und Woche 1 – 48 (siehe Tabelle 4-36)</li> <li>• <b>Analyse 2:</b> Mittlere Änderung der täglichen Dosis Eisenchelator im Vergleich zur Baseline; Woche 9 – 24 und Woche 33 – 48 (siehe Tabelle 4-37)</li> </ul>
ANCOVA = <i>analysis of covariance</i> ; KI = Konfidenzintervall; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; sFerritin = Serum-Ferritin	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-35: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Reduktion der Eisenüberladung in MEDALIST

Studie MEDALIST Endpunkt- kategorien	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
sFerritin	niedrig	ja	ja	ja	ja	niedrig
Bedarf an Eisenchelatoren	niedrig	ja	ja	ja	ja	niedrig
sFerritin = Serum-Ferritin						

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Morbidität – Reduktion der Eisenüberladung wurde verblindet erhoben. Das ITT-Prinzip wurde adäquat umgesetzt. Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Morbidität – Reduktion der Eisenüberladung und eine hohe Aussagesicherheit der Ergebnisse.

*Stellen Sie die Ergebnisse für den Endpunkt Morbidität – Reduktion der Eisenüberladung für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

**sFerritin-Spiegel**

Tabelle 4-36: Ergebnisse zum Endpunkt Morbidität – Reduktion der Eisenüberladung aus MEDALIST – Mittlere Änderung des sFerritin-Spiegels

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studienbeginn	MW (SD) zum Auswertungszeitpunkt <sup>(3)</sup>	N <sup>(2)</sup>	MW (SD) zu Studienbeginn	MW (SD) zum Auswertungszeitpunkt <sup>(3)</sup>	LS-MWD [95 %-KI]	p-Wert <sup>(4)</sup>	SMD [95 %-KI]
<b>Mittlere Änderung in Woche 1 – 24</b>								
153	1348,02 (971,24)	1265,12 (1034,15)	76	1503,84 (1242,94)	1593,11 (1352,36)	-169,05 [-303,60; -34,50]	0,014	-0,29 [-0,56; -0,01]
<b>Mittlere Änderung in Woche 1 – 48</b>								
153	1348,02 (971,24)	1164,69 (1000,66)	76	1503,84 (1242,94)	1524,25 (1279,34)	-209,30 [-343,28; -75,31]	0,002	-0,36 [-0,63; -0,08]
Datenschnitt: 08.05.2018; Population: ITT-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline. Einheit: sFerritin [ng/ml]/16 Wochen <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> Anzahl Patienten mit nicht-fehlendem Wert zu Studienbeginn und Auswertungszeitpunkt. <sup>(3)</sup> Bestes 16-Wochen-Intervall innerhalb des Analysezeitraums. <sup>(4)</sup> Ergebnisse für die Veränderung zur Baseline aus ANCOVA-Modell mit Behandlung und Stratifizierungsfaktoren als Kovariablen.  ANCOVA = <i>analysis of covariance</i> ; BSC = Best-Supportive-Care; ITT = Intention-to-Treat; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; sFerritin = Serum-Ferritin; SMD = standardisierte Mittelwertdifferenz nach Hedges' g Quelle: [18]								

Im Studienzeitraum der primären Behandlungsphase (Woche 1 – 24) zeigte sich hinsichtlich der Reduktion der Eisenüberladung, gemessen anhand des sFerritin-Spiegels, ein deutlicher und statistisch signifikanter Vorteil für Luspatercept im Vergleich zur Placebo-Behandlung ( $p = 0,014$ ). Während es unter Luspatercept zu einer Reduktion des sFerritin-Spiegels kam, zeigte sich im Placebo-Arm eine leichte Erhöhung des sFerritin-Spiegels. Der sFerritin-Spiegel ließ sich durch Luspatercept somit um durchschnittlich 169,05 ng/ml (95 %-KI: [-303,60; -34,50]) mehr senken als unter Placebo (siehe Tabelle 4-36).

Auch bei Betrachtung der langfristigen Effekte über den Studienzeitraum bis Woche 48 setzte sich dieser statistisch signifikante Vorteil weiter fort ( $p = 0,002$ ). Dies deutet darauf hin, dass bei längerer Behandlung der sFerritin-Spiegel weiter gesenkt wird und Luspatercept das Potential hat, langfristig eine Reduktion der Eisenüberladung zu erreichen und Siderosebedingte Morbidität und Mortalität zu reduzieren.

**Bedarf an Eisenchelatoren**

Tabelle 4-37: Ergebnisse zum Endpunkt Morbidität – Reduktion der Eisenüberladung aus MEDALIST – Mittlere Änderung der täglichen Dosis Eisenchelator

Studie MEDALIST							
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
N <sup>(2)</sup>	MW (SD) zu Studienbeginn <sup>(3)</sup>	LS-MW (SE) Veränderung zum Auswertungszeitpunkt	N <sup>(2)</sup>	MW (SD) zu Studienbeginn <sup>(3)</sup>	LS-MW (SE) Veränderung zum Auswertungszeitpunkt	LS-MWD (95 %-KI)	p-Wert <sup>(4)</sup>
<b>Mittlere Änderung in Woche 9 – 24</b>							
128	550,0 (991,09)	10,0 (29,25)	68	498,8 (1010,87)	51,0 (35,92)	-41,0 (-120,3; 38,2)	0,309
<b>Mittlere Änderung in Woche 33 – 48</b>							
78	550,0 (991,09)	-148,8 (46,13)	12	498,8 (1010,87)	-123,8 (92,19)	-24,9 (-210,7; 160,8)	0,7903
Datenschnitt: 08.05.2018; Population: ITT-Population; Einheit: tägliche Dosis Eisenchelator [mg]/16 Wochen <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> Anzahl Patienten mit nicht fehlendem Wert zu Studienbeginn und Auswertungszeitpunkt). <sup>(3)</sup> Studienbeginn ist definiert als durchschnittliche tägliche Dosis Eisenchelator [mg] in den 16 Wochen vor Behandlungsbeginn. <sup>(4)</sup> ANCOVA-Modell stratifiziert nach durchschnittlicher Transfusionslast und dem IPSS-R-Score zur Baseline.  ANCOVA = <i>analysis of covariance</i> ; BSC = Best-Supportive-Care; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz Quelle: [1, 18]							

Im Studienzeitraum der primären Behandlungsphase (Woche 1 – 24) zeigte sich hinsichtlich der Reduktion der Eisenüberladung, gemessen anhand der Änderung der täglichen Dosis Eisenchelator, ein numerischer Vorteil für Luspatercept im Vergleich zu Placebo. Die tägliche Dosis Eisenchelator stieg unter Placebo im Mittel um 51,0 mg (SE = 35,92), während es unter Luspatercept lediglich zu einer geringen Erhöhung um durchschnittlich 10,0 mg (SE = 29,25) kam. Dies entspricht einem um durchschnittlich 41,0 mg (95 %-KI: [-120,3; 38,2]) geringeren Anstieg der täglichen Dosis Eisenchelator unter Luspatercept als unter Placebo (siehe Tabelle 4-37). Da die Anpassung der Dosis Eisenchelator in Abhängigkeit der Transfusionslast und der damit einhergehenden Eisenüberladung stattfindet, entspricht eine geringere Dosis Eisenchelator einer geringeren Eisenbelastung des Patienten.

Auswertungen bis Woche 48 wurden ergänzend berichtet. Konsistent zu den Ergebnissen der Änderung der täglichen Dosis Eisenchelator in den ersten 24 Wochen der Behandlungsphase, zeigte sich auch in der Analyse bis Woche 48 ein numerischer Vorteil für Luspatercept. So konnte die tägliche Dosis Eisenchelator unter Luspatercept um 24,9 mg (95 %-KI: [-210,7; 160,8]) stärker gesenkt werden als unter Placebo. Langfristig kann der Einsatz von Eisenchelatoren durch Luspatercept somit potentiell eingespart werden, was wiederum einer Reduktion der Eisenüberladung an sich und einer Reduktion der Belastung durch die zusätzliche Therapie und damit verbundenen Nebenwirkungen entsprechen würde.

### **Zusammenfassung der Ergebnisse zum Endpunkt Reduktion der Eisenüberladung**

Die Auswertungen zum Endpunkt Reduktion der Eisenüberladung sind klinisch relevante Endpunkte im vorliegenden Anwendungsgebiet und werden daher ergänzend berichtet, jedoch nicht zur Ableitung eines Zusatznutzens herangezogen.

Insgesamt spiegeln die Ergebnisse das Potential von Luspatercept wider, durch Vermeidung von EK-Transfusionen die transfusionsbedingte Eisenüberladung und das damit einhergehende Risiko Siderose-bedingter Morbidität und Mortalität zu reduzieren. So wurde der sFerritin-Spiegel unter Luspatercept im Vergleich zu Placebo statistisch signifikant reduziert und auch hinsichtlich des Bedarfs an Eisenchelatoren zeichnet sich bereits ein positiver Trend ab.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

### **Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

### **Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Morbidität – Reduktion der Eisenüberladung liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.1.5 Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens

*Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-38: Operationalisierung des Endpunkts Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Die Symptomatik wurde anhand der Symptomskalen des EORTC QLQ-C30-Fragebogens (Version 3.0) erhoben:</p> <ul style="list-style-type: none"> <li>• Fatigue (Item 10, 12, 18)</li> <li>• Übelkeit und Erbrechen (Item 14, 15)</li> <li>• Schmerz (Item 9, 19)</li> <li>• Dyspnoe (Item 8)</li> <li>• Schlaflosigkeit (Item 11)</li> <li>• Appetitminderung (Item 13)</li> <li>• Obstipation (Item 16)</li> <li>• Durchfall (Item 17)</li> <li>• Finanzielle Schwierigkeiten (Item 28)</li> </ul> <p>Die Symptomatik wurde zur Baseline (Dosis 1 Tag 1), zu jedem zweiten Zyklus (Zyklus 1, 3, 5, 7) und der Woche-25-Visite, jeweils vor Verabreichung der Studienmedikation erhoben. Für Patienten, die nach der Woche 25-Visite weiterhin doppelblind behandelt wurden, wurde die Symptomatik weiterhin zu jedem zweiten Zyklus vor Verabreichung der Studienmedikation erhoben.</p> <p>In die Analysen zur Verbesserung der Symptomatik gehen nur solche Patienten ein, die zur Baseline und zu mindestens einem Auswertungszeitpunkt im Studienverlauf einen ausgefüllten EORTC QLQ-C30-Fragebogen aufwiesen (HRQoL-Population). Ein Fragebogen galt dann als ausgefüllt, wenn mindestens die Hälfte aller Fragen beantwortet wurden, d. h. mindestens 15 Fragen des EORTC QLQ-C30-Fragebogens ausgefüllt waren.</p> <p>Auswertungen zu den Rücklaufquoten des Fragebogens beziehen sich auf die ITT-Population. Höhere Antwortwerte entsprechen in den Symptomskalen des EORTC QLQ-C30 einer stärkeren Symptomatik.</p> <p>Skalen, für die weniger als die Hälfte der Items beantwortet waren, wurden entsprechend der Vorgaben im Handbuch des EORTC QLQ-C30-Fragebogens als fehlende Werte bewertet. Eine Imputation fehlender Werte wurde nicht durchgeführt.</p> <p>Als klinisch relevante Änderung wurde die vom G-BA anerkannte und im Studienprotokoll präspezifizierte MCID von zehn Punkten herangezogen.</p> <p>In den Responder-Analysen (Hauptanalyse) wurden Unterschiede zwischen den Behandlungsarmen mittels CMH-Methode, stratifiziert nach Transfusionslast zur Baseline und IPSS-R-Score zur Baseline, berechnet und als RR, OR und ARR, jeweils inkl. 95 %-KI, dargestellt. Das OR wurde mittels logistischem Regressionsmodell mit Behandlungsarm als Haupteffekt berechnet. Für die Berechnung des RR und der 95 %-KI wurde eine Normalverteilung angenommen. Patienten mit fehlenden Werten zu Woche 48 wurden als Patienten ohne klinisch relevante Verbesserung der Lebensqualität gewertet (Non-Responder).</p>

Studie	Operationalisierung
	<p>Ereigniszeitanalysen, die Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in Woche 1 – 48, werden ergänzend dargestellt. Patienten ohne klinisch relevante Verbesserung wurden zum Zeitpunkt der Analyse zensiert. Die mediane Zeit bis zum Erreichen einer klinisch relevanten Verbesserung wurde mittels Kaplan-Meier-Methode geschätzt. Kaplan-Meier-Schätzer für die mediane Zeit bis zur klinisch relevanten Verbesserung sowie das HR, berechnet mittels Cox-Modell mit Behandlungsarm und Baseline-Werten als Kovariate, jeweils inkl. 95 %-KI werden dargestellt.</p> <p>Änderungen der Symptomatik zu den Erhebungszeitpunkten im Vergleich zur Baseline werden grafisch dargestellt, inkl. tabellarischer, deskriptiver Listung der Werte im Studienverlauf. Unterschiede zwischen den Behandlungsarmen wurden mittels ANOVA berechnet. Statistisch signifikante Unterschiede (<math>p &lt; 0,05</math>) wurden in den Grafiken markiert.</p> <p>Änderungen der Symptomatik im Studienverlauf wurden mittels MMRM berechnet. In das Modell flossen Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und ein Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte ein. SMD inkl. 95 %-KI in Form von Hedges' g werden ergänzend dargestellt, um klinisch relevante Effekte zu bewerten.</p> <p>Auswertungen zu den Datenschnitten aus dem Jahr 2019 werden nicht dargestellt, da die Rücklaufquoten sehr gering waren (insbesondere im Placebo-Arm hatten <math>&lt; 10</math> Patienten einen ausgefüllten Fragebogen) und somit weitere Aussagen zum Zusatznutzen von Luspatercept nicht möglich sind.</p> <p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Hauptanalyse:</b> Anteil der Patienten mit einer Verbesserung der Symptomatik um mindestens zehn Punkte; Woche 1 – 24 (siehe Tabelle 4-41)</li> <li>• <b>Zusatzanalyse 1:</b> Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte; Woche 1 – 48 (siehe Tabelle 4-42)</li> <li>• <b>Zusatzanalyse 2:</b> Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten; Woche 1 – 48 (siehe Abbildung 4-24 ff. &amp; Tabelle 4-43)</li> <li>• <b>Zusatzanalyse 3:</b> Mittlere Änderung der Symptomatik im Studienverlauf (MMRM-Analyse); Woche 1 – 24 und Woche 1 – 48 (siehe Tabelle 4-44 f.)</li> </ul>
<p>ARR = absolute Risikodifferenz; CMH = Cochran-Mantel-Haenszel; EORTC QLQ-C30 = <i>European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module</i>; HRQoL-Population = Population mit auswertbarer gesundheitsbezogener Lebensqualität; KI = Konfidenzintervall; MCID = minimale klinische Relevanzschwelle; MMRM = gemischtes Modell für wiederholte Messungen (<i>mixed-model repeated measure</i>); OR = Odds Ratio; RR = Relatives Risiko; SMD = standardisierte Mittelwertdifferenz</p>	

*Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.*

*Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.*

Tabelle 4-39: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens in MEDALIST

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
MEDALIST	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens wurde verblindet erhoben. In die Analyse gingen nur solche Patienten ein, die den EORTC QLQ-C30-Fragebogen zu Studienbeginn und zu mindestens einem Zeitpunkt im Studienverlauf ausgefüllt hatten (HRQoL-Population). Zwischen der HRQoL-Population und der ITT-Population liegen keine wesentlichen Unterschiede hinsichtlich der Baseline-Charakteristika vor [1, 95]. Darüber hinaus umfasst die HRQoL-Population > 95 % der Patienten der ITT-Population (Luspatercept: 149/153 (97,4 %); Placebo: 76/76 (100,0 %); siehe Tabelle 4-6), sodass von einer guten Vergleichbarkeit der HRQoL- und ITT-Population auszugehen ist. Die Rücklaufquoten bezogen auf die ITT-Population liegen für den EORTC QLQ-C30-Fragebogen bis Woche 25 über den vom G-BA geforderten 70 % und sind damit ausreichend hoch, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen (siehe Tabelle 4-40). Ergebnisse nach Woche 25 wurden aufgrund geringerer Rücklaufquoten (< 70 %) nur ergänzend bis Woche 48 berichtet. Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens und eine hohe Aussage-sicherheit der Ergebnisse.

Tabelle 4-40: Rücklaufquoten des EORTC QLQ-C30-Fragebogens aus MEDALIST

Studie MEDALIST Rücklaufquoten des EORTC QLQ-C30-Frage- bogens	Luspatercept + BSC <sup>(1)</sup> N = 153		Placebo + BSC <sup>(1)</sup> N = 76	
	Patienten unter Therapie zum jeweiligen Erhebungs- zeitpunkt	Patienten mit ausgefülltem Fragebogen zum jeweiligen Erhebungs- zeitpunkt <sup>(2)</sup> n (%)	Patienten unter Therapie zum jeweiligen Erhebungs- zeitpunkt	Patienten mit ausgefülltem Fragebogen zum jeweiligen Erhebungs- zeitpunkt <sup>(2)</sup> n (%)
Baseline / Zyklus 1	153	150 (98,0)	76	76 (100,0)
Woche 7 / Zyklus 3	151	136 (88,9)	76	70 (92,1)
Woche 13 / Zyklus 5	146	132 (86,3)	73	66 (86,8)
Woche 19 / Zyklus 7	134	112 (73,2)	70	60 (78,9)
Woche 24	127	112 (73,2)	68	54 (71,1)
Zyklus 1 der EP	100	51 (33,3)	23	9 (11,8)
Zyklus 3 der EP	94	82 (53,6)	19	16 (21,1)
Zyklus 5 der EP	89	80 (52,3)	16	15 (19,7)
Zyklus 7 der EP	86	75 (49,0)	14	13 (17,1)
Zyklus 9 der EP	75	68 (44,4)	10	8 (10,5)
Zyklus 11 der EP	60	52 (34,0)	7	7 (9,2)
Zyklus 13 der EP	47	43 (28,1)	5	5 (6,6)
Zyklus 15 der EP	33	25 (16,3)	4	4 (5,3)
Zyklus 17 der EP	29	25 (16,3)	3	2 (2,6)
Zyklus 19 der EP	18	15 (9,8)	1	1 (1,3)
Zyklus 21 der EP	12	12 (7,8)	1	1 (1,3)
Zyklus 23 der EP	5	5 (3,3)	1	1 (1,3)
Zyklus 25 der EP	3	3 (2,0)	0	0
Zyklus 27 der EP	2	2 (1,3)	0	0
Zyklus 29 der EP	2	2 (1,3)	0	0
Zyklus 31 der EP	2	2 (1,3)	0	0
EOT	71	47 (30,7)	54	28 (36,8)

Datenschnitt: 08.05.2018; Population: ITT-Population  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> Prozentualer Anteil bezogen auf die ITT-Population.

BSC = Best-Supportive-Care; EOT = *end of treatment*; EP = Extensionsphase; ITT = Intention-to-Treat  
Quelle: [1, 95]

Stellen Sie die Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

### Hauptanalyse: Anteil der Patienten mit einer Verbesserung der Symptomatik um mindestens zehn Punkte

Tabelle 4-41: Ergebnisse zum Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Anteil der Patienten mit einer Verbesserung der Symptomatik um mindestens zehn Punkte; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI]	RR [95 %-KI] <sup>(3)</sup>	ARR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Symptomskala des EORTC QLQ-C30- Fragebogens	Patienten mit Ereignis <sup>(2)</sup> n/N (%)	Patienten mit Ereignis <sup>(2)</sup> n/N (%)				
Fatigue	32/149 (21,5)	24/76 (31,6)	0,59 (0,32; 1,10)	0,68 (0,43; 1,07)	-10,1 (-22,5; 2,3)	0,099
Übelkeit und Erbrechen	17/149 (11,4)	5/76 (6,6)	1,83 (0,65; 5,16)	1,73 (0,67; 4,52)	4,8 (-2,7; 12,4)	0,254
Schmerz	25/149 (16,8)	14/76 (18,4)	0,89 (0,43; 1,84)	0,91 (0,50; 1,65)	-1,6 (-12,2; 8,9)	0,758
Dyspnoe	24/149 (16,1)	16/76 (21,1)	0,72 (0,36; 1,45)	0,77 (0,43; 1,35)	-4,9 (-15,8; 6,0)	0,360
Schlaflosigkeit	28/149 (18,8)	18/76 (23,7)	0,75 (0,38; 1,46)	0,79 (0,47; 1,34)	-4,9 (-16,3; 6,5)	0,390
Appetit- minderung	22/149 (14,8)	9/76 (11,8)	1,29 (0,56; 2,96)	1,25 (0,60; 2,57)	2,9 (-6,3; 12,2)	0,548
Obstipation	31/149 (20,8)	13/76 (17,1)	1,27 (0,62; 2,61)	1,22 (0,68; 2,18)	3,7 (-7,0; 14,4)	0,509
Durchfall	11/149 (7,4)	6/76 (7,9)	0,93 (0,33; 2,62)	0,94 (0,36; 2,43)	-0,5 (-7,9; 6,9)	0,891
Finanzielle Schwierigkeiten	13/149 (8,7)	9/76 (11,8)	0,71 (0,29; 1,75)	0,74 (0,33; 1,65)	-3,1 (-11,7; 5,4)	0,458
Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> Patienten mit Ereignis waren definiert als Patienten mit einer Verbesserung der Symptomatik um mindestens zehn Punkte; Patienten mit fehlenden Werten zu Woche 24 wurden als Patienten ohne Ereignis gezählt. <sup>(3)</sup> RR und ARR wurden mittels Normal-Approximation berechnet. <sup>(4)</sup> OR und p-Wert berechnet mittels logistischem Regressionsmodell mit Behandlungsarm als Haupteffekt.  ARR = Absolute Risikodifferenz; HRQoL = <i>health-related quality of life</i> ; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko Quelle: [19]						

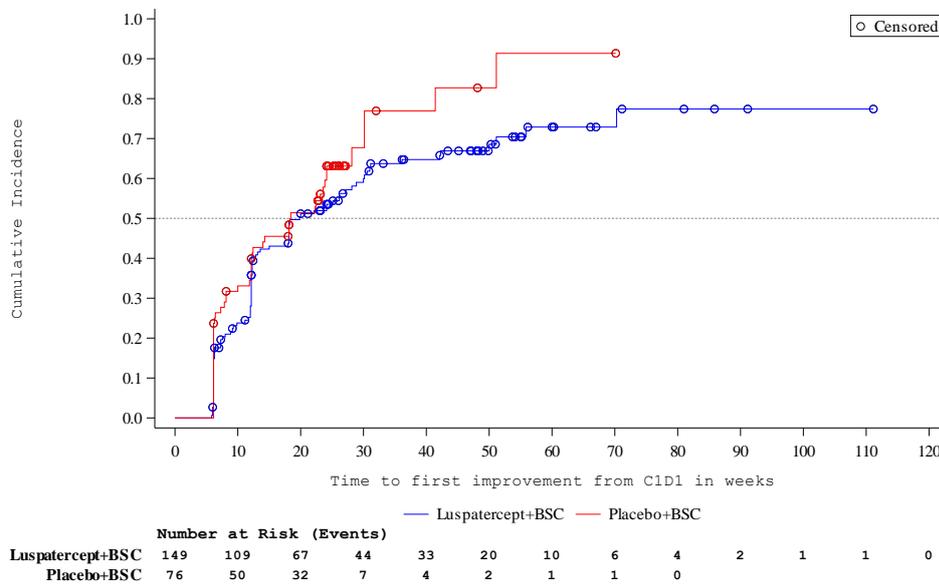
In den ersten 24 Wochen der Behandlungsphase zeigte sich für den Anteil der Patienten mit einer klinisch relevanten Verbesserung der Symptomatik, gemessen als Verbesserung der Symptomskalen des EORTC QLQ-C30 um mindestens zehn Punkte, kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen. Über alle Symptomskalen hinweg, traten klinisch relevante Verbesserungen in beiden Behandlungsarmen zu etwa gleichen Teilen auf. Dies trifft für alle erhobenen Symptome zu (siehe Tabelle 4-41).

### Zusatzanalyse 1: Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte

Tabelle 4-42: Zusatzanalyse 1: Ergebnisse zum Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte; Woche 1 – 48

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Symptomskala des EORTC QLQ-C30- Fragebogens	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>
Fatigue	94/149 (63,1)	19,9 (13,1; 28,9)	50/76 (65,8)	18,4 (12,1; 24,1)	0,886 (0,622; 1,261)	0,501
Übelkeit und Erbrechen	28/149 (18,8)	N.E. (N.E.; N.E.)	11/76 (14,5)	N.E. (N.E.; N.E.)	1,059 (0,506; 2,221)	0,879
Schmerz	55/149 (36,9)	N.E. (78,0; N.E.)	27/76 (35,5)	42,1 (30,1; N.E.)	1,069 (0,665; 1,721)	0,782
Dyspnoe	61/149 (40,9)	73,3 (31,0; N.E.)	38/76 (50,0)	24,1 (18,1; N.E.)	0,782 (0,510; 1,198)	0,258
Schlaflosigkeit	59/149 (39,6)	N.E. (30,0; N.E.)	33/76 (43,4)	36,9 (24,1; 85,1)	0,938 (0,604; 1,456)	0,774
Appetit- minderung	37/149 (24,8)	N.E. (N.E.; N.E.)	20/76 (26,3)	N.E. (N.E.; N.E.)	0,464 (0,255; 0,844)	0,012
Obstipation	52/149 (34,9)	N.E. (N.E.; N.E.)	20/76 (26,3)	N.E. (30,1; N.E.)	1,024 (0,599; 1,751)	0,930
Durchfall	25/149 (16,8)	N.E. (N.E.; N.E.)	14/76 (18,4)	N.E. (N.E.; N.E.)	0,745 (0,363; 1,529)	0,422
finanzielle Schwierigkeiten	22/149 (14,8)	N.E. (N.E.; N.E.)	11/76 (14,5)	N.E. (N.E.; N.E.)	1,384 (0,623; 3,075)	0,425

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
Symptomskala des EORTC QLQ-C30- Fragebogens	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline (1) BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). (2) Schätzung nach Kaplan-Meier-Methode. (3) HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell mit Behandlung und Baseline-Wert als Kovariable. (4) Der p-Wert des HR basiert auf einem stratifizierten Log-Rank-Test.  BSC = Best-Supportive-Care; HR = Hazard-Ratio; HRQoL = <i>health-related quality of life</i> ; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar Quelle: [19]						

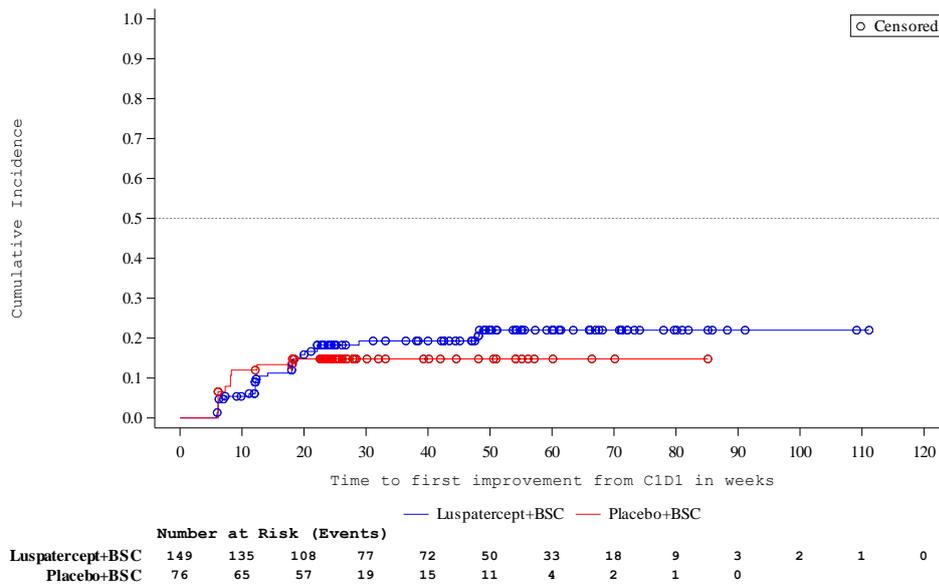


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-15: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Fatigue; Woche 1 – 48

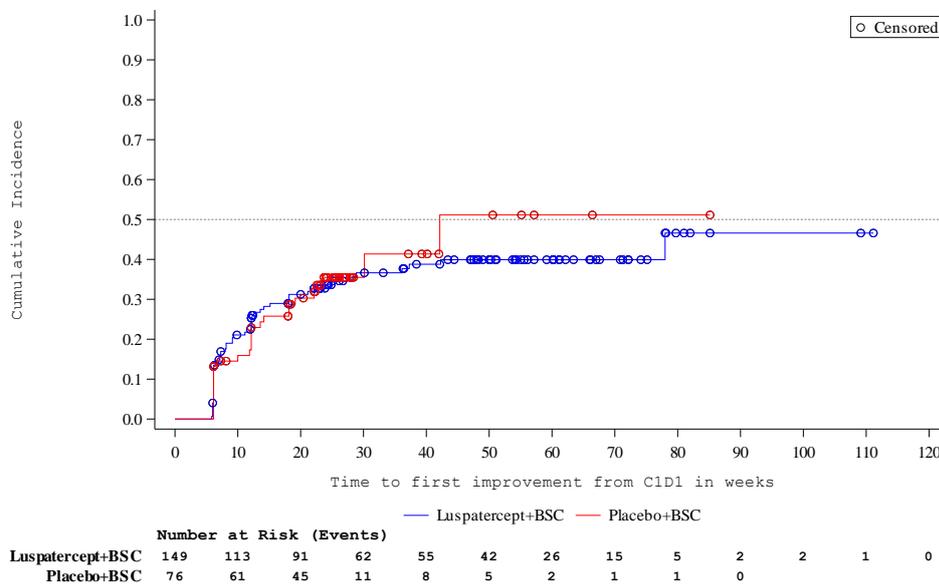


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-16: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Übelkeit und Erbrechen; Woche 1 – 48

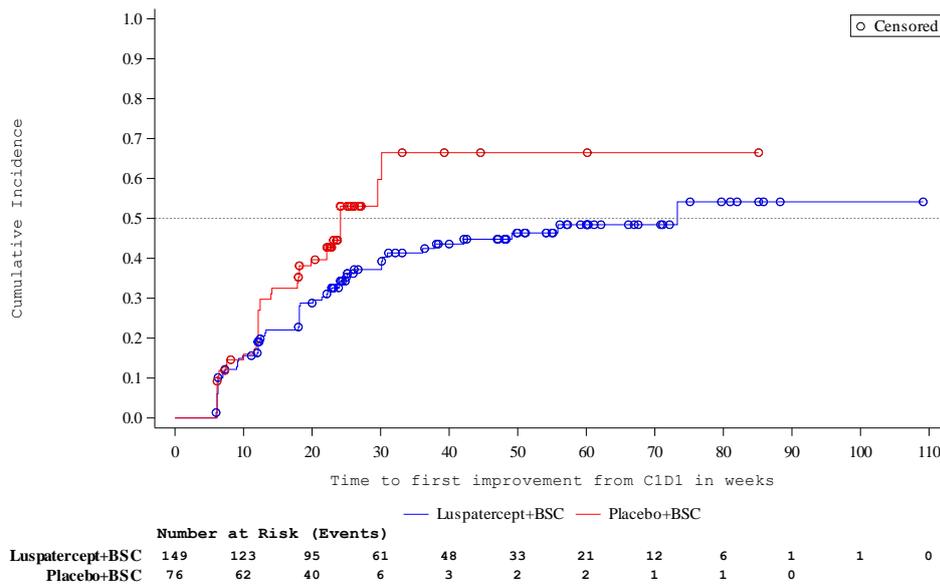


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-17: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Schmerz; Woche 1 – 48

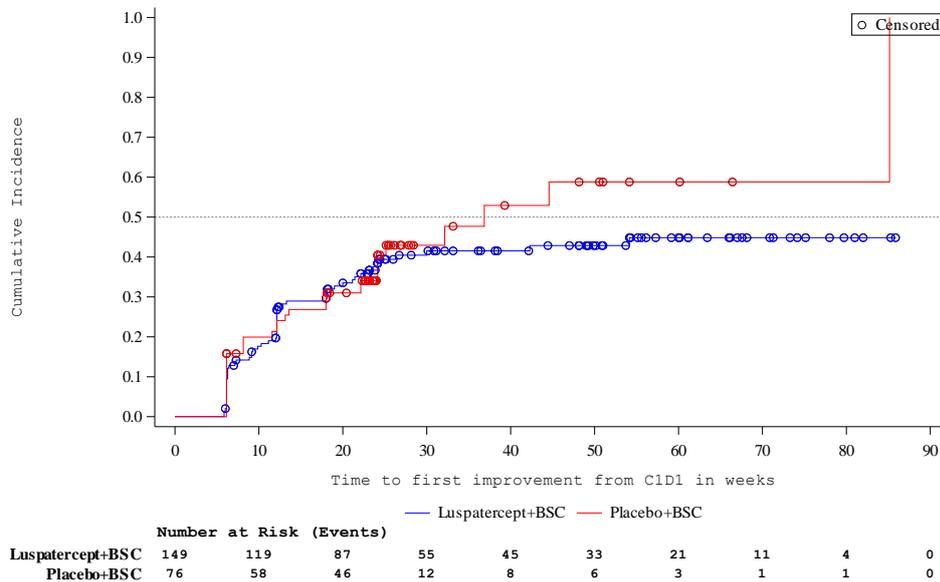


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-18: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Dyspnoe; Woche 1 – 48

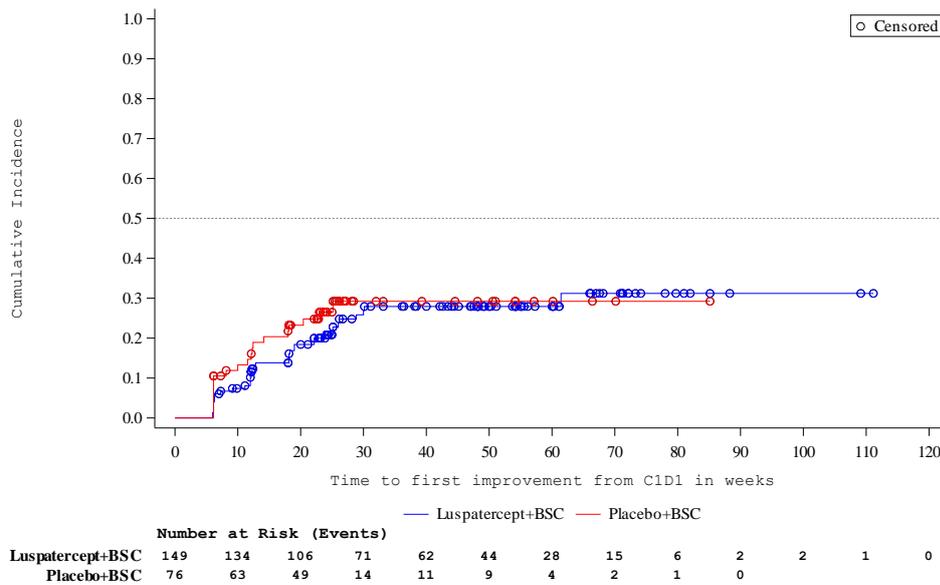


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-19: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Schlaflosigkeit; Woche 1 – 48

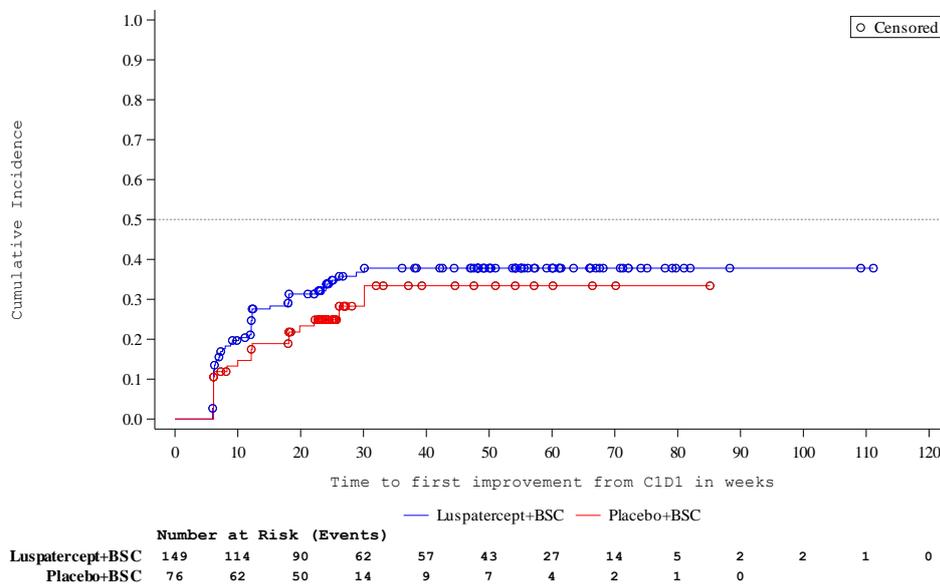


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-20: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Appetitminderung; Woche 1 – 48

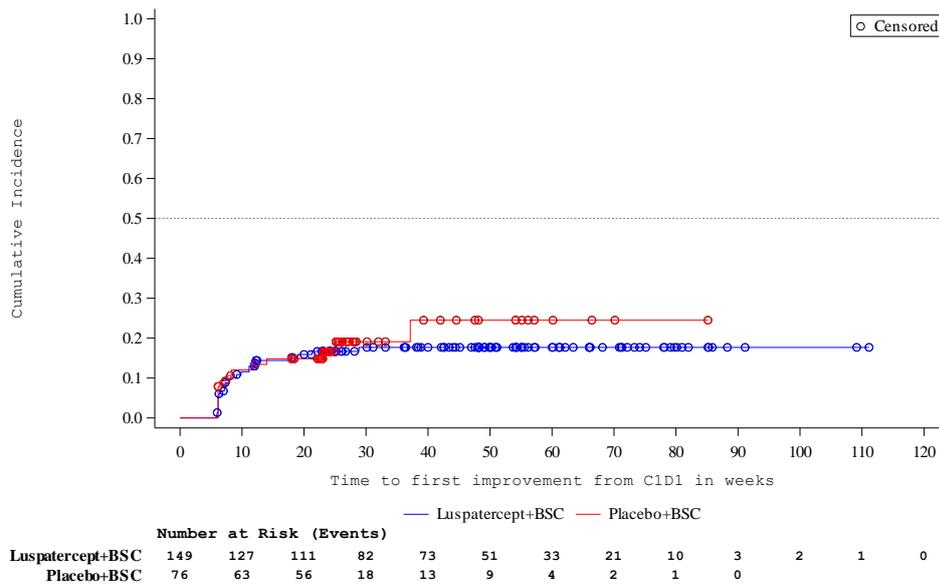


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-21: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Obstipation; Woche 1 – 48

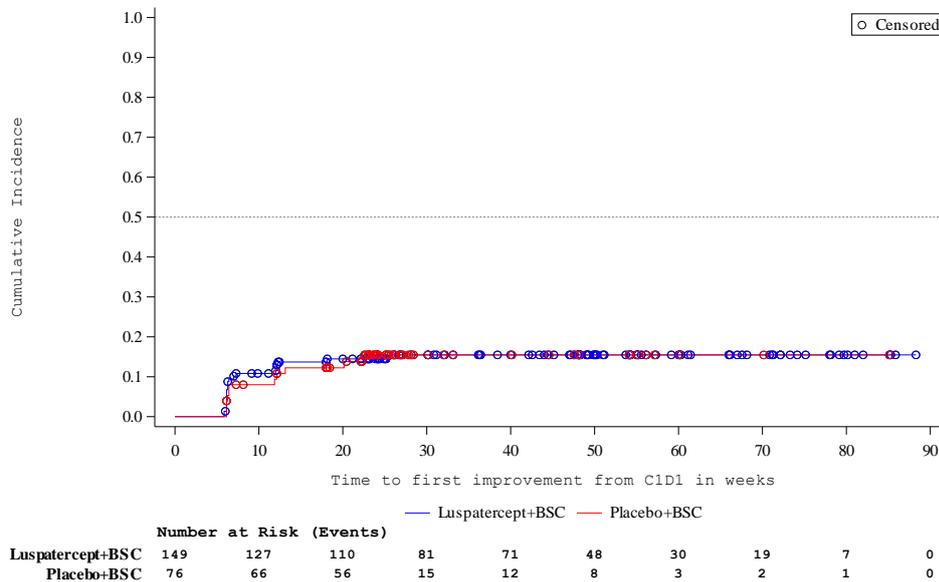


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-22: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Diarrhö; Woche 1 – 48



Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

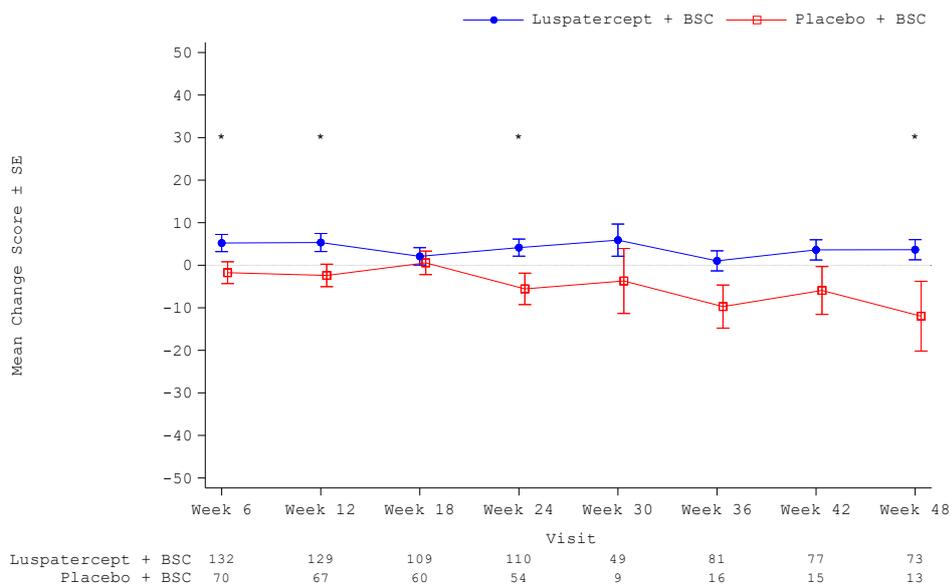
Quelle: [19]

Abbildung 4-23: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der Symptomatik um mindestens zehn Punkte in MEDALIST – Finanzielle Schwierigkeiten; Woche 1 – 48

Auswertungen über den Studienzeitraum bis Woche 48 werden aufgrund der Unsicherheit hinsichtlich der Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten sowie Rücklaufquoten < 70 % nach Woche 24 ergänzend als Ereigniszeitanalysen dargestellt.

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit einer Verbesserung der Symptomatik um mindestens zehn Punkte in den ersten 24 Wochen der Behandlungsphase, zeigten sich auch in der Analyse der Zeit bis zum Erreichen einer klinisch relevanten Verbesserung der Symptomatik um mindestens zehn Punkte – mit Ausnahme der Symptomskala Appetitminderung – keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen. Der Anteil der Patienten mit einer klinisch relevanten Verbesserung der Symptomatik war in beiden Behandlungsarmen für alle Symptomskalen etwa gleich groß, die mediane Zeit bis zur Verbesserung der Symptomatik wurde in den meisten Fällen nicht erreicht (siehe Tabelle 4-42 & Abbildung 4-15 ff.).

### Zusatzanalyse 2: Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten

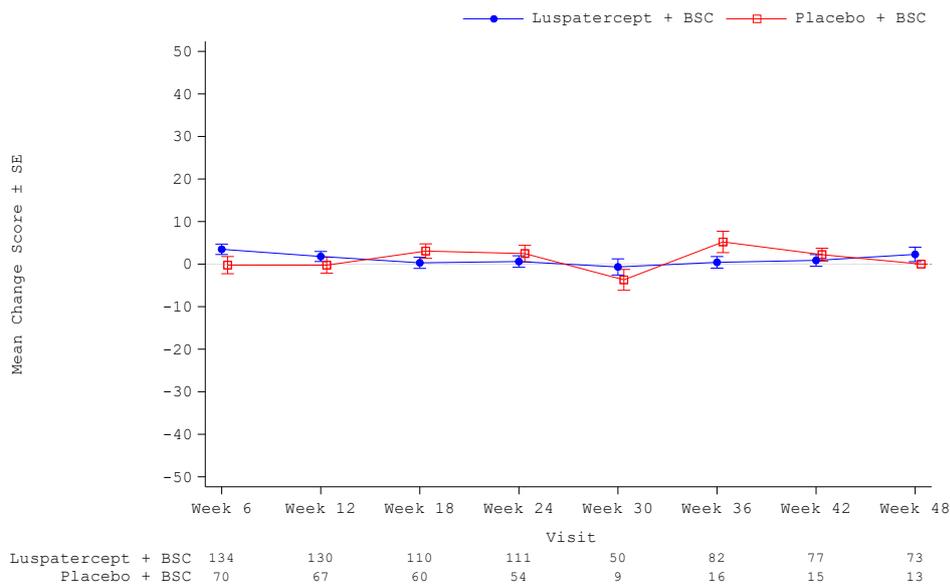


Datenschnitt: 08.05.2018;

\* = statistisch signifikanter Effekt, berechnet mittels ANOVA (*analysis of variance*)

Quelle: [19]

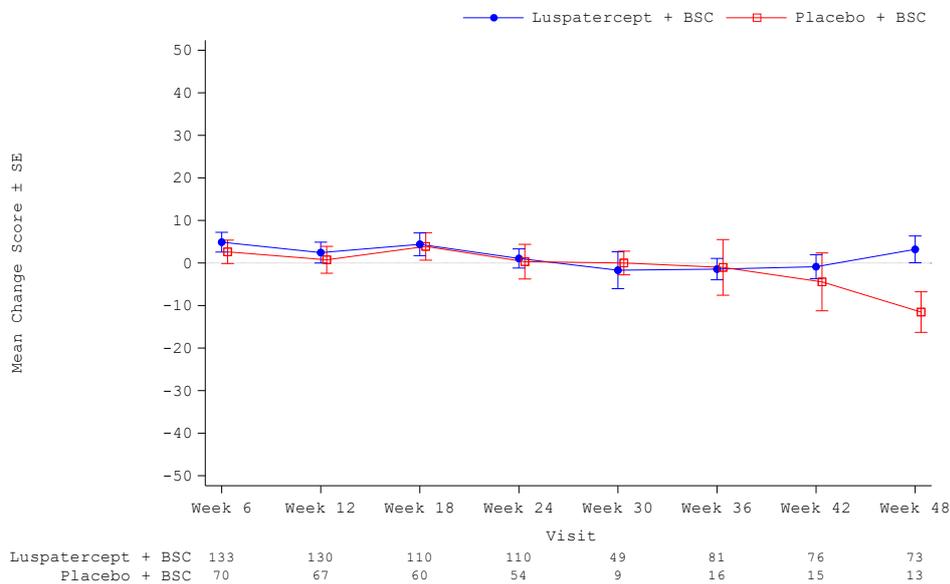
Abbildung 4-24: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Fatigue; Woche 1 – 48



Datenschnitt: 08.05.2018

Quelle: [19]

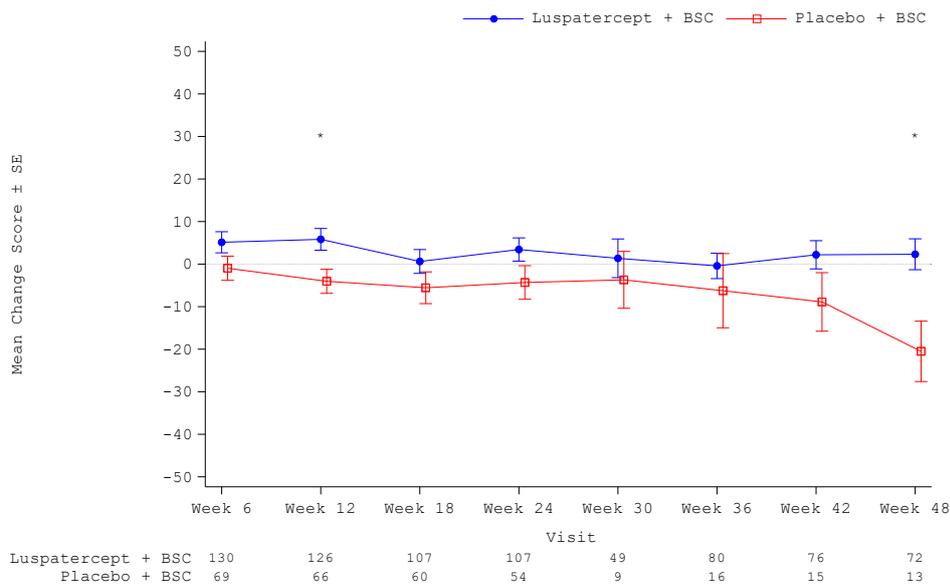
Abbildung 4-25: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Übelkeit und Erbrechen; Woche 1 – 8



Datenschnitt: 08.05.2018

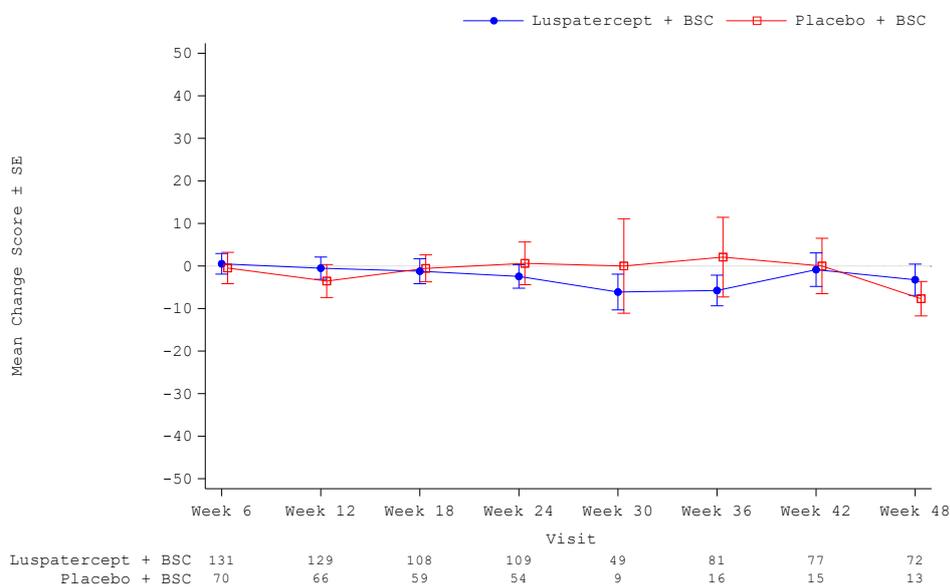
Quelle: [19]

Abbildung 4-26: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Schmerz; Woche 1 – 48



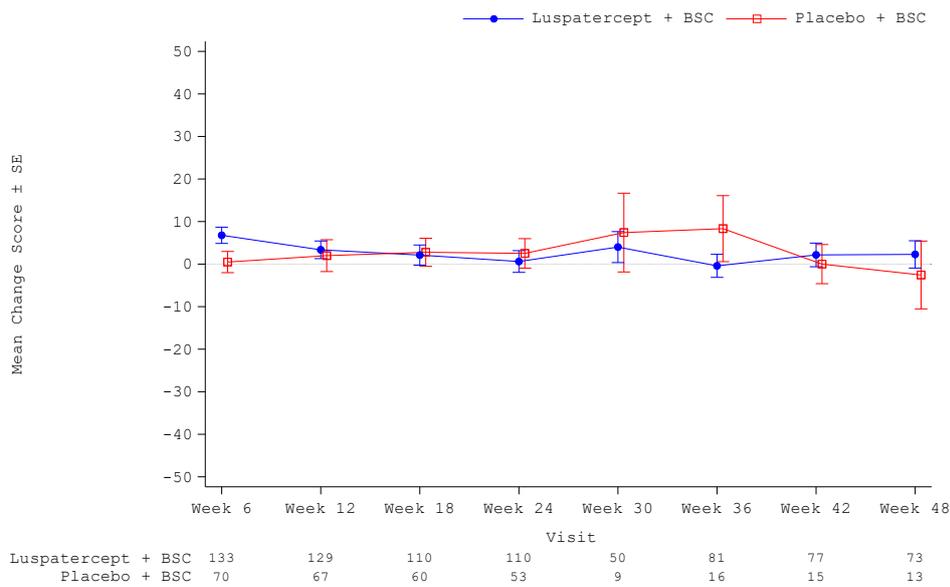
Datenschnitt: 08.05.2018; \* = statistisch signifikanter Effekt, berechnet mittels ANOVA (*analysis of variance*)  
 Quelle: [19]

Abbildung 4-27: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Dyspnoe; Woche 1 – 48



Datenschnitt: 08.05.2018  
 Quelle: [19]

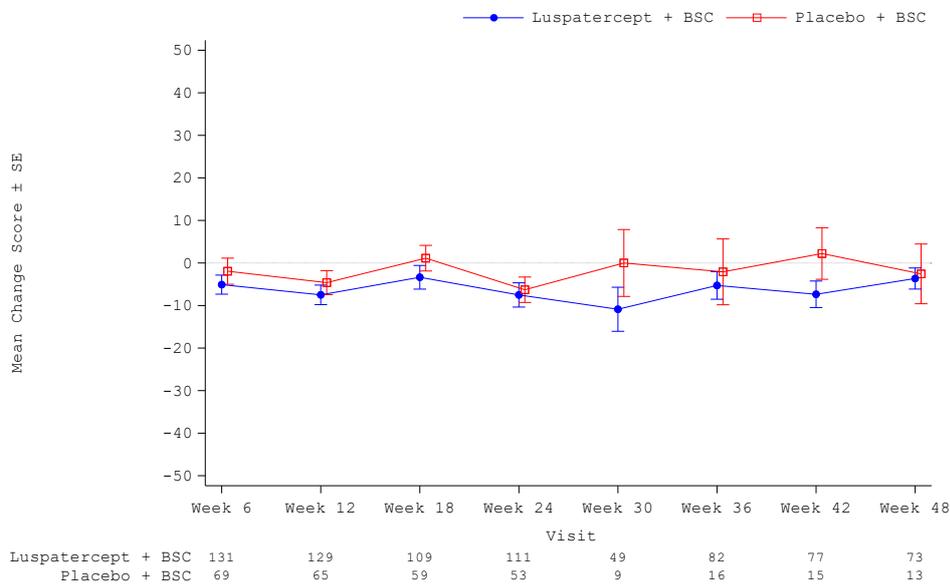
Abbildung 4-28: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Schlaflosigkeit; Woche 1 – 48



Datenschnitt: 08.05.2018

Quelle: [19]

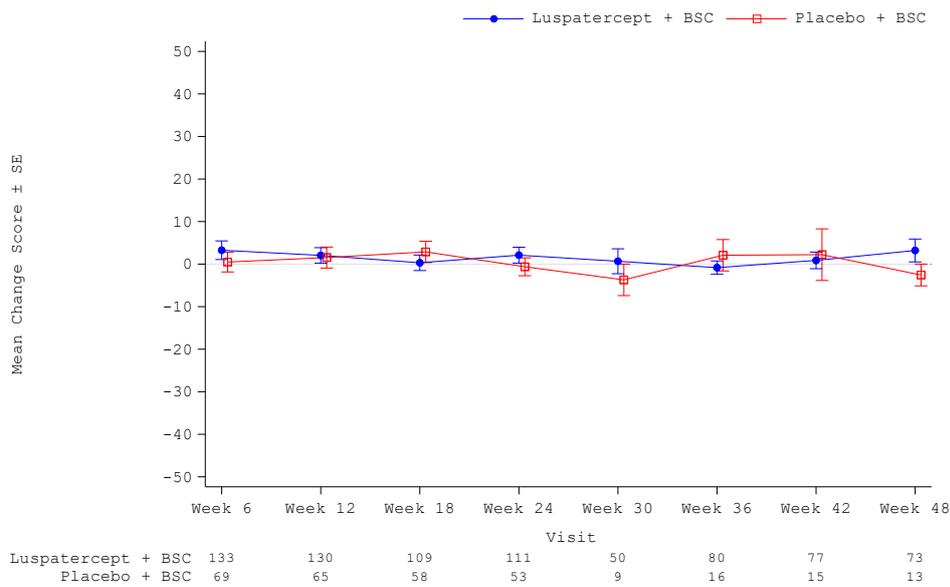
Abbildung 4-29: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Appetitminderung; Woche 1 – 48



Datenschnitt: 08.05.2018

Quelle: [19]

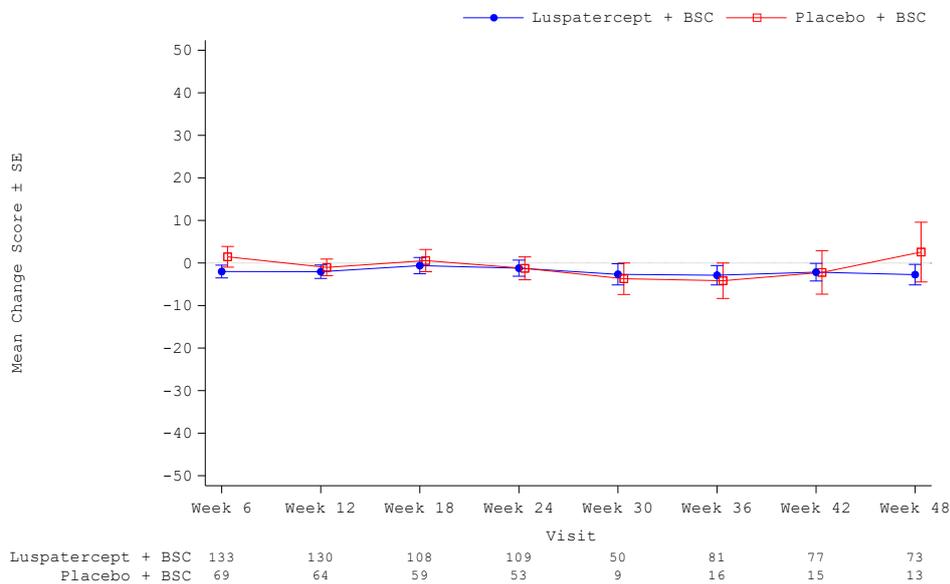
Abbildung 4-30: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Obstipation; Woche 1 – 48



Datenschnitt: 08.05.2018

Quelle: [19]

Abbildung 4-31: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Durchfall; Woche 1 – 48



Datenschnitt: 08.05.2018

Quelle: [19]

Abbildung 4-32: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik zu den Erhebungszeitpunkten – Finanzielle Schwierigkeiten; Woche 1 – 48

Tabelle 4-43: Zusatzanalyse 2: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Werte der Symptomatik zu den Erhebungszeitpunkten; Woche 1 – 48

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
<b><i>Fatigue</i></b>				
Baseline	148 (99,3)	41,89 (24,80)	76 (100)	45,18 (24,70)
Woche 6	133 (89,3)	45,61 (27,04)	70 (92,1)	44,29 (25,41)
Woche 12	130 (87,2)	46,75 (26,83)	67 (88,2)	42,21 (23,87)
Woche 18	110 (73,8)	42,37 (27,06)	60 (78,9)	43,70 (25,87)
Woche 24	111 (74,5)	43,19 (24,30)	54 (71,1)	38,48 (23,91)
Woche 30	50 (33,6)	44,56 (25,21)	9 (11,8)	34,57 (11,71)
Woche 36	82 (55,0)	42,07 (26,89)	16 (21,1)	34,03 (27,06)
Woche 42	77 (51,7)	43,29 (25,97)	15 (19,7)	34,81 (21,77)
Woche 48	73 (49,0)	42,31 (33,33)	13 (17,1)	35,04 (26,00)
<b><i>Übelkeit und Erbrechen</i></b>				
Baseline	149 (100)	4,92 (11,87)	76 (100)	5,48 (13,98)
Woche 6	134 (89,9)	7,71 (14,26)	70 (92,1)	5,71 (10,19)
Woche 12	130 (87,2)	7,05 (13,43)	67 (88,2)	4,98 (9,63)
Woche 18	110 (73,8)	5,30 (12,35)	60 (78,9)	6,67 (13,45)
Woche 24	111 (74,5)	5,71 (12,00)	54 (71,1)	6,79 (14,67)
Woche 30	50 (33,6)	5,33 (9,18)	9 (11,8)	0,00 (0,00)
Woche 36	82 (55,0)	5,08 (9,33)	16 (21,1)	5,21 (10,03)
Woche 42	77 (51,7)	5,41 (11,61)	15 (19,7)	2,22 (5,86)
Woche 48	73 (49,0)	6,62 (12,95)	13 (17,1)	0,00 (0,00)
<b><i>Schmerz</i></b>				
Baseline	148 (99,3)	18,47 (25,43)	76 (100)	19,96 (23,57)
Woche 6	134 (89,9)	23,63 (24,55)	70 (92,1)	21,90 (28,17)
Woche 12	131 (87,9)	20,87 (26,34)	67 (88,2)	20,40 (28,10)
Woche 18	110 (73,8)	21,52 (27,12)	60 (78,9)	24,44 (27,35)
Woche 24	111 (74,5)	17,57 (21,65)	54 (71,1)	21,30 (28,12)
Woche 30	50 (33,6)	18,67 (23,24)	9 (11,8)	20,37 (23,24)
Woche 36	82 (55,0)	17,68 (21,01)	16 (21,1)	19,79 (20,38)
Woche 42	77 (51,7)	16,45 (24,70)	15 (19,7)	14,44 (17,67)
Woche 48	73 (49,0)	19,18 (25,71)	13 (17,1)	8,97 (14,62)

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
<b><i>Dyspnoe</i></b>				
Baseline	146 (98,0)	33,79 (29,03)	76 (100)	39,47 (30,17)
Woche 6	133 (89,3)	37,09 (31,42)	69 (90,8)	37,20 (30,00)
Woche 12	129 (86,6)	37,21 (31,08)	66 (86,8)	33,84 (27,11)
Woche 18	110 (73,8)	30,00 (27,45)	60 (78,9)	31,67 (27,05)
Woche 24	109 (73,2)	32,72 (29,39)	54 (71,1)	31,48 (27,79)
Woche 30	50 (33,6)	36,00 (32,88)	9 (11,8)	25,93 (32,39)
Woche 36	82 (55,0)	32,11 (28,42)	16 (21,1)	27,08 (34,89)
Woche 42	77 (51,7)	32,47 (31,05)	15 (19,7)	26,67 (28,73)
Woche 48	73 (49,0)	31,96 (29,10)	13 (17,1)	20,51 (21,68)
<b><i>Schlaflosigkeit</i></b>				
Baseline	148 (99,3)	27,03 (30,70)	76 (100)	28,07 (30,33)
Woche 6	132 (88,6)	27,53 (29,84)	70 (92,1)	28,57 (30,18)
Woche 12	130 (87,2)	26,15 (30,78)	66 (86,8)	24,24 (25,85)
Woche 18	109 (73,2)	25,38 (31,06)	59 (77,6)	25,42 (25,02)
Woche 24	109 (73,2)	23,85 (29,78)	54 (71,1)	27,16 (27,53)
Woche 30	49 (32,9)	25,85 (29,86)	9 (11,8)	33,33 (33,33)
Woche 36	82 (55,0)	24,39 (29,19)	16 (21,1)	31,25 (25,73)
Woche 42	77 (51,7)	29,87 (31,34)	15 (19,7)	24,44 (23,46)
Woche 48	72 (48,3)	25,93 (29,19)	13 (17,1)	12,82 (16,88)
<b><i>Appetitminderung</i></b>				
Baseline	149 (100)	15,21 (24,96)	76 (100)	13,16 (23,14)
Woche 6	133 (89,3)	20,55 (26,18)	70 (92,1)	14,29 (20,91)
Woche 12	129 (86,6)	18,09 (24,65)	67 (88,2)	15,92 (26,82)
Woche 18	110 (73,8)	17,27 (26,22)	60 (78,9)	15,00 (24,10)
Woche 24	110 (73,8)	15,45 (22,43)	53 (69,7)	13,84 (23,96)
Woche 30	50 (33,6)	16,67 (22,59)	9 (11,8)	14,81 (29,40)
Woche 36	81 (54,4)	14,81 (22,36)	16 (21,1)	18,75 (27,13)
Woche 42	77 (51,7)	17,32 (25,14)	15 (19,7)	11,11 (20,57)
Woche 48	73 (49,0)	15,53 (26,11)	13 (17,1)	10,26 (21,01)
<b><i>Obstipation</i></b>				
Baseline	149 (100)	19,24 (27,99)	75 (98,7)	14,22 (24,61)
Woche 6	131 (87,9)	13,99 (24,79)	70 (92,1)	13,33 (24,34)
Woche 12	129 (86,6)	12,14 (21,63)	66 (86,8)	11,11 (22,14)

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
Woche 18	109 (73,2)	14,68 (26,62)	60 (78,9)	13,89 (20,62)
Woche 24	111 (74,5)	12,61 (22,04)	54 (71,1)	8,64 (16,10)
Woche 30	49 (32,9)	14,97 (23,63)	9 (11,8)	14,81 (24,22)
Woche 36	82 (55,0)	13,01 (22,68)	16 (21,1)	14,58 (20,97)
Woche 42	77 (51,7)	9,96 (17,99)	15 (19,7)	17,78 (30,52)
Woche 48	73 (49,0)	10,05 (19,79)	13 (17,1)	12,82 (21,68)
<b>Durchfall</b>				
Baseline	149 (100)	8,95 (19,23)	75 (98,7)	9,33 (16,94)
Woche 6	133 (89,3)	11,78 (22,92)	70 (92,1)	9,52 (16,19)
Woche 12	130 (87,2)	11,03 (20,07)	66 (86,8)	12,63 (20,06)
Woche 18	109 (73,2)	7,65 (18,50)	59 (77,6)	14,12 (20,72)
Woche 24	111 (74,5)	8,41 (19,83)	54 (71,1)	8,64 (17,36)
Woche 30	50 (33,6)	6,67 (15,06)	9 (11,8)	3,70 (11,11)
Woche 36	80 (53,7)	5,00 (15,09)	16 (21,1)	6,25 (13,44)
Woche 42	77 (51,7)	7,36 (18,44)	15 (19,7)	6,67 (18,69)
Woche 48	73 (49,0)	10,05 (22,00)	13 (17,1)	0,00 (0,00)
<b>finanzielle Schwierigkeiten</b>				
Baseline	149 (100)	11,19 (23,45)	75 (98,7)	10,22 (21,20)
Woche 6	133 (89,3)	9,52 (20,34)	70 (92,1)	10,95 (23,89)
Woche 12	130 (87,2)	10,26 (22,29)	65 (85,5)	8,72 (17,95)
Woche 18	108 (72,5)	12,35 (23,94)	60 (78,9)	9,44 (20,44)
Woche 24	109 (73,2)	10,40 (20,13)	54 (71,1)	10,49 (21,30)
Woche 30	50 (33,6)	9,33 (20,25)	9 (11,8)	0,00 (0,00)
Woche 36	81 (54,4)	10,70 (25,18)	16 (21,1)	4,17 (11,39)
Woche 42	77 (51,7)	11,26 (25,71)	15 (19,7)	6,67 (13,80)
Woche 48	73 (49,0)	9,13 (19,45)	13 (17,1)	12,82 (25,60)
Datenschnitt: 08.05.2018; Population: HRQoL-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i> ; MW = Mittelwert; SD = Standardabweichung Quelle: [19]				

Die mittlere Änderung der Symptomatik zu den einzelnen Erhebungszeitpunkten ist für die einzelnen Symptomskalen des EORTC QLQ-C30 über den gesamten Studienverlauf bis Woche 48 grafisch dargestellt. Infolge unterschiedlich langer Behandlungszeiten und Rücklaufquoten < 70 % nach Woche 25 ist die Interpretierbarkeit der Ergebnisse limitiert.

Insgesamt bestätigen die grafischen Darstellungen der mittleren Änderung der Symptomatik die Ergebnisse der Hauptanalyse. Für keine der Symptomskalen zeigte sich ein über die Behandlungsphase bis Woche 24 konstant statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Abbildung 4-24 – Abbildung 4-28 und Tabelle 4-43). Die Änderungen der Symptomatik sind innerhalb der beiden Behandlungsarme im Vergleich zur Baseline vorwiegend als geringfügige Effekte einzustufen (mittlere Änderung < 10 Punkte [96]).

### Zusatzanalyse 3: Mittlere Änderung der Symptomatik im Studienverlauf (MMRM-Analyse)

Tabelle 4-44: Zusatzanalyse 3: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik im Studienverlauf (MMRM-Analyse); Woche 1 – 24

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>Fatigue</b>								
111	41,89 (24,80)	43,19 (24,30)	54	45,18 (24,70)	38,48 (23,91)	6,84 (0,55; 13,12)	0,033	0,28 (0,00; 0,55)
<b>Übelkeit und Erbrechen</b>								
111	4,92 (11,87)	5,71 (12,00)	54	5,48 (13,98)	6,79 (14,67)	-0,62 (-4,45; 3,21)	0,749	-0,05 (-0,32; 0,23)
<b>Schmerz</b>								
111	18,47 (25,43)	17,57 (21,65)	54	19,96 (23,57)	21,30 (28,12)	-0,55 (-7,15; 6,05)	0,870	-0,02 (-0,30; 0,25)
<b>Dyspnoe</b>								
109	33,79 (29,03)	32,72 (29,39)	54	39,47 (30,17)	31,48 (27,79)	5,58 (-2,00; 13,17)	0,148	0,19 (-0,09; 0,47)

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>Schlaflosigkeit</b>								
109	27,03 (30,70)	23,85 (29,78)	54	28,07 (30,33)	27,16 (27,53)	-0,79 (-8,21; 6,64)	0,834	-0,03 (-0,30; 0,25)
<b>Appetitminderung</b>								
110	15,21 (24,96)	15,45 (22,43)	53	13,16 (23,14)	13,84 (23,96)	0,22 (-6,99; 7,42)	0,953	0,01 (-0,27; 0,28)
<b>Obstipation</b>								
111	19,24 (27,99)	12,61 (22,04)	54	14,22 (24,61)	8,64 (16,10)	3,92 (-2,38; 10,22)	0,221	0,15 (-0,13; 0,42)
<b>Durchfall</b>								
111	8,95 (19,23)	8,41 (19,83)	54	9,33 (16,94)	8,64 (17,36)	-1,02 (-6,17; 4,12)	0,695	-0,06 (-0,33; 0,22)
<b>Finanzielle Schwierigkeiten</b>								
109	11,19 (23,45)	10,40 (20,13)	54	10,22 (21,20)	10,49 (21,30)	0,60 (-4,24; 5,44)	0,807	0,03 (-0,25; 0,30)
<p>Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline.</p> <p><sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).</p> <p><sup>(2)</sup> Anzahl Patienten mit nicht-fehlendem Wert zu Woche 24.</p> <p><sup>(3)</sup> Werte zu Studienbeginn beziehen sich auf die gesamte HRQoL-Population (N = 149 im Luspatercept-Arm; N = 76 im Placebo-Arm).</p> <p><sup>(4)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarme, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.</p> <p>BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i>; IPSS-R = <i>International Prognostic Scoring System-revised</i>; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate (<i>least square</i>); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz nach Hedges' g</p> <p>Quelle: [19]</p>								

Tabelle 4-45: Zusatzanalyse 3: Ergebnisse für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der Symptomatik im Studienverlauf (MMRM-Analyse); Woche 1 – 48

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>Fatigue</b>								
73	41,89 (24,80)	43,19 (24,30)	13	45,18 (24,70)	38,48 (23,91)	9,62 (0,66; 18,58)	0,036	0,39 (0,11; 0,67)
<b>Übelkeit und Erbrechen</b>								
73	4,92 (11,87)	5,71 (12,00)	13	5,48 (13,98)	6,79 (14,67)	1,56 (-3,60; 6,73)	0,550	0,12 (-0,15; 0,40)
<b>Schmerz</b>								
73	18,47 (25,43)	17,57 (21,65)	13	19,96 (23,57)	21,30 (28,12)	0,26 (-9,84; 10,37)	0,959	0,01 (-0,27; 0,29)
<b>Dyspnoe</b>								
73	33,79 (29,03)	32,72 (29,39)	13	39,47 (30,17)	31,48 (27,79)	6,80 (-4,49; 18,08)	0,237	0,23 (-0,05; 0,51)
<b>Schlaflosigkeit</b>								
72	27,03 (30,70)	23,85 (29,78)	13	28,07 (30,33)	27,16 (27,53)	-0,55 (-10,77; 9,67)	0,915	-0,02 (-0,29; 0,26)
<b>Appetitminderung</b>								
73	15,21 (24,96)	15,45 (22,43)	13	13,16 (23,14)	13,84 (23,96)	0,04 (-10,49; 10,56)	0,995	0,00 (-0,27; 0,28)
<b>Obstipation</b>								
73	19,24 (27,99)	12,61 (22,04)	13	14,22 (24,61)	8,64 (16,10)	0,36 (-8,26; 8,98)	0,935	0,01 (-0,26; 0,29)
<b>Durchfall</b>								
73	8,95 (19,23)	8,41 (19,83)	13	9,33 (16,94)	8,64 (17,36)	1,20 (-6,51; 8,91)	0,759	0,06 (-0,21; 0,34)
<b>Finanzielle Schwierigkeiten</b>								
73	11,19 (23,45)	10,40 (20,13)	13	10,22 (21,20)	10,49 (21,30)	-0,45 (-7,80; 6,90)	0,905	-0,02 (-0,30; 0,26)

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<p>Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline.</p> <p><sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).</p> <p><sup>(2)</sup> Anzahl Patienten mit nicht-fehlerndem Wert zu Woche 24.</p> <p><sup>(3)</sup> Werte zu Studienbeginn beziehen sich auf die gesamte HRQoL-Population (N = 149 im Luspatercept-Arm; N = 76 im Placebo-Arm).</p> <p><sup>(4)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarme, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.</p> <p>BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i>; IPSS-R = <i>International Prognostic Scoring System-revised</i>; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate (<i>least square</i>); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz nach Hedges' g</p> <p>Quelle: [19]</p>								

Auch bei Betrachtung der Ergebnisse zur mittleren Änderung der Symptomatik unter Berücksichtigung aller Werte im Studienverlauf ergaben sich für die Symptomskalen des EORTC QLQ-C30, mit Ausnahme der Symptomskala Fatigue, keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen. Dies trifft sowohl auf die primäre Behandlungsphase bis Woche 24 als auch die weitere Behandlung bei den zu Woche 25 selektierten Patienten bis Woche 48 zu (siehe Tabelle 4-44 & Tabelle 4-45).

Einzig für die Symptomskala Fatigue zeigte sich ein statistisch signifikanter Unterschied zuungunsten von Luspatercept (LS-MWD(Woche 1 – 24) = 6,84, 95 %-KI: [0,55; 13,12], p = 0,033; LS-MWD(Woche 1 – 48) = 9,62; 95 %-KI: [0,66; 18,58]; p = 0,036), der jedoch nicht klinisch relevant war (SMD(Woche 1 – 24) = 0,28; 95 %-KI: [0,00; 0,55]; SMD(Woche 1 – 48) = 0,39; 95 %-KI: [0,11; 0,67]).

Insgesamt liegen somit konsistent zur Hauptanalyse keine relevanten Unterschiede zwischen den beiden Behandlungsarmen hinsichtlich einer Änderung der Symptomatik vor.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

### **Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

### **Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.1.6 Morbidität – Hospitalisierungen

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-46: Operationalisierung des Endpunkts Morbidität – Hospitalisierungen

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Hospitalisierungen wurden erfasst als Gesamthospitalisierungen (Hospitalisierungen jeglicher Ursache) und Hospitalisierungen aufgrund von UE und/oder Transfusionen, die seit Randomisierung bis zum Auswertungszeitpunkt aufgetreten sind.</p> <p>Unterschiede zwischen den Behandlungsarmen werden mittels CMH-Tests, stratifiziert nach durchschnittlicher Transfusionslast und IPSS-R-Risikoklassifizierung zur Baseline, ermittelt. Dargestellt werden die Ergebnisse als OR, stratifiziert berechnet mittels Normal-Approximation, RR und ARR, unstratifiziert berechnet mittels Normal-Approximation, jeweils inkl. zugehörigem 95 %-KI.</p> <p>Auswertungen der Zeit bis zur ersten Hospitalisierung während der Behandlungs- und Extensionsphase bis zum Datenschnitt vom 01.07.2019 werden ergänzend dargestellt. Patienten, die bis zum Datenschnitt vom 01.07.2019 nicht stationär aufgenommen wurden, wurden zensiert.</p> <p>Die mediane Zeit bis zur Hospitalisierung wurde mittels Kaplan-Meier-Methode geschätzt und inkl. zugehörigem 95 %-KI, welche über eine Log-Log-Transformation nach Brookmeyer und Crowley berechnet wurden, dargestellt. Unterschiede zwischen den Behandlungsarmen wurden mittels stratifiziertem Log-Rank-Test analysiert. Das HR inklusive zweiseitigem 95 %-KI für Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, wurde anhand eines stratifizierten Cox-Proportional-Hazards-Modells mit Behandlung als Kovariable geschätzt.</p> <p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Hauptanalyse:</b> Anteil der Patienten mit mindestens einer Hospitalisierung; Woche 1 – 24 (siehe Tabelle 4-48)</li> <li>• <b>Sensitivitätsanalyse:</b> Anteil der Patienten mit mindestens einer Hospitalisierung aufgrund von UE und/oder Transfusionen; Woche 1 – 24 (siehe Tabelle 4-50)</li> <li>• <b>Zusatzanalyse 1:</b> Zeit bis zur ersten Hospitalisierung; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-52)</li> </ul>
ARR = absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran-Mantel-Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; UE = unerwünschtes Ereignis	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-47: Bewertung des Verzerrungspotenzials für den Endpunkt Morbidität – Hospitalisierungen in MEDALIST

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
MEDALIST	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Morbidität – Hospitalisierungen wurde verblindet erhoben. Das ITT-Prinzip wurde adäquat umgesetzt. Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Morbidität – Hospitalisierungen und eine hohe Aussagesicherheit der Ergebnisse.

*Stellen Sie die Ergebnisse für den Endpunkt Morbidität – Hospitalisierungen für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

**Hauptanalyse: Anteil der Patienten mit mindestens einer Hospitalisierung**

Tabelle 4-48: Ergebnisse für den Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Anteil der Patienten mit mindestens einer Hospitalisierung; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(3)</sup>	ARR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Endpunkt	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)				
Hospitalisierung jeglicher Ursache	34/153 (22,2)	17/76 (22,4)	0,99 [0,50; 1,95]	0,99 [0,60; 1,64]	-0,17 [-11,41; 11,08]	0,977
Datenschnitt: 08.05.2018; Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittlicher Transfusionslast zur Baseline. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR berechnet mittels Normal-Approximation stratifiziert nach Stratifizierungsfaktoren. <sup>(3)</sup> RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(4)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.  ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; ITT = Intention-to-Treat; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko Quelle: [18]						

Während der ersten 24 Wochen der Behandlungsphase zeigte sich kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen. 22,2 % (34/153) der Patienten im Luspatercept-Arm und 22,4 % (17/76) der Patienten im Placebo-Arm wurden während der primären Studienphase stationär aufgenommen (siehe Tabelle 4-48).

Hauptgrund der stationären Aufnahme war das Auftreten eines UE. Andere Ursachen, darunter auch Transfusionen, traten lediglich vereinzelt auf (siehe Tabelle 4-49).

Tabelle 4-49: Ergebnisse für den Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Gründe für Hospitalisierungen; Woche 1 – 24

<b>Studie MEDALIST Gründe für Hospitalisierungen n (%)</b>	<b>Luspatercept + BSC<sup>(1)</sup> N = 153</b>	<b>Placebo + BSC<sup>(1)</sup> N = 76</b>
UE	31 (20,3)	13 (17,1)
elektive Behandlung einer vorbestehenden Erkrankung	1 (0,7)	3 (3,9)
Nicht Protokoll-bedingte Untersuchung	1 (0,7)	2 (2,6)
Behandlung, die vor Unterzeichnung der Einverständniserklärung geplant war	2 (1,3)	0
Protokoll-bedingte Behandlung	1 (0,7)	0
Soziale, technische oder praktische Gründe in Abwesenheit eines UE	1 (0,7)	0
Transfusion	0	1 (1,3)
Datenschnitt: 08.05.2018; Population: ITT-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).  BSC = Best-Supportive-Care; ITT = Intention-to-Treat; UE = unerwünschtes Ereignis Quelle: [18]		

***Sensitivitätsanalyse: Anteil der Patienten mit mindestens einer Hospitalisierung aufgrund von UE und/oder Transfusionen***

Tabelle 4-50: Sensitivitätsanalyse: Ergebnisse zum Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Anteil der Patienten mit mindestens einer Hospitalisierung aufgrund von UE und/oder Transfusionen; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept +BSC vs. Placebo + BSC <sup>(1)</sup>			
Endpunkt	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(3)</sup>	ARR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Hospitalisierung aufgrund von UE und/oder Transfusionen	31/153 (20,3)	13/76 (17,1)	1,24 [0,60; 2,58]	1,19 [0,67; 2,11]	3,20 [-7,26; 13,66]	0,560
Datenschnitt: 08.05.2018; Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittlicher Transfusionslast zur Baseline. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR berechnet mittels Normal-Approximation stratifiziert nach Stratifizierungsfaktoren. <sup>(3)</sup> RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(4)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.  ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; ITT = Intention-to-Treat; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; UE = unerwünschtes Ereignis Quelle: [18]						

Auch bei Betrachtung der Ergebnisse zu Hospitalisierungen aufgrund von UE und/oder Transfusionen zeigte sich konsistent zur Hauptanalyse kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (RR = 1,19; 95 %-KI: [0,67; 2,11]; siehe Tabelle 4-50).

Häufigste Ursache einer Hospitalisierung aufgrund von UE waren Ereignisse der SOC „Infektionen und parasitäre Erkrankungen“ (siehe Tabelle 4-51). Für den Großteil der Hospitalisierungen aufgrund eines UE wurde vom Prüfarzt nicht von einem Zusammenhang mit der Studienmedikation ausgegangen; bei den meisten Hospitalisierungen ist das zugrundeliegende UE zudem ohne Folgeschäden abgeklungen.

Tabelle 4-51: Ergebnisse für den Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Hospitalisierungen aufgrund von UE auf Ebene der SOC und PT

SOC PT	Luspatercept + BSC <sup>(2)</sup>			Placebo + BSC <sup>(2)</sup>		
	Grad 1-2	Grad 3-5	Status	Grad 1-2	Grad 3-5	Status
<b>SOC: Infektionen und parasitäre Erkrankungen</b>	<b>6 (1)</b>	<b>15 (8)</b>		<b>1 (1)</b>	<b>7 (5)</b>	
PT: Sepsis	0	2 [4, 5]	nicht abgeklungen/ tödlich	0	(1 [4])	(abgeklungen)
PT: Pneumonie	2 [2]	1 [3] (3 [3])	Folgeschaden/ abgeklungen/ (nicht abgeklungen)	0	2 [3]	abgeklungen
PT: Pneumonie durch Staphylokokken	0	1 [4]	nicht abgeklungen	0	0	/
PT: Septischer Schock	0	1 [4]	nicht abgeklungen	0	(1 [3])	(abgeklungen)
PT: Endokarditis	0	1 [4]	nicht abgeklungen	0	0	/
PT: Infektion lokalisiert	0	1 [3]	Folgeschaden	0	0	/
PT: Staphylokokken-Infektion	0	1 [3]	nicht abgeklungen	0	0	/
PT: Weichteilinfektion	0	1 [3]	nicht abgeklungen	0	0	/
PT: Abdominelle Infektion	1 [2]	0	abgeklungen	0	0	/
PT: Gastroenteritis	0	1 [3]	abgeklungen	0	0	/
PT: Orchitis	0	1 [3]	abgeklungen	0	0	/
PT: Harnwegsinfektion	0 (1 [2])	1 [3] (1 [3])	abgeklungen (nicht abgeklungen)	(1 [2])	1 [3]	abgeklungen
PT: Harnwegsinfektion durch Escherichia	0	1 [3]	abgeklungen	0	0	/
PT: Infektion der unteren Atemwege	1 [2]	0	abgeklungen	0	0	/
PT: Infektion der oberen Atemwege	0	1 [3]	abgeklungen	0	0	/
PT: Pharyngitis	1 [2]	1 [3]	abgeklungen	0	0	/
PT: Cellulitis	1 [2]	0	abgeklungen	0	1 [3]	abgeklungen

SOC PT	Luspatercept + BSC <sup>(2)</sup>			Placebo + BSC <sup>(2)</sup>		
	Grad 1-2	Grad 3-5	Status	Grad 1-2	Grad 3-5	Status
PT: Urosepsis	0	(1 [4])	(abgeklungen)	0	1 [4]	nicht abgeklungen
PT: Divertikulitis	0	0	/	0	1 [3]	abgeklungen
PT: Bakteriämie	0	0	/	1 [2]	0	abgeklungen
PT: Infektion mit Metapneumo- virus	0	0	/	0	1 [3]	abgeklungen
PT: Epididymitis	0	(1 [3])	(abgeklungen)	0	0	/
PT: Lungen- infektion	0	(1 [3])	(nicht abgeklungen)	0	0	/
PT: Bronchitis	0	(1 [3])	(abgeklungen)	0	0	/
PT: Bakterielle Harnwegs- infektion	0	0	/	0	(1 [3])	(abgeklungen)
PT: Infektiöse Enterokolitis	0	0	/	0	(1 [3])	(abgeklungen)
PT: Strepto- kokken-Infektion	0	0	/	0	(1 [3])	(abgeklungen)
<b>SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Kompli- kationen</b>	<b>2 (2)</b>	<b>11 (3)</b>		<b>0 (1)</b>	<b>4 (3)</b>	
PT: Sturz	1 [1] (2 [2])	2 [3] (1 [3])	abgeklungen	0	(2 [3])	(abgeklungen)
PT: Femurfraktur	1 [2]	1 [3] (1 [3])	abgeklungen	0	0	/
PT: Fraktur der Hüfte	0	1 [3]	abgeklungen	0	2 [3] (1 [3])	nicht abgeklungen/ Folgeschaden (abgeklungen)
PT: Fraktur des Schlüsselbeins	0	1 [3]	abgeklungen	0	0	/
PT: Humerus- fraktur	0	1 [3]	abgeklungen	0	1 [3]	Folgeschaden
PT: Gelenks- luxation	0	1 [3]	abgeklungen	0	0	/
PT: Kopf- verletzung	0	1 [3]	abgeklungen	0	0	/
PT: Verletzung des Skelets	0	1 [3]	abgeklungen	0	0	/
PT: Gelenks- verletzung	0	1 [3]	abgeklungen	0	0	/

SOC PT	Luspatercept + BSC <sup>(2)</sup>			Placebo + BSC <sup>(2)</sup>		
	Grad 1-2	Grad 3-5	Status	Grad 1-2	Grad 3-5	Status
PT: Verletzung der Wirbelsäule	0	1 [3]	abgeklungen	0	0	/
PT: Fraktur eines Brustwirbels	0	0	/	0	1 [3]	abgeklungen
PT: Knöchelfraktur	0	0	/	(1 [2])	0	(nicht abgeklungen)
PT: Rippenfraktur	0	(1 [3])	(nicht abgeklungen)	0	0	/
<b>SOC: Herzerkrankungen</b>	<b>2 (3)</b>	<b>5 (4)</b>		<b>0</b>	<b>1</b>	
PT: Herzinsuffizienz	0 (1 [2])	2 [3]	abgeklungen/ nicht abgeklungen	0	0	/
PT: Atrioventrikulärer Block	0	2 [3, 4] (1 [3])	abgeklungen	0	0	/
PT: Angina pectoris	1 [2] (1 [1])	0 (1 [3])	abgeklungen	0	0	/
PT: Aortenklappenstenose	0	1 [3]	abgeklungen	0	0	/
PT: Vorhofflimmern	1 [2] (1 [2])	0	abgeklungen	0	0	/
PT: Herzinfarkt	0	0	/	0	1 [3]	abgeklungen
PT: Tachykardie supraventrikulär	0	(1 [3])	(abgeklungen)	0	0	/
PT: Koronare Herzerkrankung	0	(1 [3])	(abgeklungen)			
<b>SOC: Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen</b>	<b>4 (2)</b>	<b>3 (1)</b>		<b>0</b>	<b>0</b>	
PT: Rheumatoide Arthritis	2 [2]	0	abgeklungen	0	0	/
PT: Rückenschmerzen	0 (1 [2])	1 [3]	abgeklungen	0	0	/
PT: Muskuläre Schwäche	0	1 [3]	abgeklungen	0	0	/
PT: Polymyalgia rheumatica	1 [2]	0	Folgeschaden			
PT: Myositis	0	1 [3]	abgeklungen	0	0	/
PT: Chondritis	1 [2]	0	abgeklungen	0	0	/
PT: Flankenschmerz	(1 [2])	0	(nicht abgeklungen)	0	0	/

SOC PT	Luspatercept + BSC <sup>(2)</sup>			Placebo + BSC <sup>(2)</sup>		
	Grad 1-2	Grad 3-5	Status	Grad 1-2	Grad 3-5	Status
PT: Muskelblutung	0	(1 [3])	(abgeklungen)	0	0	/
<b>SOC: Erkrankungen der Atemwege, des Brustraums und Mediastinums</b>	<b>0 (2)</b>	<b>5 (1)</b>		<b>0</b>	<b>1</b>	
PT: Dyspnoe	0 (1 [2])	2 [3]	abgeklungen	0	0	/
PT: Belastungsdyspnoe	0	1 [3]	abgeklungen	0	0	/
PT: Akutes Atemnotsyndrom	0	1 [4]	nicht abgeklungen	0	0	/
PT: Lungenfibrose	0	1 [3]	abgeklungen <sup>(1)</sup>	0	0	/
PT: Hypoxie	0	(1 [3])	(nicht abgeklungen)	0	1 [3]	abgeklungen
PT: Epistaxis	(1 [2])	0	(abgeklungen)	0	0	/
<b>SOC: Gastrointestinale Erkrankungen</b>	<b>2</b>	<b>3 (3)</b>		<b>0</b>	<b>0</b>	
PT: Ulkus duodeni	0	1 [4]	nicht abgeklungen	0	0	/
PT: Ulkus duodeni mit Blutung	0	1 [3]	abgeklungen	0	0	/
PT: Schmerzen Oberbauch	1 [2]	0	nicht abgeklungen	0	0	/
PT: akute Pankreatitis	0	1 [4]	abgeklungen	0	0	/
PT: Dysphagie	1 [2]	0	abgeklungen	0	0	/
PT: Übelkeit	0	(1 [3])	(abgeklungen)	0	0	/
PT: Erbrechen	0	(1 [3])	(abgeklungen)	0	0	/
PT: Aszites	0	(1 [3])	(abgeklungen)	0	0	/
<b>SOC: Erkrankungen des Nervensystems</b>	<b>0</b>	<b>4 (3)</b>		<b>0</b>	<b>0 (3)</b>	
PT: Synkope	0	2 [3] (2 [3])	abgeklungen	0	0	/
PT: Hirnblutung	0	1 [4] (1 [5])	abgeklungen/ (tödlich)	0	0	/
PT: Krampfanfall	0	1 [3]	abgeklungen	0	0	/

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

SOC PT	Luspatercept + BSC <sup>(2)</sup>			Placebo + BSC <sup>(2)</sup>		
	Grad 1-2	Grad 3-5	Status	Grad 1-2	Grad 3-5	Status
PT: intrakranielle Blutung	0	0	/	0	(1 [3])	(abgeklungen)
PT: Zerebrale Ischämie	0	0	/	0	(1 [4])	(Folgeschaden)
PT: Austritt von Liquor	0	0	/	0	(1 [4])	(abgeklungen)
<b>SOC: Gefäß- erkrankungen</b>	<b>1</b>	<b>3</b>		<b>0</b>	<b>0</b>	
PT: hämorrhagischer Schock	0	2 [4, 5]	nicht abgeklungen/ tödlich	0	0	/
PT: Orthostase-syndrom	0	1 [3]	Folgeschaden	0	0	/
PT: Granulomatose mit Polyangiitis	1 [2]	0	abgeklungen	0	0	/
<b>SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	<b>2 (1)</b>	<b>2 (1)</b>		<b>1 (1)</b>	<b>2</b>	
PT: Generalisiertes Ödem	0	2 [4]	nicht abgeklungen	0	0	/
PT: Brustkorbschmerz	1 [2]	0	abgeklungen	0	0	/
PT: Thoraxschmerz nicht kardialen Ursprungs	1 [2]	0	abgeklungen	0	0	/
PT: Generelle Verschlechterung des physischen Gesundheitsstatus	0	0	/	0	2 [4, 5]	nicht abgeklungen/ tödlich
PT: Fieber	0	0	/	1 [1] (1 [2])	0	abgeklungen
PT: Gangstörung	(1 [2])	0	(abgeklungen)	0	0	/
PT: Asthenie	0	(1 [3])	(nicht abgeklungen)	0	0	/
<b>SOC: Erkrankungen des Blutes und des Lymphsystems</b>	<b>0</b>	<b>3 (1)</b>		<b>0</b>	<b>0</b>	
PT: Anämie	0	2 [3] (1 [3])	abgeklungen	0	0	/

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

SOC PT	Luspatercept + BSC <sup>(2)</sup>			Placebo + BSC <sup>(2)</sup>		
	Grad 1-2	Grad 3-5	Status	Grad 1-2	Grad 3-5	Status
PT: Disseminierte intravaskuläre Gerinnung	0	1 [4]	nicht abgeklungen	0	0	/
<b>SOC: Erkrankungen der Niere und Harnwege</b>	<b>0</b>	<b>2</b>		<b>0</b>	<b>1 (1)</b>	
PT: Akute Nierenschädigung	0	2 [3]	abgeklungen/ nicht abgeklungen <sup>(1)</sup>	0	0	/
PT: Harnretention	0	0	/	0	1 [3]	Folgeschaden
PT: Nierenschädigung	0	0	/	0	(1 [3])	(abgeklungen)
<b>SOC: Leber- und Gallenerkrankungen</b>	<b>1</b>	<b>1</b>		<b>1</b>	<b>0</b>	
PT: Hyperbilirubinämie	1 [2]	0	Folgeschaden	0	0	/
PT: Cholelithiasis	0	1 [3]	nicht abgeklungen	0	0	/
PT: Cholangitis	0	0	/	1 [2]	0	/
<b>SOC: Stoffwechsel- und Ernährungsstörungen</b>	<b>0 (1)</b>	<b>1 (1)</b>		<b>0</b>	<b>0</b>	
PT: Laktatazidose	0	1 [3]	abgeklungen	0	0	/
PT: Diabetes mellitus Typ 2	0	(1 [3])	(nicht abgeklungen)	0	0	/
PT: Hypoglykämie	(1 [2])	0	(nicht abgeklungen)	0	0	/
<b>SOC: Untersuchungen</b>	<b>0</b>	<b>1</b>		<b>0</b>	<b>0</b>	
PT: Allgemeinzustand anormal	0	1 [3]	abgeklungen	0	0	/
<b>SOC: Probleme mit Medizinprodukten</b>	<b>1</b>	<b>0</b>		<b>0</b>	<b>0</b>	
PT: Thrombose in einem Medizinprodukt	1 [2]	0	Folgeschaden	0	0	/

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

SOC PT	Luspatercept + BSC <sup>(2)</sup>			Placebo + BSC <sup>(2)</sup>		
	Grad 1-2	Grad 3-5	Status	Grad 1-2	Grad 3-5	Status
<b>SOC: Psychiatrische Erkrankungen</b>	<b>0</b>	<b>0 (1)</b>		<b>0</b>	<b>1 (1)</b>	
PT: Suizidversuch	0	0	/	0	1 [4]	abgeklungen
PT: Verwirr- heitszustand	0	(1 [3])	(abgeklungen)	0	0	/
PT: Delirium	0	0	/	0	(1 [3])	(abgeklungen)

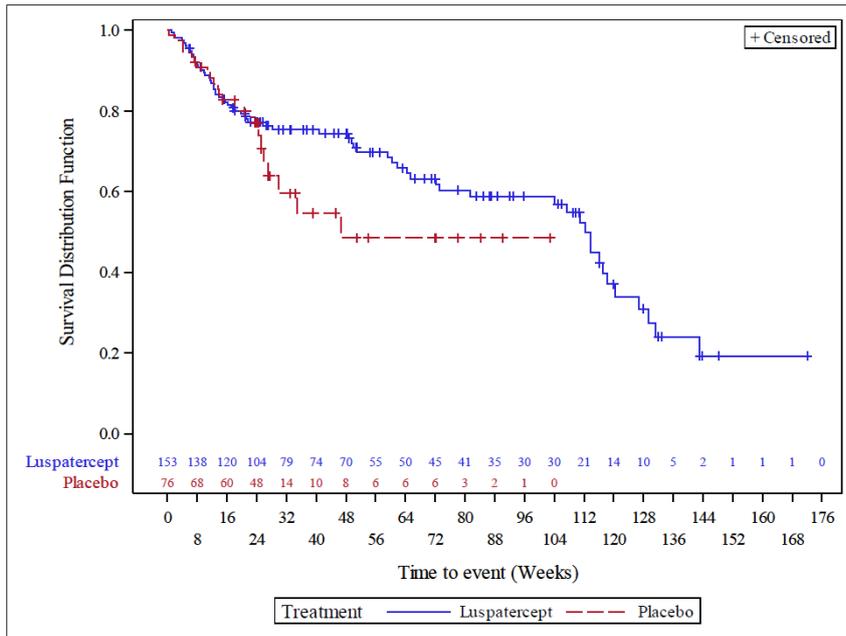
Datenschnitt: 07.01.2019; Population: ITT-Population  
<sup>(1)</sup> Verdacht auf möglichen Zusammenhang zur Studienmedikation nach Ermessen des Prüfarztes  
<sup>(2)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).  
 ( ) = aufgetretene Hospitalisierung aufgrund von UE nach Woche 24; [ ] = Schweregrad nach CTCAE; CTCAE = *Common Terminology Criteria for Adverse Events*; ITT = Intention-to-Treat; PT = bevorzugter Begriff (*preferred term*); SOC = Systemorganklasse (*system organ class*)  
 Quelle: [94]

### Zusatzanalyse 1: Zeit bis zur ersten Hospitalisierung

Tabelle 4-52: Zusatzanalyse 1: Ergebnisse zum Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Zeit bis zur ersten Hospitalisierung; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Hospitalisierung jeglicher Ursache	63/153 (41,2)	112,3 [81,4; 118,1]	24/76 (31,6)	46,7 [27,0; N.E.]	0,77 [0,46; 1,27]	0,305
Hospitalisierung aufgrund von UE und/oder Transfu- sionen	51/153 (33,3)	118,1 [113,7; 143,0]	19/76 (25,0)	N.E. [45,3; N.E.]	0,88 [0,50; 1,53]	0,643

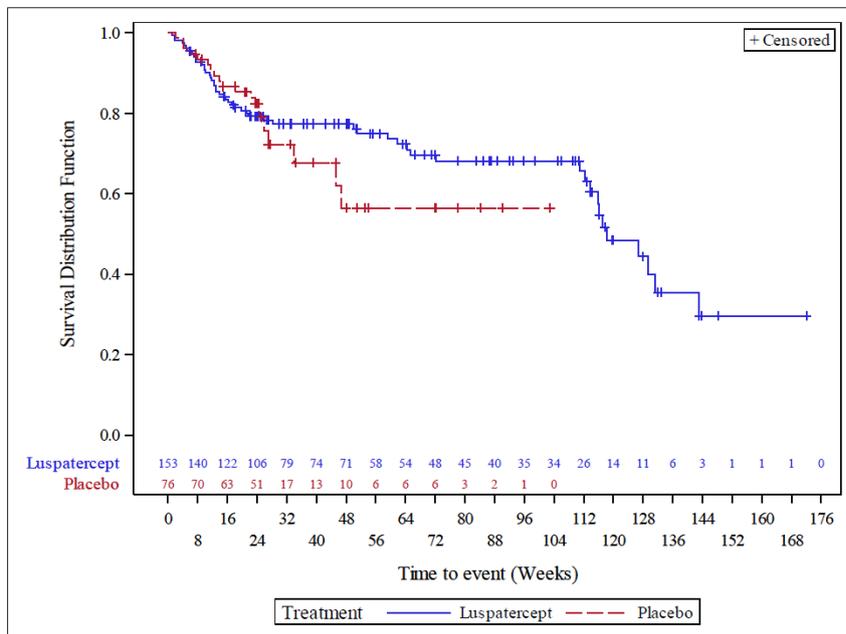
Datenschnitt: 01.07.2019; Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.  
<sup>(3)</sup> HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.  
<sup>(4)</sup> Der p-Wert des HR basiert auf einem stratifizierten Log-Rank-Test.  
 BSC = Best-Supportive-Care; HR = Hazard-Ratio; ITT = Intention-to-Treat; IPSS-R = International Prognostic Scoring System-revised; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar  
 Quelle: [20]



Datenschnitt: 01.07.2019

Quelle: [20]

Abbildung 4-33: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur ersten Hospitalisierung in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019

Quelle: [20]

Abbildung 4-34: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zur ersten Hospitalisierung aufgrund von UE und/oder Transfusionen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

Auswertungen über den Studienzeitraum bis zum Datenschnitt vom 01.07.2019 werden aufgrund der Unsicherheit hinsichtlich der Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalysen dargestellt.

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit mindestens einer Hospitalisierung in den ersten 24 Wochen der Behandlungsphase, zeigte sich auch in der Analyse der Zeit bis zur ersten Hospitalisierung jeglicher Ursache kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (HR = 0,77; 95 %-KI: [0,46; 1,27]; p = 0,305). Dies trifft auch für die Ergebnisse der Zeit bis zur ersten Hospitalisierung aufgrund von UE und/oder Transfusionen zu (HR = 0,88; 95 %-KI: [0,50; 1,53]; p = 0,643; siehe Tabelle 4-52 & Abbildung 4-33 f.).

Analog zur Hauptanalyse war auch bei Betrachtung der Ereignisse über Woche 25 hinaus der Hauptgrund für eine stationäre Aufnahme das Auftreten eines UE. Andere Ursachen, darunter auch Transfusionen, traten lediglich vereinzelt auf (siehe Tabelle 4-53).

Tabelle 4-53: Zusatzanalyse 1: Ergebnisse zum Endpunkt Morbidität – Hospitalisierungen aus MEDALIST – Gründe für Hospitalisierungen; Woche eins bis zum Datenschnitt vom 01.07.2019

<b>Studie MEDALIST</b> <b>Gründe für Hospitalisierungen n (%)</b>	<b>Luspatercept + BSC<sup>(1)</sup></b> <b>N = 153</b>	<b>Placebo + BSC<sup>(1)</sup></b> <b>N = 76</b>
UE	51 (33,3)	19 (25,0)
elektive Behandlung einer vorbestehenden Erkrankung	6 (3,9)	5 (6,6)
Nicht Protokoll-bedingte Untersuchung	7 (4,6)	3 (3,9)
Behandlung, die vor Unterzeichnung der Einverständniserklärung geplant war	2 (1,3)	0
Protokoll-bedingte Behandlung	2 (1,3)	2 (2,6)
Soziale, technische oder praktische Gründe in Abwesenheit eines UE	1 (0,7)	0
Transfusion	2 (1,3)	1 (1,3)
Datenschnitt: 01.07.2019; Population: ITT-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).  BSC = Best-Supportive-Care; ITT = Intention-to-Treat; UE = unerwünschtes Ereignis Quelle: [20]		

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

### **Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

### **Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Morbidität – Hospitalisierungen liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.1.7 Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-54: Operationalisierung des Endpunkts Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Die gesundheitsbezogene Lebensqualität wurde mittels der Funktionsskalen des generischen Fragebogens EORTC QLQ-C30 (Version 3.0) erhoben:</p> <ul style="list-style-type: none"> <li>• Körperliche Funktion (Item 1 bis 5)</li> <li>• Rollenfunktion (Item 6, 7)</li> <li>• Emotionale Funktion (Item 21 bis 24)</li> <li>• Kognitive Funktion (Item 20, 25)</li> <li>• Soziale Funktion (Item 26, 27)</li> <li>• Gesundheitlicher Status insgesamt (Item 29, 30)</li> </ul> <p>Die gesundheitsbezogene Lebensqualität wurde zur Baseline (Dosis 1 Tag 1), zu jedem zweiten Zyklus (Zyklus 1, 3, 5, 7) und der Woche-25-Visite, jeweils vor Verabreichung der Studienmedikation erhoben. Für Patienten, die nach der Woche 25-Visite weiterhin doppelblind behandelt wurden, wurde die gesundheitsbezogene Lebensqualität weiterhin zu jedem zweiten Zyklus vor Verabreichung der Studienmedikation erhoben.</p> <p>In die Analysen zur Verbesserung der gesundheitsbezogenen Lebensqualität gehen nur solche Patienten ein, die zur Baseline und zu mindestens einem Auswertungszeitpunkt im Studienverlauf einen ausgefüllten EORTC QLQ-C30-Fragebogen aufweisen (HRQoL-Population). Ein Fragebogen galt dann als ausgefüllt, wenn mindestens die Hälfte aller Fragen beantwortet wurden, d. h. mindestens 15 Fragen des EORTC QLQ-C30-Fragebogens ausgefüllt waren.</p> <p>Auswertungen zu den Rücklaufquoten des Fragebogens beziehen sich auf die ITT-Population. Höhere Antwortwerte der Funktionsskalen des EORTC QLQ-C30 entsprechen einer höheren Lebensqualität.</p> <p>Skalen, für die weniger als die Hälfte der Items beantwortet waren, wurden entsprechend der Vorgaben im Handbuch des EORTC QLQ-C30 als fehlende Werte bewertet. Eine Imputation fehlender Werte wurde nicht durchgeführt.</p> <p>Als klinisch relevante Änderung wurde die vom G-BA anerkannte und im Studienprotokoll präspezifizierte MCID von zehn Punkten herangezogen.</p> <p>In den Responder-Analysen (Hauptanalyse) wurden Unterschiede zwischen den Behandlungsarmen mittels CMH-Methode, stratifiziert nach Transfusionslast zur Baseline und IPSS-R-Score zur Baseline, berechnet und als RR, OR und ARR, jeweils inkl. 95 %-KI, dargestellt. Das OR wurde mittels logistischem Regressionsmodell mit Behandlungsarm als Haupteffekt berechnet. Für die Berechnung des RR und der 95 %-KI wurde eine Normalverteilung angenommen. Patienten mit fehlenden Werten zu Woche 48 wurden als Patienten ohne klinisch relevante Verbesserung der Lebensqualität gewertet (Non-Responder).</p> <p>Ereigniszeitanalysen, die Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in Woche 1 – 48, wird ergänzend dargestellt. Patienten ohne klinisch relevante Verbesserung wurden zum Zeitpunkt der Analyse zensiert. Die mediane Zeit bis zum Erreichen einer klinisch relevanten Verbesserung wurde mittels Kaplan-Meier-Methode geschätzt. Kaplan-Meier-Schätzer für die mediane Zeit bis zur Verbesserung der gesundheitsbezogenen Lebensqualität sowie das HR, berechnet mittels Cox-Modell mit Behandlungsarm und Baseline-Werten als Kovariate, jeweils inkl. 95 %-KI werden dargestellt.</p>

Studie	Operationalisierung
	<p>Änderungen der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten im Vergleich zur Baseline werden grafisch dargestellt, inkl. tabellarischer, deskriptiver Listung der Werte im Studienverlauf. Unterschiede zwischen den Behandlungsarmen wurden mittels ANOVA berechnet. Statistisch signifikante Unterschiede (<math>p &lt; 0,05</math>) wurden in den Grafiken markiert.</p> <p>Änderungen der gesundheitsbezogenen Lebensqualität im Studienverlauf wurden mittels MMRM berechnet. In das Modell flossen Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und ein Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte ein. SMD inkl. 95 %-KI in Form von Hedges' g werden ergänzend dargestellt, um klinisch relevante Effekte zu bewerten.</p> <p>Auswertungen zu den Datenschnitten aus dem Jahr 2019 werden nicht dargestellt, da die Rücklaufquoten sehr gering waren (insbesondere im Placebo-Arm hatten <math>&lt; 10</math> Patienten einen ausgefüllten Fragebogen) und somit weitere Aussagen zum Zusatznutzen von Luspatercept nicht möglich sind.</p> <p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Hauptanalyse:</b> Anteil der Patienten mit einer Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte; Woche 1 – 24 (siehe Tabelle 4-56)</li> <li>• <b>Zusatzanalyse 1:</b> Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte; Woche 1 – 48 (siehe Tabelle 4-57)</li> <li>• <b>Zusatzanalyse 2:</b> Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten; Woche 1 – 48 (siehe Abbildung 4-41 ff. &amp; Tabelle 4-58)</li> <li>• <b>Zusatzanalyse 3:</b> Mittlere Änderung der gesundheitsbezogenen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 24 und Woche 1 – 48 (siehe Tabelle 4-59 f.)</li> </ul>
	<p>ANOVA = <i>analysis of variance</i>; ARR = absolute Risikodifferenz; CMH = Cochran-Mantel-Haenszel; EORTC QLQ-C30 = <i>European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module</i>; HRQoL-Population = Population mit auswertbarer gesundheitsbezogener Lebensqualität; IPSS-R = <i>International Prognostic Scoring System-revised</i>; KI = Konfidenzintervall; MCID = minimale klinische Relevanzschwelle; MMRM = gemischtes Modell für wiederholte Messungen (<i>mixed-model repeated measure</i>); OR = Odds Ratio; RR = Relatives Risiko; SMD = standardisierte Mittelwertdifferenz</p>

*Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.*

*Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.*

Tabelle 4-55: Bewertung des Verzerrungspotenzials für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens in MEDALIST

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
MEDALIST	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens wurde verblindet erhoben. In die Analyse gingen nur solche Patienten ein, die zu Studienbeginn und mindestens einem Zeitpunkt im Studienverlauf einen EORTC QLQ-C30-Fragebogen ausgefüllt hatten (HRQoL-Population). Zwischen der HRQoL-Population und der ITT-Population liegen keine wesentlichen Unterschiede hinsichtlich der Baseline-Charakteristika vor [1, 95]. Darüber hinaus umfasst die HRQoL-Population > 95 % der Patienten der ITT-Population (Luspatercept: 149/153 (97,4 %); Placebo: 76/76 (100,0 %)), sodass von einer guten Vergleichbarkeit der HRQoL- und ITT-Population auszugehen ist (siehe Tabelle 4-6). Die Rücklaufquoten bezogen auf die ITT-Population liegen für den EORTC QLQ-C30 bis Woche 25 über den vom G-BA geforderten 70 % und sind damit ausreichend hoch, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen (siehe Tabelle 4-40). Ergebnisse nach Woche 25 wurden aufgrund geringerer Rücklaufquoten (< 70 %) nur bis Woche 48 ergänzend berichtet. Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens und eine hohe Aussagesicherheit der Ergebnisse.

*Stellen Sie die Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

### Hauptanalyse: Anteil der Patienten mit einer Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte

Tabelle 4-56: Ergebnisse zum Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Anteil der Patienten mit einer Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI]	RR [95 %-KI] <sup>(3)</sup>	ARR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Funktionsskala des EORTC QLQ-C30- Fragebogens	Patienten mit Ereignis <sup>(2)</sup> n/N (%)	Patienten mit Ereignis <sup>(2)</sup> n/N (%)				
Körperliche Funktion	25/149 (16,8)	18/76 (23,7)	0,65 (0,33; 1,28)	0,71 (0,41; 1,21)	-6,9 (-18,2; 4,4)	0,215
Rollenfunktion	30/149 (20,1)	18/76 (23,7)	0,81 (0,42; 1,58)	0,85 (0,51; 1,42)	-3,5 (-15,1; 8,0)	0,539
Emotionale Funktion	18/149 (12,1)	11/76 (14,5)	0,81 (0,36; 1,82)	0,83 (0,42; 1,68)	-2,4 (-11,9; 7,1)	0,613
Kognitive Funktion	29/149 (19,5)	14/76 (18,4)	1,07 (0,53; 2,17)	1,06 (0,59; 1,88)	1,0 (-9,7; 11,8)	0,851
Soziale Funktion	26/149 (17,4)	16/76 (21,1)	0,79 (0,40; 1,59)	0,83 (0,47; 1,45)	-3,6 (-14,6; 7,4)	0,512
Gesundheitlicher Status insgesamt	31/149 (20,8)	12/76 (15,8)	1,40 (0,67; 2,91)	1,32 (0,72; 2,42)	5,0 (-5,5; 15,5)	0,367
Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> Patienten mit Ereignis waren definiert als Patienten mit einer Verbesserung der Symptomatik um mindestens zehn Punkte; Patienten mit fehlenden Werten zu Woche 24 wurden als Patienten ohne Ereignis gezählt. <sup>(3)</sup> RR und ARR wurden mittels Normal-Approximation berechnet. <sup>(4)</sup> OR und p-Wert berechnet mittels logistischem Regressionsmodell mit Behandlungsarm als Haupteffekt.  ARR = Absolute Risikodifferenz; HRQoL = <i>health-related quality of life</i> ; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko Quelle: [19]						

In den ersten 24 Wochen der Behandlungsphase zeigte sich für den Anteil der Patienten mit einer klinisch relevanten Verbesserung der gesundheitsbezogenen Lebensqualität, gemessen über die Verbesserung der Funktionsskalen des EORTC QLQ-C30 um mindestens zehn Punkte, kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen. Über alle Symptomskalen hinweg traten in beiden Behandlungsarmen zu etwa gleichen Teilen klinisch relevante Verbesserungen auf (siehe Tabelle 4-56).

### Zusatzanalyse 1: Zeit bis zur Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte

Tabelle 4-57: Zusatzanalyse 1: Ergebnisse zum Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Zeit bis zur Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte; Woche 1 – 48

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Körperliche Funktion	59/149 (39,6)	60,1 (36,1; N.E.)	38/76 (50,0)	27,1 (19,9; 54,1)	0,660 (0,434; 1,003)	0,052
Rollenfunktion	68/149 (45,6)	42,4 (24,1; N.E.)	41/76 (53,9)	24,1 (18,0; 34,0)	0,674 (0,452; 1,007)	0,054
Emotionale Funktion	56/149 (37,6)	73,3 (44,1; N.E.)	32/76 (42,1)	N.E. (21,7; N.E.)	0,693 (0,442; 1,089)	0,112
Kognitive Funktion	59/149 (39,6)	73,3 (32,3; N.E.)	28/76 (36,8)	N.E. (36,3; N.E.)	0,969 (0,612; 1,534)	0,893
Soziale Funktion	61/149 (40,9)	N.E. (32,4; N.E.)	42/76 (55,3)	23,1 (12,1; 34,0)	0,588 (0,390; 0,888)	0,012
Gesundheitlicher Status insgesamt	64/149 (43,0)	50,1 (30,0; N.E.)	33/76 (43,4)	29,6 (22,1; N.E.)	0,643 (0,411; 1,005)	0,053

Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline

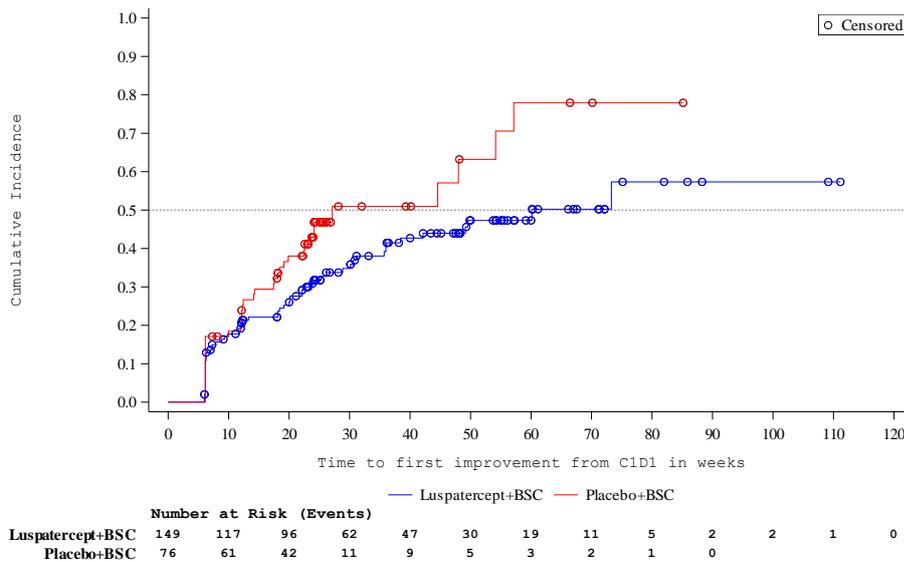
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

<sup>(2)</sup> Schätzung nach Kaplan-Meier-Methode.

<sup>(3)</sup> HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell mit Behandlung und Baseline-Wert als Kovariable.

<sup>(4)</sup> Der p-Wert des HR basiert auf einem stratifizierten Log-Rank-Test.

BSC = Best-Supportive-Care; HR = Hazard-Ratio; HRQoL = *health-related quality of life*; IPSS-R = *International Prognostic Scoring System-revised*; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar  
Quelle: [19]

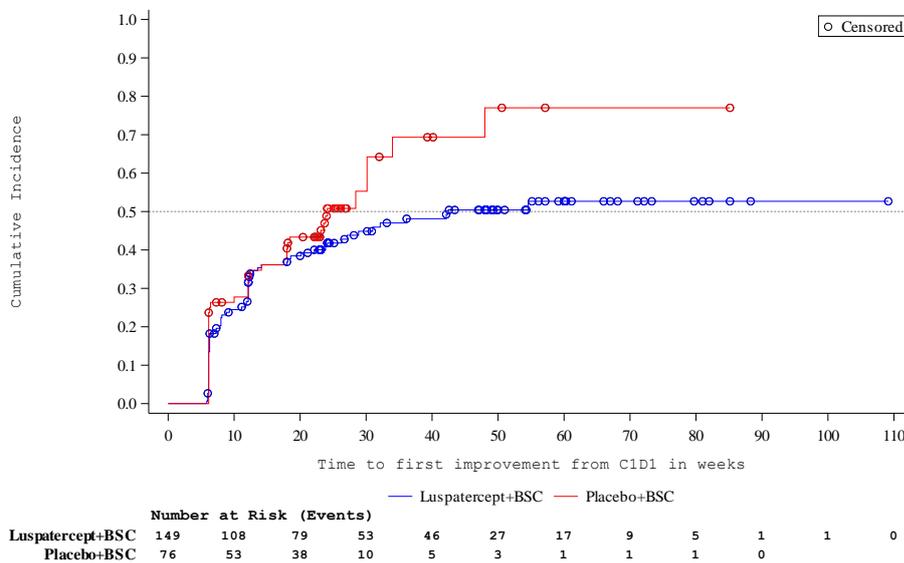


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

Quelle: [19]

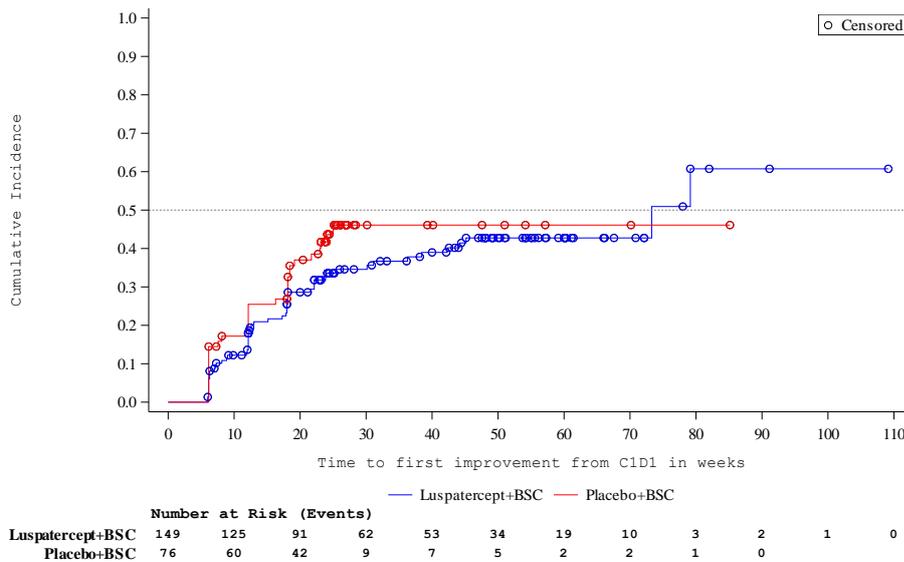
Abbildung 4-35: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Körperliche Funktion; Woche 1 – 48



Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). Quelle: [19]

Abbildung 4-36: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Rollenfunktion; Woche 1 – 48

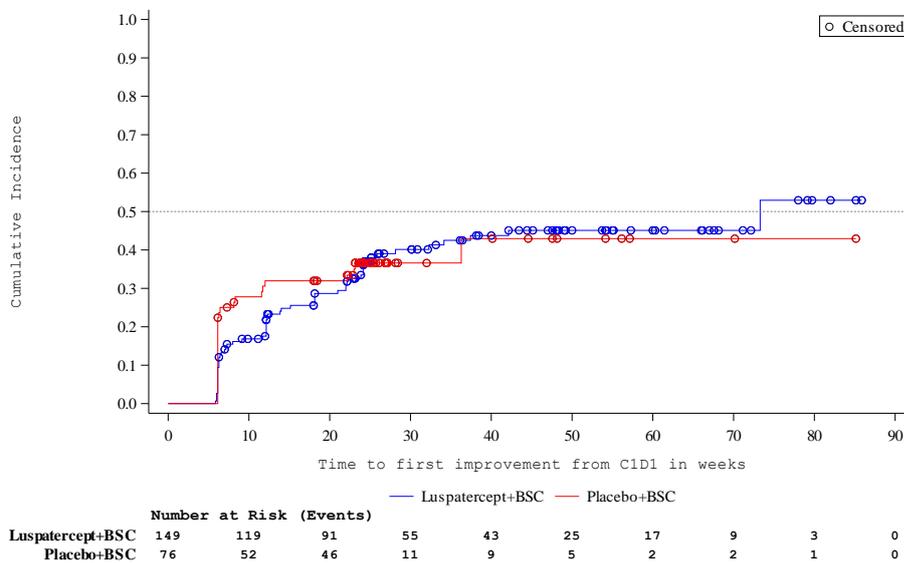


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-37: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Emotionale Funktion; Woche 1 – 48

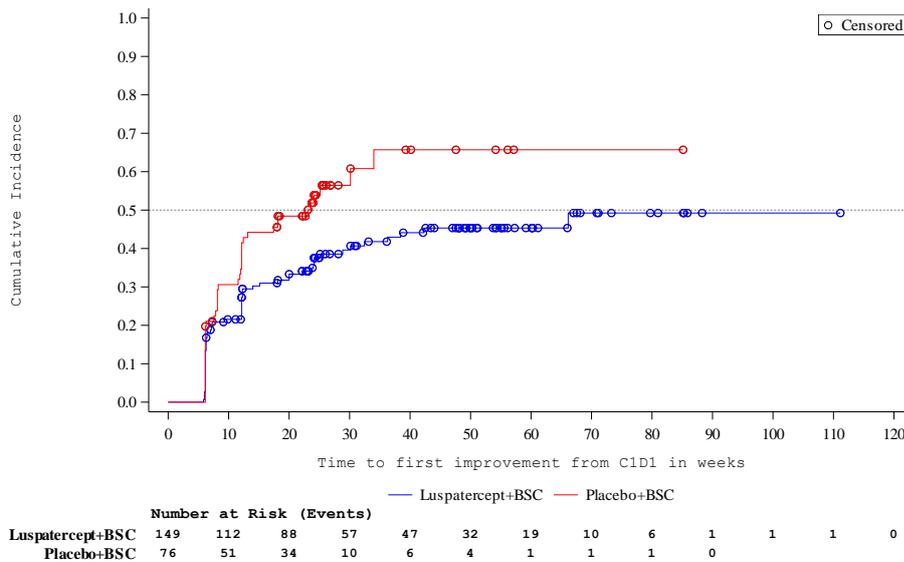


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-38: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Kognitive Funktion; Woche 1 – 48

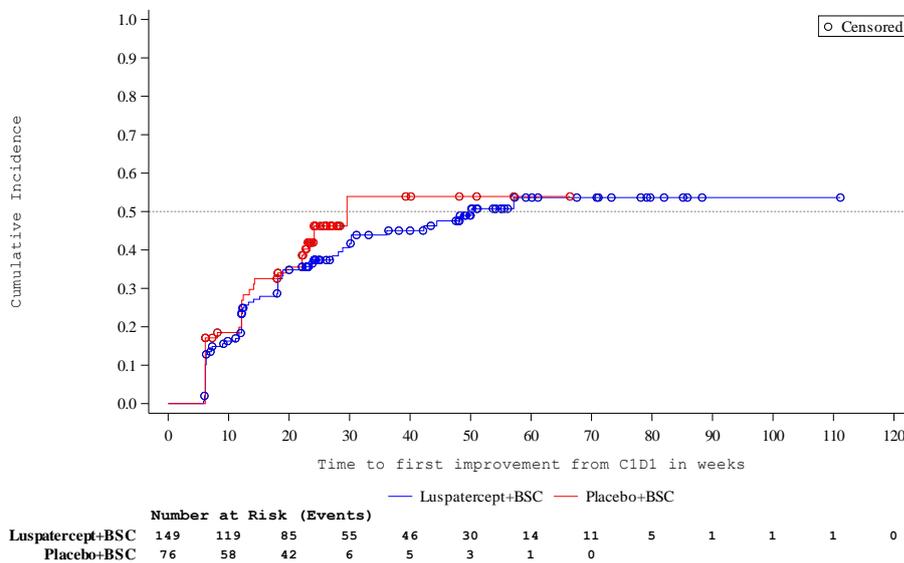


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-39: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Soziale Funktion; Woche 1 – 48



Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

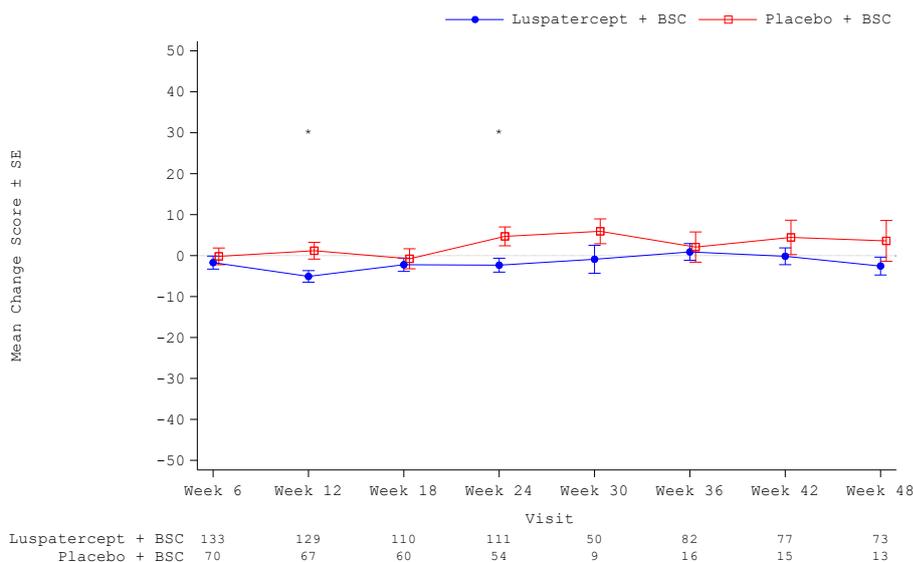
Quelle: [19]

Abbildung 4-40: Zusatzanalyse 1: Kumulative Inzidenz der Zeit bis zur ersten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in MEDALIST – Gesundheitlicher Status insgesamt; Woche 1 – 48

Auswertungen über den Studienzeitraum bis Woche 48 werden aufgrund der Unsicherheit hinsichtlich der Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten sowie Rücklaufquoten < 70 % nach Woche 24 ergänzend als Ereigniszeitanalysen dargestellt.

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit einer Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte in den ersten 24 Wochen der Behandlungsphase, zeigten sich auch in der Analyse der Zeit bis zum Erreichen einer klinisch relevanten Verbesserung der gesundheitsbezogenen Lebensqualität um mindestens zehn Punkte – mit Ausnahme der Funktionsskala soziale Funktion – keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen. Der Anteil der Patienten mit einer klinisch relevanten Verbesserung der gesundheitsbezogenen Lebensqualität war in beiden Behandlungsarmen für alle Funktionsskalen etwa gleich groß (siehe Tabelle 4-57 & Abbildung 4-35 ff.).

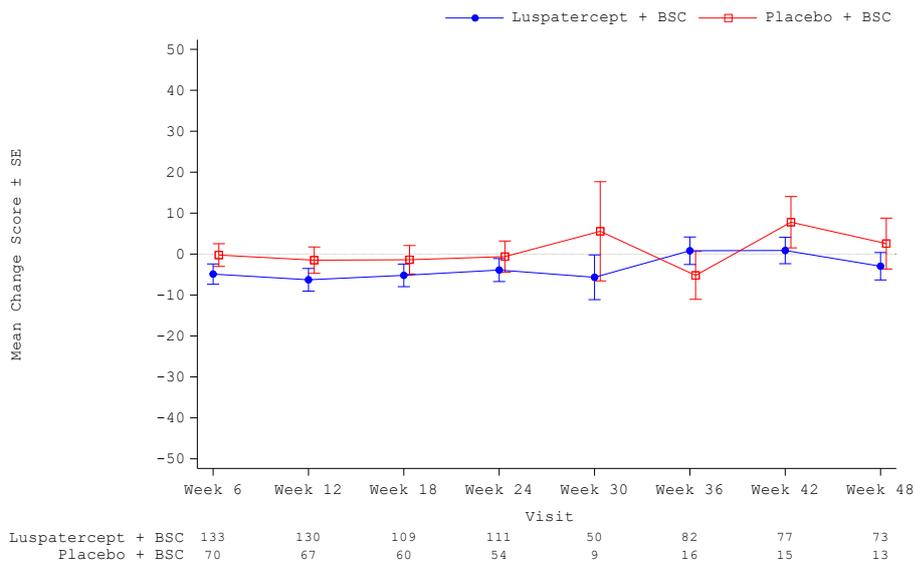
### Zusatzanalyse 2: Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten



Datenschnitt: 08.05.2018; \* = statistisch signifikanter Unterschied, berechnet mittels ANOVA (*analysis of variance*); BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-41: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – körperliche Funktion; Woche 1 – 48

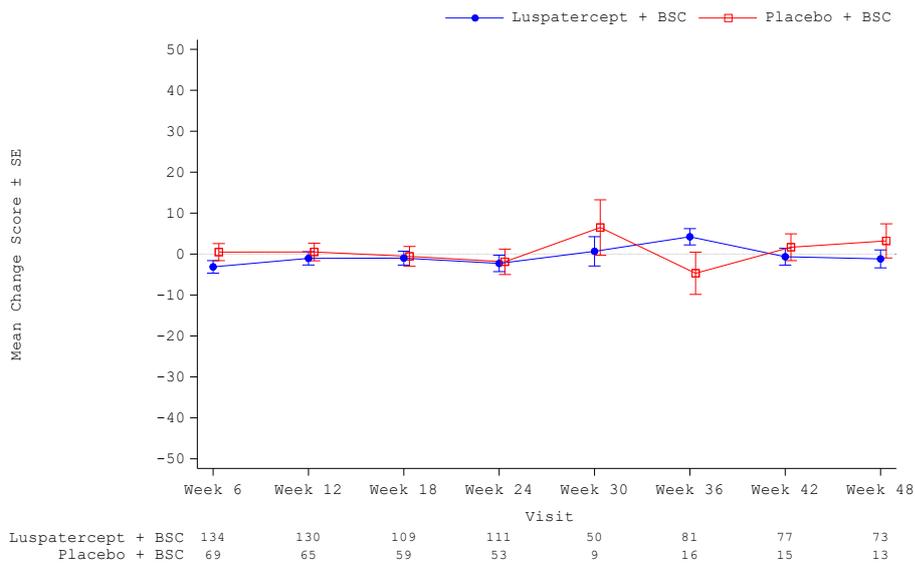


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-42: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – Rollenfunktion; Woche 1 – 48

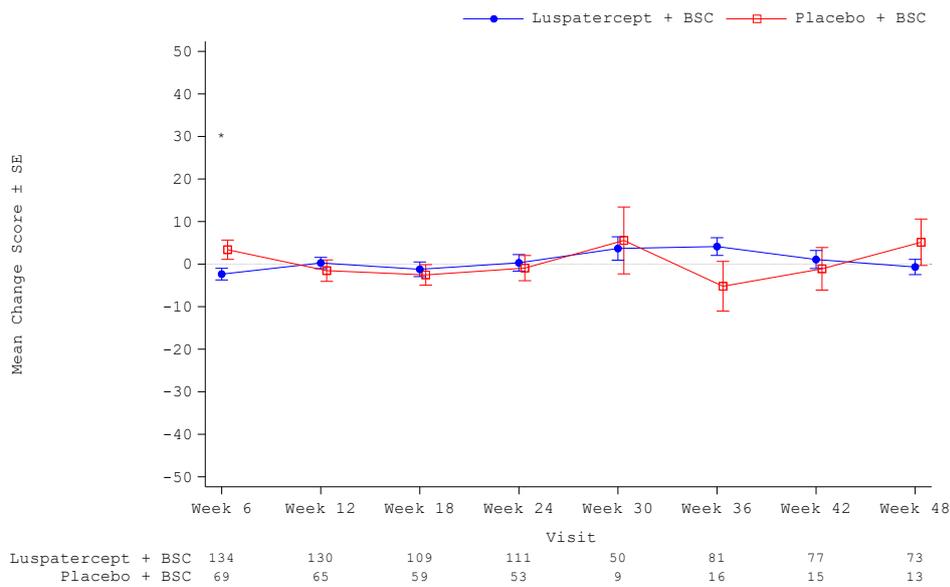


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-43: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – emotionale Funktion; Woche 1 – 48

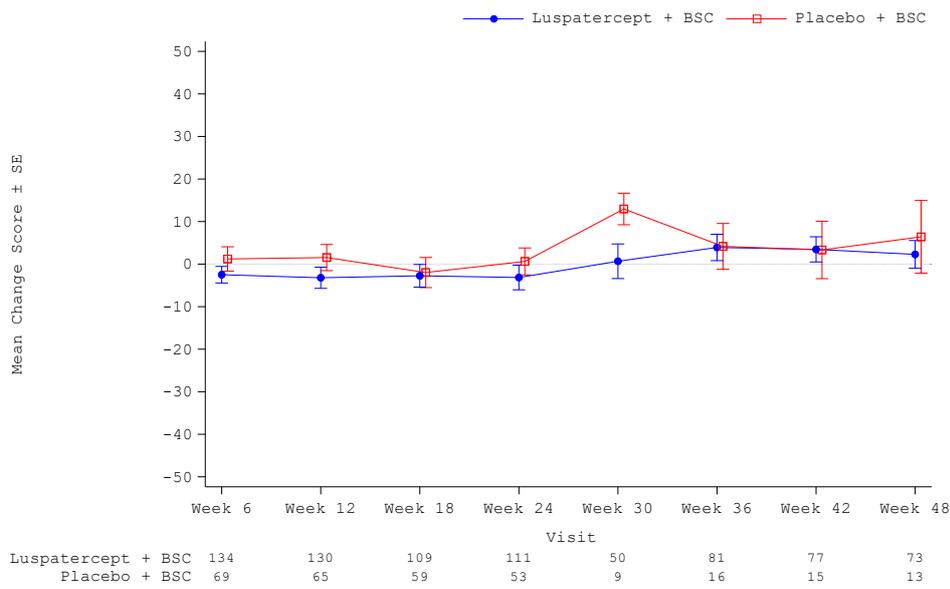


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-44: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – kognitive Funktion; Woche 1 – 48

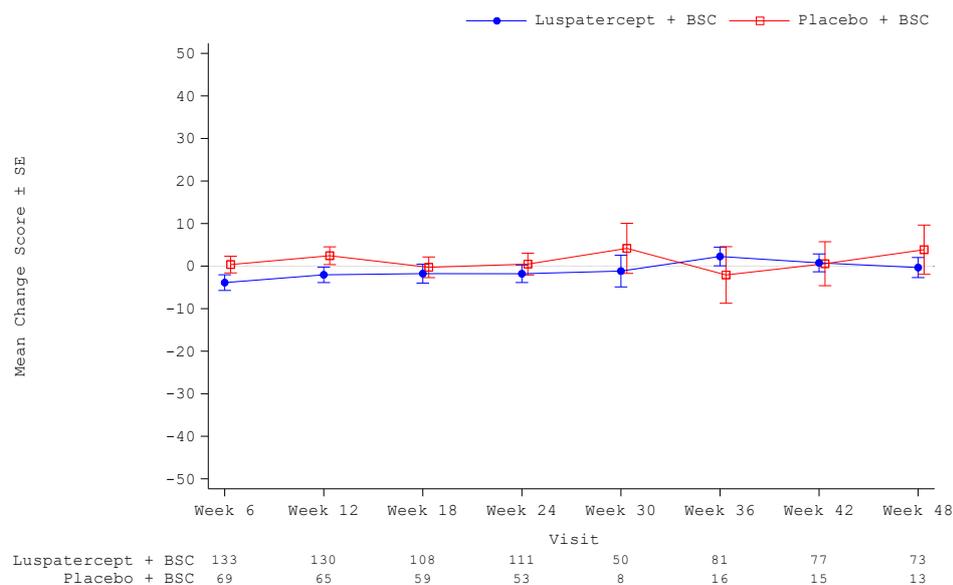


Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-45: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – soziale Funktion; Woche 1 – 48



Datenschnitt: 08.05.2018;

BSC = Best-Supportive-Care; BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

Quelle: [19]

Abbildung 4-46: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten – gesundheitlicher Status insgesamt; Woche 1 – 48

Tabelle 4-58: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Werte der gesundheitsbezogenen Lebensqualität zu den Erhebungszeitpunkten; Woche 1 – 48

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
<b>Körperliche Funktion</b>				
Baseline	149 (100)	66,85 (21,82)	76 (100)	65,26 (19,99)
Woche 6	133 (89,3)	65,21 (21,11)	70 (92,1)	65,71 (19,27)
Woche 12	129 (86,6)	62,21 (22,68)	67 (88,2)	66,87 (20,10)
Woche 18	110 (73,8)	66,68 (22,05)	60 (78,9)	66,67 (20,39)
Woche 24	111 (74,5)	66,91 (21,44)	54 (71,1)	70,99 (18,28)
Woche 30	50 (33,6)	65,23 (23,04)	9 (11,8)	71,11 (14,91)
Woche 36	82 (55,0)	68,62 (22,41)	16 (21,1)	69,17 (24,81)
Woche 42	77 (51,7)	68,83 (20,73)	15 (19,7)	72,00 (22,28)
Woche 48	73 (49,0)	67,58 (21,00)	13 (17,1)	64,10 (26,46)

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
<b><i>Rollenfunktion</i></b>				
Baseline	149 (100)	65,66 (30,46)	76 (100)	64,25 (27,72)
Woche 6	133 (89,3)	62,03 (30,11)	70 (92,1)	64,05 (29,30)
Woche 12	130 (87,2)	60,77 (31,67)	67 (88,2)	63,43 (30,46)
Woche 18	109 (73,2)	63,30 (31,08)	60 (78,9)	63,89 (33,22)
Woche 24	111 (74,5)	65,47 (30,29)	54 (71,1)	65,43 (29,65)
Woche 30	50 (33,6)	60,33 (31,21)	9 (11,8)	74,07 (26,50)
Woche 36	82 (55,0)	66,46 (31,26)	16 (21,1)	59,38 (27,20)
Woche 42	77 (51,7)	69,70 (30,07)	15 (19,7)	76,67 (26,58)
Woche 48	73 (49,0)	66,21 (27,21)	13 (17,1)	64,10 (31,80)
<b><i>Emotionale Funktion</i></b>				
Baseline	149 (100)	76,81 (20,17)	75 (98,7)	76,78 (19,15)
Woche 6	134 (89,9)	73,51 (23,16)	70 (92,1)	76,19 (20,72)
Woche 12	130 (87,2)	75,64 (22,03)	66 (86,8)	75,88 (18,73)
Woche 18	109 (73,2)	76,40 (22,64)	60 (78,9)	76,81 (19,83)
Woche 24	111 (74,5)	75,23 (22,03)	54 (71,1)	73,30 (22,45)
Woche 30	50 (33,6)	75,83 (22,48)	9 (11,8)	79,63 (12,58)
Woche 36	81 (54,4)	80,14 (19,70)	16 (21,1)	78,13 (19,92)
Woche 42	77 (51,7)	76,08 (21,94)	15 (19,7)	85,56 (9,16)
Woche 48	73 (49,0)	77,47 (20,93)	13 (17,1)	85,26 (11,86)
<b><i>Kognitive Funktion</i></b>				
Baseline	149 (100)	82,89 (19,66)	75 (98,7)	80,22 (21,52)
Woche 6	134 (89,9)	81,34 (22,11)	70 (92,1)	82,62 (19,54)
Woche 12	130 (87,2)	83,46 (20,07)	66 (86,8)	79,04 (23,81)
Woche 18	109 (73,2)	83,03 (22,56)	60 (78,9)	80,83 (23,33)
Woche 24	111 (74,5)	83,78 (19,00)	54 (71,1)	79,94 (18,14)
Woche 30	50 (33,6)	84,67 (14,22)	9 (11,8)	88,89 (18,63)
Woche 36	81 (54,4)	85,60 (18,59)	16 (21,1)	75,00 (24,34)
Woche 42	77 (51,7)	83,33 (17,52)	15 (19,7)	83,33 (18,90)
Woche 48	73 (49,0)	82,88 (17,34)	13 (17,1)	84,62 (15,90)
<b><i>Soziale Funktion</i></b>				
Baseline	149 (100)	75,95 (27,01)	75 (98,7)	71,56 (28,44)
Woche 6	134 (89,9)	75,25 (27,02)	70 (92,1)	72,14 (28,48)
Woche 12	130 (87,2)	72,69 (27,45)	66 (86,8)	75,25 (22,69)

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
Woche 18	109 (73,2)	74,77 (27,65)	60 (78,9)	73,61 (28,51)
Woche 24	111 (74,5)	74,62 (27,42)	54 (71,1)	71,30 (30,96)
Woche 30	50 (33,6)	75,67 (25,69)	9 (11,8)	77,78 (27,64)
Woche 36	81 (54,4)	76,34 (24,13)	16 (21,1)	75,00 (27,22)
Woche 42	77 (51,7)	77,92 (26,55)	15 (19,7)	80,00 (19,11)
Woche 48	73 (49,0)	77,17 (26,13)	13 (17,1)	75,64 (24,17)
<b>Gesundheitlicher Status insgesamt</b>				
Baseline	149 (100)	57,72 (21,00)	75 (98,7)	58,67 (19,49)
Woche 6	133 (89,3)	55,64 (20,91)	70 (92,1)	58,21 (17,40)
Woche 12	130 (87,2)	56,03 (21,27)	66 (86,8)	60,86 (17,72)
Woche 18	108 (72,5)	58,10 (20,77)	60 (78,9)	59,03 (21,16)
Woche 24	111 (74,5)	59,08 (20,55)	54 (71,1)	60,34 (18,94)
Woche 30	50 (33,6)	56,00 (22,27)	8 (10,5)	60,42 (17,68)
Woche 36	81 (54,4)	60,39 (18,71)	16 (21,1)	55,21 (19,69)
Woche 42	77 (51,7)	60,39 (19,55)	15 (19,7)	61,11 (15,00)
Woche 48	73 (49,0)	59,82 (20,19)	13 (17,1)	59,62 (16,96)
Datenschnitt: 08.05.2018; Population: HRQoL-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i> ; MW = Mittelwert; SD = Standardabweichung Quelle: [19]				

Die mittlere Änderung der gesundheitsbezogenen Lebensqualität zu den einzelnen Erhebungszeitpunkten ist für die einzelnen Funktionsskalen des EORTC QLQ-C30 über den gesamten Studienverlauf bis Woche 48 grafisch dargestellt. Durch die Selektion der Patienten mit einem Therapieerfolg zu Woche 25, ist die Interpretierbarkeit der Ergebnisse ab Woche 25 mit Unsicherheit behaftet. Darüber hinaus liegen die Rücklaufquoten nach Woche 25 < 70 %, sodass die Aussagekraft der Ergebnisse limitiert ist.

Insgesamt bestätigen die grafischen Darstellungen der mittleren Änderung der gesundheitsbezogenen Lebensqualität die Ergebnisse der Hauptanalyse. Für keine der Funktionsskalen zeigte sich ein über die Behandlungsphase bis Woche 24 konstant statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Abbildung 4-41 – Abbildung 4-46 und Tabelle 4-58). Die Änderungen der gesundheitsbezogenen Lebensqualität sind innerhalb der beiden Behandlungsarme im Vergleich zur Baseline vorwiegend als geringfügige Effekte einzustufen (mittlere Änderung < 10 Punkte [96]).

### Zusatzanalyse 3: Mittlere Änderung der gesundheitsbezogenen Lebensqualität im Studienverlauf (MMRM-Analyse)

Tabelle 4-59: Zusatzanalyse 3: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 24

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studienbeginn <sup>(3)</sup>	MW (SD) zu Woche 24	N <sup>(2)</sup>	MW (SD) zu Studienbeginn <sup>(3)</sup>	MW (SD) zu Woche 24	LS-MWD (95 %-KI)	p-Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>Körperliche Funktion</b>								
111	66,85 (21,82)	66,91 (21,44)	54	65,26 (19,99)	70,99 (18,28)	-6,92 (-11,80; -2,04)	0,006	-0,32 (-0,60; -0,05)
<b>Rollenfunktion</b>								
111	65,66 (30,46)	65,47 (30,29)	54	64,25 (27,72)	65,43 (29,65)	-5,26 (-13,39; 2,86)	0,203	-0,18 (-0,45; 0,10)
<b>Emotionale Funktion</b>								
111	76,81 (20,17)	75,23 (22,03)	54	76,78 (19,15)	73,30 (22,45)	-1,06 (-6,70; 4,58)	0,710	-0,05 (-0,33; 0,22)
<b>Kognitive Funktion</b>								
111	82,89 (19,66)	83,78 (19,00)	54	80,22 (21,52)	79,94 (18,14)	1,31 (-4,12; 6,73)	0,635	0,06 (-0,21; 0,34)
<b>Soziale Funktion</b>								
111	75,95 (27,01)	74,642 (27,42)	54	71,56 (28,44)	71,30 (30,96)	-3,43 (-11,01; 4,15)	0,373	-0,12 (-0,40; 0,15)
<b>Gesundheitlicher Status insgesamt</b>								
111	57,72 (21,00)	59,08 (20,55)	54	58,67 (19,49)	60,34 (18,94)	-3,82 (-9,44; 1,79)	0,181	-0,19 (-0,46; 0,09)

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline.								
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).								
<sup>(2)</sup> Anzahl Patienten mit nicht-fehlerndem Wert zu Woche 24.								
<sup>(3)</sup> Werte zu Studienbeginn beziehen sich auf die gesamte HRQoL-Population (N = 149 im Luspatercept-Arm; N = 76 im Placebo-Arm).								
<sup>(4)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.								
BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i> ; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz nach Hedges' g								
Quelle: [19]								

Tabelle 4-60: Zusatzanalyse 3: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Mittlere Änderung der gesundheitsbezogenen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 48

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>Körperliche Funktion</b>								
73	66,85 (21,82)	66,91 (21,44)	13	65,26 (19,99)	70,99 (18,28)	-5,88 (-12,88; 1,13)	0,100	-0,28 (-0,55; 0,00)
<b>Rollenfunktion</b>								
73	65,66 (30,46)	65,47 (30,29)	13	64,25 (27,72)	65,43 (29,65)	-2,12 (-12,66; 8,43)	0,693	-0,07 (-0,35; 0,20)
<b>Emotionale Funktion</b>								
73	76,81 (20,17)	75,23 (22,03)	13	76,78 (19,15)	73,30 (22,45)	-1,93 (-10,30; 6,44)	0,649	-0,10 (-0,37; 0,18)

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>Kognitive Funktion</b>								
73	82,89 (19,66)	83,78 (19,00)	13	80,22 (21,52)	79,94 (18,14)	3,05 (-3,95; 10,05)	0,391	0,15 (-0,13; 0,43)
<b>Soziale Funktion</b>								
73	75,95 (27,01)	74,62 (27,42)	13	71,56 (28,44)	71,30 (30,96)	-3,96 (-14,30; 6,38)	0,450	-0,14 (-0,42; 0,13)
<b>Gesundheitlicher Status insgesamt</b>								
73	57,72 (21,00)	59,08 (20,55)	13	58,67 (19,49)	60,34 (18,94)	-0,28 (-7,81; 7,24)	0,941	-0,01 (-0,29; 0,26)
Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline.								
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).								
<sup>(2)</sup> Anzahl Patienten mit nicht-fehlendem Wert zu Woche 48.								
<sup>(3)</sup> Werte zu Studienbeginn beziehen sich auf die gesamte HRQoL-Population (N = 149 im Luspatercept-Arm; N = 76 im Placebo-Arm).								
<sup>(4)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.								
BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i> ; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz nach Hedges' g								
Quelle: [19]								

Auch bei Betrachtung der Ergebnisse zur mittleren Änderung der Lebensqualität unter Berücksichtigung aller Werte im Studienverlauf ergaben sich für die Funktionsskalen des EOTC QLQ-C30, mit Ausnahme der Funktionsskala körperliche Funktion, keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen. Dies trifft sowohl auf die primäre Behandlungsphase bis Woche 24 als auch auf die weitere Behandlung bis Woche 48 zu (siehe Tabelle 4-59 & Tabelle 4-60).

Einzig für die Funktionsskala körperliche Funktion zeigte sich in Woche 1 – 24 ein statistisch signifikanter Unterschied zuungunsten von Luspatercept (LS-MWD(Woche 1 – 24) = -6,92; 95 %-KI: [-11,80; -2,04]; p = 0,006), der jedoch nicht klinisch relevant war (SMD(Woche 1 – 24) = -0,32; 95 %-KI: [-0,60; -0,05]).

**Zusammenfassung der Ergebnisse zum Endpunkt Lebensqualität – EORTC QLQ-C30**

Der EORTC QLQ-C30 bildet die Lebensqualität bei onkologischen Patienten indikationsübergreifend ab. Im Kontext der MDS werden über diesen Fragebogen insbesondere Aspekte der Grunderkrankung wie der Einfluss von Fatigue auf das Wohlbefinden und die Ausführung alltäglicher Aufgaben erhoben. Im Rahmen der Studie MEDALIST wurde in beiden Behandlungsarmen entsprechend der derzeitigen Therapieempfehlungen eine adäquate Behandlung im Sinne einer BSC sichergestellt. Erkrankungsbezogene Symptome und damit verbundene Beeinträchtigungen wurden somit weitestgehend vermieden. Ziel der Therapie war somit eine stabile Aufrechterhaltung der Lebensqualität – auch bei Reduktion der supportiven Maßnahmen wie der Transfusionstherapie. Dieses Therapieziel wurde unter Luspatercept über den gesamten Studienverlauf hinweg erreicht.

Zusammenfassend wurde die generische Lebensqualität in beiden Behandlungsarmen und über den gesamten Studienverlauf hinweg konstant gut erhalten. Klinisch relevante Unterschiede zwischen den beiden Behandlungsarmen auf Ebene der einzelnen Domänen lagen nicht vor.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

**Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

**Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.1.8 Lebensqualität gemessen anhand des QoL-E-Fragebogens

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-61: Operationalisierung des Endpunkts Lebensqualität gemessen anhand des QoL-E-Fragebogens

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Die Lebensqualität wurde anhand des krankheitsspezifischen Fragebogens QoL-E (Version 3.0) anhand von sechs Skalen erhoben:</p> <ul style="list-style-type: none"> <li>• körperliches Wohlbefinden (Item 3a bis 3d)</li> <li>• funktionales Wohlbefinden (Item 4a bis 4b, 5)</li> <li>• soziales/familiäres Wohlbefinden (Item 6a bis 6c, 7)</li> <li>• sexuelles Wohlbefinden (Item 8, 14f)</li> <li>• Fatigue (Item 9, 10, 11a bis 11d, 12)</li> <li>• MDS-bezogene Störungen (Item 13, 14a bis 14e, 14g)</li> </ul> <p>Aus den einzelnen Skalen werden die folgenden Summenskalen gebildet und dargestellt:</p> <ul style="list-style-type: none"> <li>• allgemeines Wohlbefinden (Summe aller Domänen ohne MDS-spezifische Störungen)</li> <li>• Behandlungseffekt (<i>Treatment Outcome Index</i>, TOI; Summe der physischen, funktionellen und MDS-spezifischen Dimensionen)</li> <li>• Gesamtsummenskala (Summe aller Domänen)</li> </ul> <p>Die krankheitsspezifische Lebensqualität wurde zur Baseline (Dosis 1 Tag 1), zu jedem zweiten Zyklus (Zyklus 1, 3, 5, 7) und der Woche 25-Visite, jeweils vor Verabreichung der Studienmedikation erhoben. Für Patienten, die nach der Woche-25-Visite weiterhin doppelblind behandelt wurden, wurde die gesundheitsbezogene Lebensqualität zu jedem zweiten Zyklus vor Verabreichung der Studienmedikation erhoben.</p> <p>In die Analysen der Änderung der krankheitsspezifischen Lebensqualität gehen nur solche Patienten ein, die zur Baseline und zu mindestens einem Auswertungszeitpunkt im Studienverlauf einen ausgefüllten QoL-E-Fragebogen aufwiesen (HRQoL-Population). Ein Fragebogen galt dann als ausgefüllt, wenn mindestens die Hälfte aller Fragen (hier &gt; 14 Fragen) beantwortet wurden.</p> <p>Auswertungen zu den Rücklaufquoten des Fragebogens beziehen sich auf die ITT-Population. Höhere Antwortwerte der Domänen und Summenscores des QoL-E entsprechen einer höheren Lebensqualität.</p> <p>Skalen, für die weniger als die Hälfte der Items beantwortet waren, wurden entsprechend der Vorgaben im Handbuch des QoL-E als fehlende Werte bewertet. Eine Imputation fehlender Werte wurde nicht durchgeführt.</p> <p>Änderungen der gesundheitsbezogenen Lebensqualität im Studienverlauf wurden mittels MMRM berechnet. In das Modell flossen Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte ein. SMD inkl. 95 %-KI in Form von Hedges' g werden ergänzend dargestellt, um klinisch relevante Effekte zu bewerten.</p> <p>Änderungen der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten im Vergleich zur Baseline werden grafisch dargestellt, inkl. tabellarischer, deskriptiver Listing der Werte im Studienverlauf. Unterschiede zwischen den Behandlungsarmen wurden mittels ANOVA berechnet. Statistisch signifikante Unterschiede (<math>p &lt; 0,05</math>) wurden in den Grafiken markiert.</p>

Studie	Operationalisierung
	<p>Auswertungen zu den Datenschnitten aus dem Jahr 2019 werden nicht dargestellt, da die Rücklaufquoten sehr gering waren (insbesondere im Placebo-Arm hatten &lt; 10 Patienten einen ausgefüllten Fragebogen) und somit weitere Aussagen zum Zusatznutzen von Luspatercept nicht möglich sind.</p> <p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Hauptanalyse:</b> Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 24 (siehe Tabelle 4-64)</li> <li>• <b>Zusatzanalyse 1:</b> Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 48 (siehe Tabelle 4-65)</li> <li>• <b>Zusatzanalyse 2:</b> Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten; Woche 1 – 48 (siehe Abbildung 4-47 ff. &amp; Tabelle 4-66)</li> </ul>
<p>ANOVA = <i>analysis of variance</i>; HRQoL-Population = Population mit auswertbarer gesundheitsbezogener Lebensqualität; IPSS-R = <i>International Prognostic Scoring System-revised</i>; KI = Konfidenzintervall; MMRM = gemischtes Modell für wiederholte Messungen (<i>mixed-model repeated measure</i>); QoL-E = <i>Quality of Life-E</i>; SMD = standardisierte Mittelwertdifferenz</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-62: Bewertung des Verzerrungspotenzials für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens in MEDALIST

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
MEDALIST	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Lebensqualität – QoL-E wurde verblindet erhoben. In die Analyse gingen nur solche Patienten ein, die zu Studienbeginn und mindestens einem Zeitpunkt im Studienverlauf einen QoL-E-Fragebogen ausgefüllt hatten (HRQoL-Population). Zwischen der HRQoL-Population und der ITT-Population liegen keine wesentlichen Unterschiede hinsichtlich der Baseline-Charakteristika vor [1, 95]. Darüber hinaus umfasst die HRQoL-Population > 95 % der Patienten der ITT-Population (Luspatercept: 149/153 (97,4 %); Placebo: 76/76 (100,0 %)), sodass von einer guten Vergleichbarkeit der HRQoL- und ITT-Population auszugehen ist (siehe Tabelle 4-6). Die Rücklaufquoten bezogen auf die ITT-Population liegen für den QoL-E bis Woche 25 über den vom G-BA geforderten 70 % und sind damit ausreichend hoch, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen (siehe Tabelle 4-63). Ergebnisse nach Woche 25 wurden aufgrund geringerer Rücklaufquoten (< 70 %) nur ergänzend berichtet. Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens und eine hohe Aussagesicherheit der Ergebnisse.

Tabelle 4-63: Rücklaufquoten des QoL-E-Fragebogens aus MEDALIST

Studie MEDALIST Rücklaufquoten des QoL-E- Fragebogens	Luspatercept + BSC <sup>(1)</sup> N = 153		Placebo + BSC <sup>(1)</sup> N = 76	
	Patienten unter Therapie zum jeweiligen Erhebungs- zeitpunkt	Patienten mit ausgefülltem Fragebogen zum jeweiligen Erhebungs- zeitpunkt <sup>(2)</sup> n (%)	Patienten unter Therapie zum jeweiligen Erhebungs- zeitpunkt	Patienten mit ausgefülltem Fragebogen zum jeweiligen Erhebungs- zeitpunkt <sup>(2)</sup> n (%)
Baseline / Zyklus 1	153	151 (98,7)	76	76 (100)
Woche 7 / Zyklus 3	151	133 (88,1)	76	70 (92,1)
Woche 13 / Zyklus 5	146	130 (89,0)	73	67 (91,8)
Woche 19 / Zyklus 7	134	111 (82,9)	70	60 (85,7)
Woche 25	127	111 (87,4)	68	53 (77,8)
Zyklus 1 der EP	100	48 (48,0)	23	9 (39,1)
Zyklus 3 der EP	94	81 (86,2)	19	16 (84,2)
Zyklus 5 der EP	89	79 (88,8)	16	15 (93,8)
Zyklus 7 der EP	86	75 (87,2)	14	13 (92,9)
Zyklus 9 der EP	75	65 (86,7)	10	8 (80,0)
Zyklus 11 der EP	60	51 (85,0)	7	7 (100,0)
Zyklus 13 der EP	47	43 (91,5)	5	5 (100,0)

Studie MEDALIST Rücklaufquoten des QoL-E- Fragebogens	Luspatercept + BSC <sup>(1)</sup> N = 153		Placebo + BSC <sup>(1)</sup> N = 76	
	Patienten unter Therapie zum jeweiligen Erhebungs- zeitpunkt	Patienten mit ausgefülltem Fragebogen zum jeweiligen Erhebungs- zeitpunkt <sup>(2)</sup> n (%)	Patienten unter Therapie zum jeweiligen Erhebungs- zeitpunkt	Patienten mit ausgefülltem Fragebogen zum jeweiligen Erhebungs- zeitpunkt <sup>(2)</sup> n (%)
Zyklus 15 der EP	33	25 (16,3)	4	4 (5,3)
Zyklus 17 der EP	29	25 (16,3)	3	2 (2,6)
Zyklus 19 der EP	18	16 (10,5)	1	1 (1,3)
Zyklus 21 der EP	12	12 (7,8)	1	1 (1,3)
Zyklus 23 der EP	5	5 (3,3)	1	1 (1,3)
Zyklus 25 der EP	3	3 (2,0)	0	0
Zyklus 27 der EP	2	2 (1,3)	0	0
Zyklus 29 der EP	2	2 (1,3)	0	0
Zyklus 31 der EP	2	2 (1,3)	0	0
EOT	71	46 (30,1)	54	28 (36,8)
Datenschnitt: 08.05.2018; Population: HRQoL-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> Prozentualer Anteil bezogen auf die ITT-Population.  BSC = Best-Supportive-Care; EOT = <i>end of treatment</i> ; EP = Extensionsphase; HRQoL = <i>health-related quality of life</i> ; ITT = Intention-to-Treat; QoL-E = <i>Quality of Life-E</i> Quelle: [19, 95]				

*Stellen Sie die Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

### Hauptanalyse: Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf (MMRM-Analyse)

Tabelle 4-64: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 24

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>körperliches Wohlbefinden</b>								
110	53,68 (20,67)	54,32 (22,68)	53	51,48 (22,91)	54,25 (23,89)	-4,95 (-11,28; 1,39)	0,125	-0,23 (-0,51; 0,05)
<b>funktionales Wohlbefinden</b>								
108	55,18 (33,69)	54,22 (32,71)	53	51,02 (29,55)	61,01 (32,44)	-6,11 (-15,16; 2,93)	0,184	-0,19 (-0,47; 0,09)
<b>soziales/familiäres Wohlbefinden</b>								
102	48,91 (37,73)	48,04 (34,77)	48	47,43 (37,69)	53,65 (37,54)	-8,14 (-16,93; 0,66)	0,070	-0,21 (-0,51; 0,08)
<b>sexuelles Wohlbefinden</b>								
105	62,12 (35,00)	65,87 (35,02)	51	62,68 (39,18)	61,93 (36,71)	-0,13 (-9,02; 8,75)	0,976	0,00 (-0,29; 0,28)
<b>Fatigue</b>								
110	75,55 (14,41)	73,33 (14,74)	53	73,68 (14,20)	77,54 (14,48)	-4,92 (-8,89; -0,96)	0,015	-0,34 (-0,62; -0,07)
<b>MDS-bezogene Störungen</b>								
109	56,71 (24,20)	60,22 (25,16)	53	57,43 (23,05)	58,89 (24,51)	-2,05 (-7,94; 3,83)	0,492	-0,09 (-0,36; 0,19)
<b>allgemeines Wohlbefinden<sup>(3)</sup></b>								
100	59,33 (21,24)	59,23 (20,34)	48	57,38 (20,78)	62,35 (21,22)	-6,16 (-11,09; -1,23)	0,015	-0,29 (-0,58; 0,00)
<b>Gesamtsummenskala<sup>(4)</sup></b>								
100	58,16 (21,23)	59,57 (21,67)	48	57,72 (21,10)	61,33 (21,75)	-5,04 (-10,33; 0,25)	0,062	-0,24 (-0,53; 0,06)

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 24	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>Behandlungseffekt<sup>(5)</sup></b>								
107	55,40 (20,89)	56,67 (22,41)	53	53,40 (20,14)	58,05 (21,82)	-4,64 (-10,15; 0,88)	0,100	-0,22 (-0,50; 0,06)
<p>Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline.</p> <p><sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).</p> <p><sup>(2)</sup> Anzahl Patienten mit nicht-fehlerndem Wert zu Woche 24.</p> <p><sup>(3)</sup> Werte zu Studienbeginn beziehen sich auf die gesamte HRQoL-Population (N = 149 im Luspatercept-Arm; N = 76 im Placebo-Arm).</p> <p><sup>(4)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.</p> <p>BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i>; IPSS-R = <i>International Prognostic Scoring System-revised</i>; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate (<i>least square</i>); MW = Mittelwert; MWD = Mittelwertdifferenz; QoL-E = <i>Quality of Life-E</i>; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz nach Hedges' g</p> <p>Quelle: [19]</p>								

In den 24 Wochen der primären Behandlungsphase zeigten sich für den Großteil der Domänen und Summenscores des krankheitsspezifischen Fragebogens QoL-E hinsichtlich der mittleren Änderung der krankheitsspezifischen Lebensqualität keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen. Einzig für die Domäne Fatigue und den Summenscore allgemeines Wohlbefinden, welcher alle Domänen mit Ausnahme der MDS-bezogenen Störungen zusammenfasst, zeigte sich in Woche 1 – 24 ein statistisch signifikanter Unterschied zuungunsten von Luspatercept (LS-MWD(Fatigue) = -4,92; 95 %-KI: [-8,89; -0,96]; p = 0,015; LS-MWD(allgemeines Wohlbefinden) = 6,16; 95 %-KI: [-11,09; -1,23]; p = 0,015), der jedoch in beiden Fällen nicht klinisch relevant war (SMD(Fatigue) = -0,34; 95 %-KI: [-0,62; -0,07]; SMD(allgemeines Wohlbefinden) = -0,29; 95 %-KI: [-0,58; 0,00]; siehe Tabelle 4-64).

### Zusatzanalyse 1: Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf (MMRM-Analyse)

Tabelle 4-65: Zusatzanalyse 1: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität im Studienverlauf (MMRM-Analyse); Woche 1 – 48

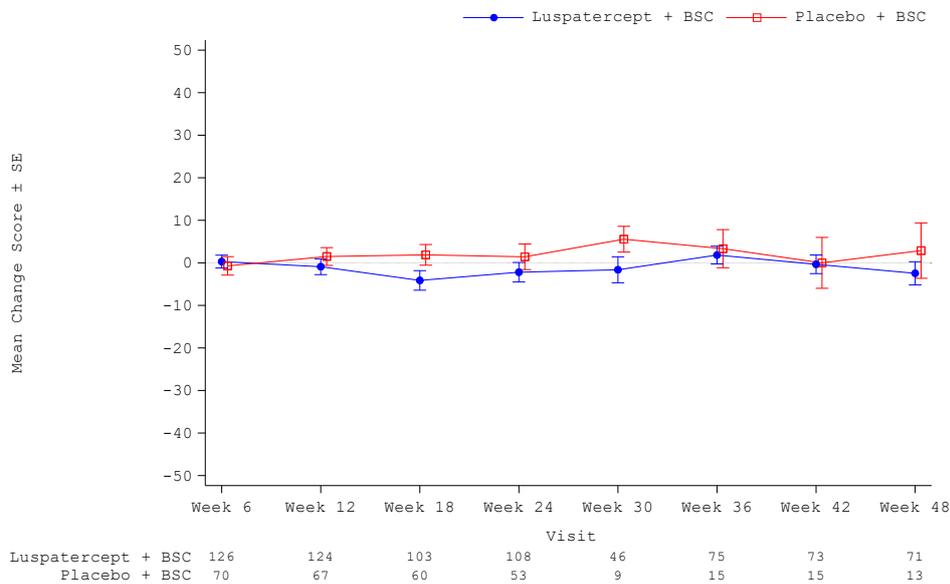
Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b><i>körperliches Wohlbefinden</i></b>								
73	53,68 (20,67)	54,32 (22,68)	13	51,48 (22,91)	54,25 (23,89)	-1,53 (-9,60; 6,55)	0,709	-0,07 (-0,35; 0,21)
<b><i>funktionales Wohlbefinden</i></b>								
73	55,18 (33,69)	54,22 (32,71)	13	51,02 (29,55)	61,01 (32,44)	-7,96 (-20,55; 4,63)	0,214	-0,25 (-0,52; 0,03)
<b><i>soziales/familiäres Wohlbefinden</i></b>								
65	48,91 (37,73)	48,04 (34,77)	13	47,43 (37,69)	53,65 (37,54)	-3,57 (-16,69; 9,56)	0,592	-0,09 (-0,38; 0,20)
<b><i>sexuelles Wohlbefinden</i></b>								
68	62,12 (35,00)	65,87 (35,02)	13	62,68 (39,18)	61,93 (36,71)	-1,09 (-14,06; 11,89)	0,869	-0,03 (-0,31; 0,25)
<b><i>Fatigue</i></b>								
73	75,55 (14,41)	73,33 (14,74)	13	73,68 (14,20)	77,54 (14,48)	-3,59 (-9,11; 1,92)	0,200	-0,25 (-0,53; 0,03)
<b><i>MDS-bezogene Störungen</i></b>								
72	56,71 (24,20)	60,22 (25,16)	13	57,43 (23,05)	58,89 (24,51)	2,79 (-5,68; 11,26)	0,517	0,12 (-0,16; 0,39)
<b><i>allgemeines Wohlbefinden<sup>(3)</sup></i></b>								
65	59,33 (21,24)	59,23 (20,34)	13	57,38 (20,78)	62,35 (21,22)	-5,49 (-12,41; 1,43)	0,119	-0,26 (-0,55; 0,03)
<b><i>Gesamtsummenskala<sup>(4)</sup></i></b>								
64	58,16 (21,23)	59,57 (21,67)	13	57,72 (21,10)	61,33 (21,75)	-3,91 (-11,18; 3,36)	0,289	-0,18 (-0,48; 0,11)

Studie MEDALIST								
Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>		
N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	N <sup>(2)</sup>	MW (SD) zu Studien- beginn <sup>(3)</sup>	MW (SD) zu Woche 48	LS-MWD (95 %-KI)	p- Wert <sup>(4)</sup>	SMD (95 %-KI)
<b>Behandlungseffekt<sup>(5)</sup></b>								
72	55,40 (20,89)	56,67 (22,41)	13	53,40 (20,14)	58,05 (21,82)	-3,84 (-11,63; 3,96)	0,332	-0,19 (-0,46; 0,09)
<p>Datenschnitt: 08.05.2018; Population: HRQoL-Population; Stratifizierung: IPSS-R Risikogruppe und durchschnittliche Transfusionslast zur Baseline.</p> <p><sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).</p> <p><sup>(2)</sup> Anzahl Patienten mit nicht-fehlerndem Wert zu Woche 48.</p> <p><sup>(3)</sup> Werte zu Studienbeginn beziehen sich auf die gesamte HRQoL-Population (N = 149 im Luspatercept-Arm; N = 76 im Placebo-Arm).</p> <p><sup>(4)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktionsterm als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.</p> <p>BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i>; IPSS-R = <i>International Prognostic Scoring System-revised</i>; KI = Konfidenzintervall; LS = Methode der kleinsten Quadrate (<i>least square</i>); MW = Mittelwert; MWD = Mittelwertdifferenz; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz nach Hedges' g</p> <p>Quelle: [19]</p>								

Auswertungen über den Studienzeitraum bis Woche 48 werden aufgrund der Unsicherheit hinsichtlich der Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten sowie Rücklaufquoten < 70 % nach Woche 24 ergänzend als Ereigniszeitanalysen dargestellt.

Konsistent zu den Ergebnissen der Hauptanalyse bis Woche 24 zeigten sich auch bei Betrachtung der Ergebnisse bis Woche 48 keine statistisch signifikanten und klinisch relevanten Unterschiede zwischen den beiden Behandlungsarmen (siehe Tabelle 4-65).

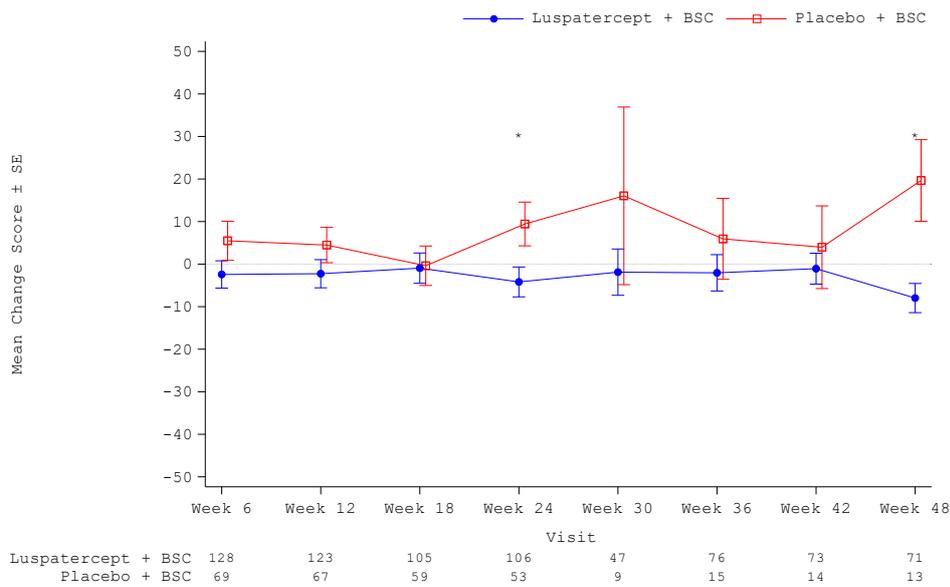
### Zusatzanalyse 2: Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten



Datenschnitt: 08.05.2018

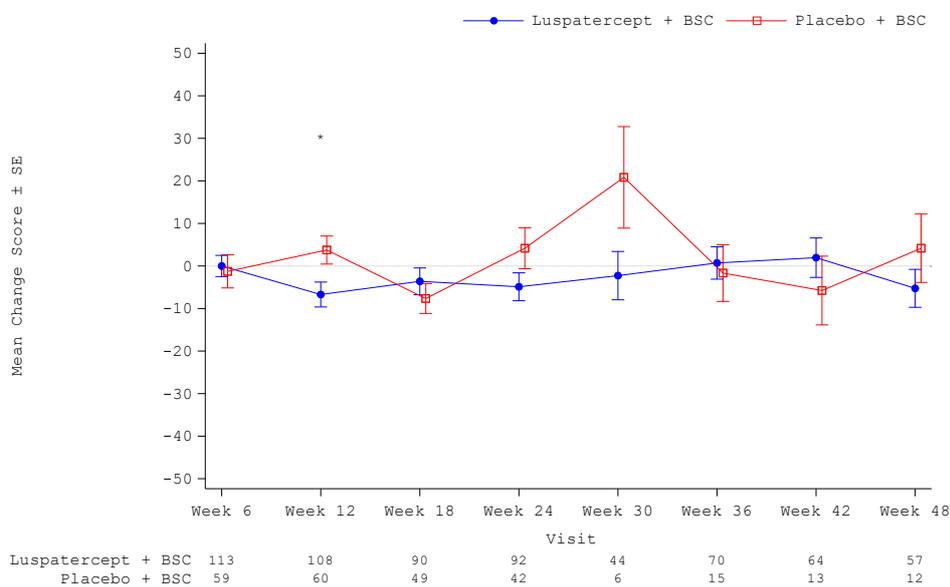
Quelle: [19]

Abbildung 4-47: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – körperliches Wohlbefinden; Woche 1 – 48



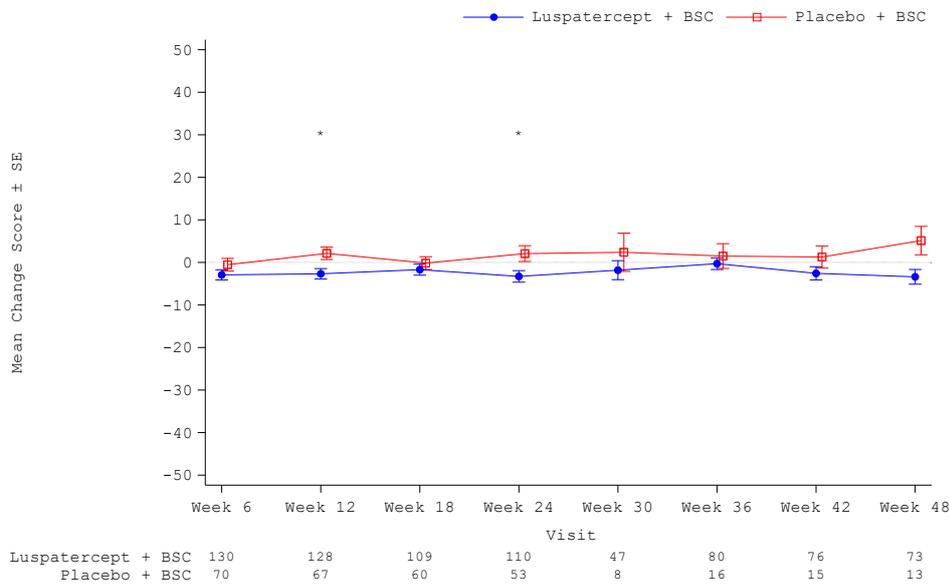
Datenschnitt: 08.05.2018; \* = statistisch signifikanter Unterschied, berechnet mittels ANOVA (*analysis of variance*)  
 Quelle: [19]

Abbildung 4-48: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – funktionales Wohlbefinden; Woche 1 – 48



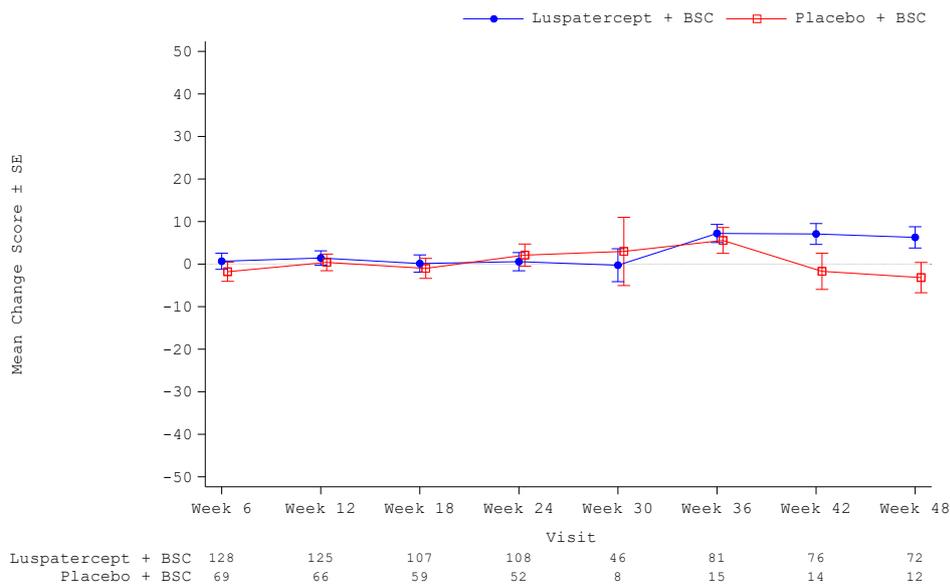
Datenschnitt: 08.05.2018; \* = statistisch signifikanter Unterschied, berechnet mittels ANOVA (*analysis of variance*)  
 Quelle: [19]

Abbildung 4-49: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – soziales/familiäres Wohlbefinden; Woche 1 – 48



Datenschnitt: 08.05.2018; \* = statistisch signifikanter Unterschied, berechnet mittels ANOVA (*analysis of variance*)  
 Quelle: [19]

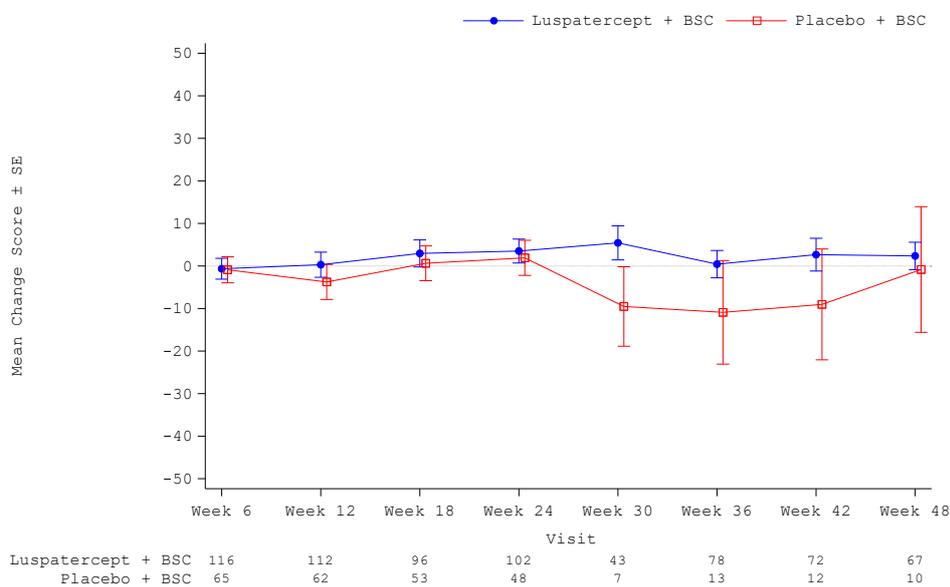
Abbildung 4-50: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – Fatigue; Woche 1 – 48



Datenschnitt: 08.05.2018

Quelle: [19]

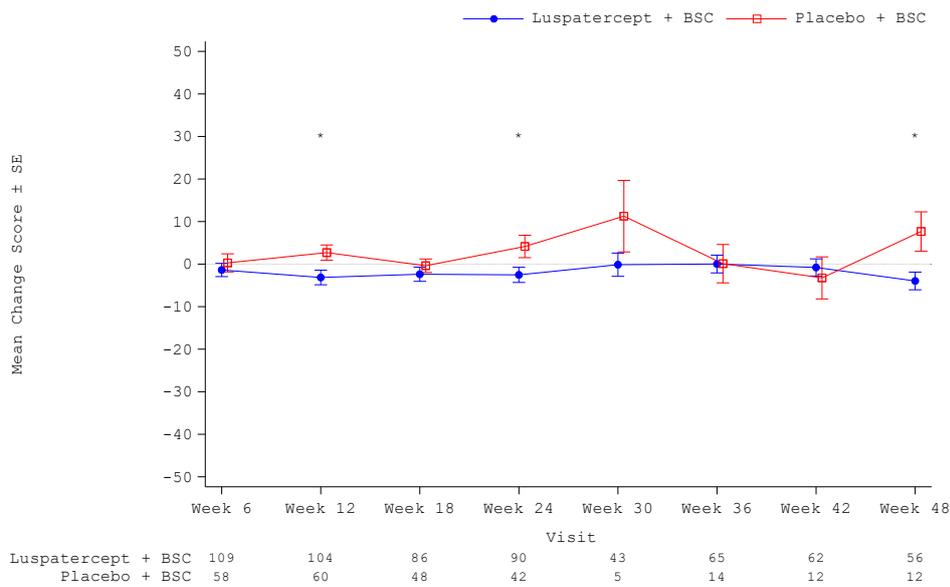
Abbildung 4-51: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – MDS-bezogene Störungen; Woche 1 – 48



Datenschnitt: 08.05.2018

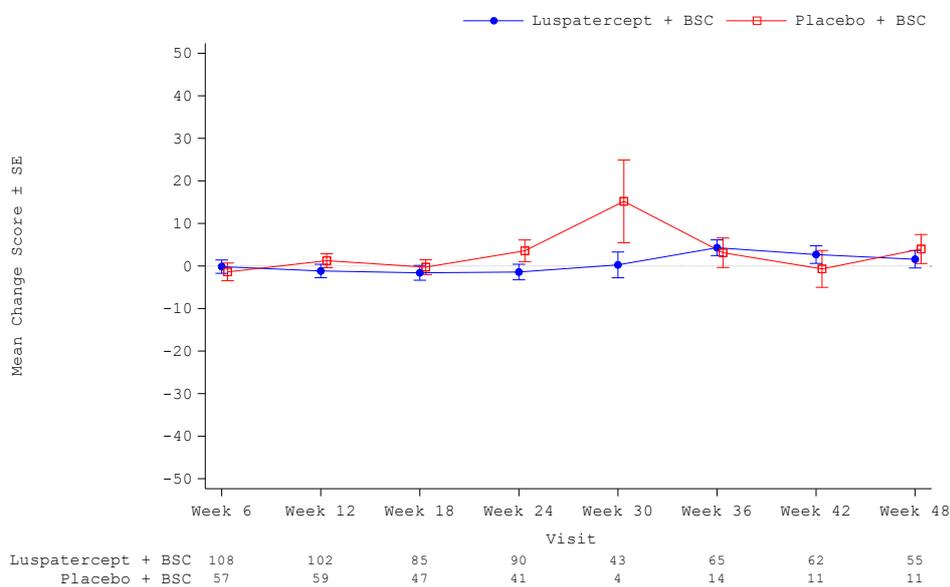
Quelle: [19]

Abbildung 4-52: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – sexuelles Wohlbefinden; Woche 1 – 48



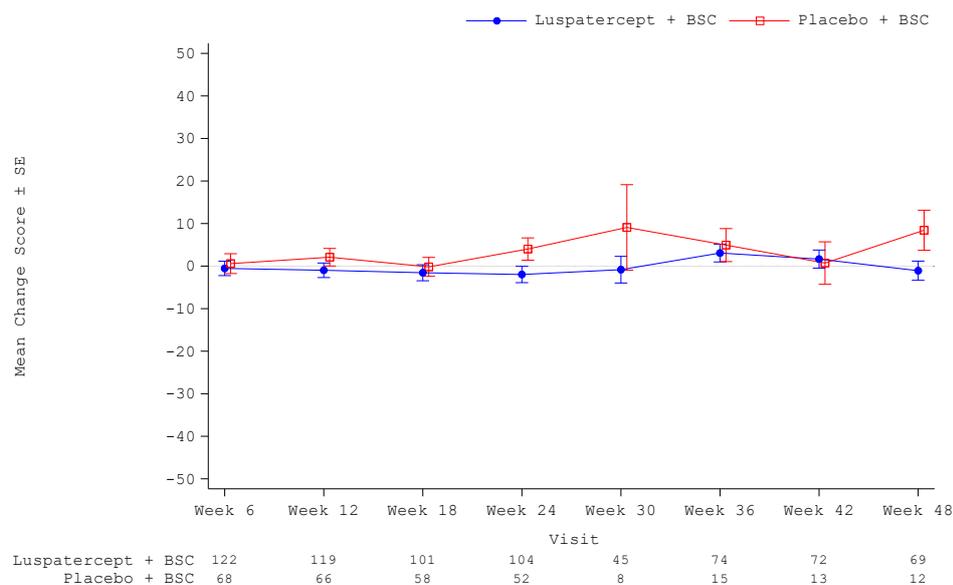
Datenschnitt: 08.05.2018; \* = statistisch signifikanter Unterschied, berechnet mittels ANOVA (*analysis of variance*)  
 Quelle: [19]

Abbildung 4-53: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – allgemeines Wohlbefinden; Woche 1 – 48



Datenschnitt: 08.05.2018  
 Quelle: [19]

Abbildung 4-54: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – Gesamtsummenskala; Woche 1 – 48



Datenschnitt: 08.05.2018

Quelle: [19]

Abbildung 4-55: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Mittlere Änderung der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten – Behandlungseffekt; Woche 1 – 48

Tabelle 4-66: Zusatzanalyse 2: Ergebnisse für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Werte der krankheitsspezifischen Lebensqualität zu den Erhebungszeitpunkten; Woche 1 – 48

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
<b>körperliches Wohlbefinden</b>				
Baseline	146 (98,0)	53,68 (20,67)	76 (100)	51,48 (22,91)
Woche 6	128 (85,9)	54,39 (20,37)	70 (92,1)	52,68 (22,61)
Woche 12	127 (85,2)	52,36 (23,29)	67 (88,2)	52,24 (22,29)
Woche 18	106 (71,1)	52,48 (25,32)	60 (78,9)	54,79 (21,48)
Woche 24	110 (73,8)	54,32 (22,68)	53 (69,7)	54,25 (23,89)
Woche 30	46 (30,9)	53,53 (22,46)	9 (11,8)	58,33 (20,73)
Woche 42	75 (50,3)	56,83 (20,48)	15 (19,7)	51,67 (26,67)
Woche 48	73 (49,0)	54,28 (21,37)	13 (17,1)	46,15 (31,20)
<b>funktionales Wohlbefinden</b>				
Baseline	146 (98,0)	55,18 (33,69)	76 (100)	51,02 (29,55)
Woche 6	130 (87,2)	54,62 (34,48)	69 (90,8)	55,23 (32,34)

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
Woche 12	125 (83,9)	54,04 (32,59)	67 (88,2)	54,73 (31,09)
Woche 18	108 (72,5)	57,61 (33,39)	59 (77,6)	51,79 (29,80)
Woche 24	108 (72,5)	54,22 (32,71)	53 (69,7)	61,01 (32,44)
Woche 30	47 (31,5)	56,03 (35,44)	9 (11,8)	71,60 (34,74)
Woche 36	78 (52,3)	55,27 (32,43)	15 (19,7)	56,30 (27,04)
Woche 42	75 (50,3)	58,07 (34,01)	14 (18,4)	49,21 (27,81)
Woche 48	73 (49,0)	51,90 (31,48)	13 (17,1)	61,54 (35,02)
<b>soziales/familiäres Wohlbefinden</b>				
Baseline	137 (91,9)	48,91 (37,73)	68 (89,5)	47,43 (37,69)
Woche 6	124 (83,2)	50,20 (36,83)	63 (82,9)	46,83 (38,23)
Woche 12	116 (77,9)	44,18 (36,62)	63 (82,9)	50,79 (36,74)
Woche 18	99 (66,4)	47,22 (38,78)	55 (72,4)	45,00 (37,70)
Woche 24	102 (68,5)	48,04 (34,77)	48 (63,2)	53,65 (37,54)
Woche 30	45 (30,2)	43,33 (41,08)	7 (9,2)	50,00 (47,87)
Woche 36	77 (51,7)	51,62 (35,67)	16 (21,1)	56,25 (38,19)
Woche 42	71 (47,7)	55,63 (35,66)	14 (18,4)	55,36 (39,44)
Woche 48	65 (43,6)	51,92 (36,53)	13 (17,1)	48,08 (38,81)
<b>Sexuelles Wohlbefinden</b>				
Baseline	141 (94,6)	62,12 (35,00)	71 (93,4)	62,68 (39,18)
Woche 6	121 (81,2)	60,74 (36,76)	68 (89,5)	62,75 (41,32)
Woche 12	116 (77,9)	61,57 (37,36)	64 (84,2)	61,46 (41,35)
Woche 18	100 (67,1)	64,67 (36,99)	56 (73,7)	63,10 (41,34)
Woche 24	105 (70,5)	65,87 (35,02)	51 (67,1)	61,93 (36,71)
Woche 30	45 (30,2)	63,70 (38,07)	7 (9,2)	59,52 (38,32)
Woche 36	78 (52,3)	61,32 (35,05)	15 (19,7)	45,56 (35,62)
Woche 42	74 (49,7)	65,99 (35,19)	14 (9,4)	47,02 (41,03)
Woche 48	68 (45,6)	69,61 (35,08)	13 (17,1)	51,28 (40,23)
<b>Fatigue</b>				
Baseline	149 (100)	75,55 (14,41)	76 (100)	73,68 (14,20)
Woche 6	130 (87,2)	73,08 (15,03)	70 (92,1)	72,72 (15,06)
Woche 12	128 (85,9)	73,18 (15,65)	67 (88,2)	75,76 (12,99)
Woche 18	109 (73,2)	74,88 (15,22)	60 (78,9)	74,44 (13,32)
Woche 24	110 (73,8)	73,33 (14,74)	53 (69,7)	77,54 (14,48)
Woche 30	47 (31,5)	74,57 (12,17)	8 (10,5)	77,98 (9,16)

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
Woche 36	80 (53,7)	75,83 (13,97)	16 (21,1)	75,00 (13,41)
Woche 42	76 (51,0)	74,69 (14,26)	15 (19,7)	75,24 (12,88)
Woche 48	73 (49,0)	73,71 (15,04)	13 (17,1)	76,92 (13,59)
<b>MDS-bezogene Störungen</b>				
Baseline	148 (99,3)	56,71 (24,20)	75 (98,7)	57,43 (23,05)
Woche 6	129 (86,6)	57,97 (24,85)	70 (92,1)	56,05 (24,01)
Woche 12	126 (84,6)	58,37 (25,83)	67 (88,2)	58,07 (22,76)
Woche 18	108 (72,5)	57,52 (25,10)	60 (78,9)	58,10 (26,68)
Woche 24	109 (73,2)	60,22 (25,16)	53 (69,7)	58,89 (24,51)
Woche 30	47 (31,5)	57,24 (28,63)	8 (10,5)	69,35 (18,42)
Woche 36	81 (54,4)	63,93 (25,31)	16 (21,1)	62,80 (21,87)
Woche 42	76 (51,0)	63,82 (25,00)	15 (19,7)	59,84 (19,49)
Woche 48	72 (48,3)	65,44 (24,33)	13 (17,1)	53,11 (24,17)
<b>allgemeines Wohlbefinden<sup>(2)</sup></b>				
Baseline	135 (90,6)	59,33 (21,24)	68 (89,5)	57,38 (20,78)
Woche 6	120 (80,5)	58,79 (21,26)	62 (84,6)	57,89 (22,46)
Woche 12	113 (75,8)	57,57 (21,44)	63 (82,9)	59,56 (20,04)
Woche 18	96 (64,4)	60,02 (22,46)	54 (71,1)	57,73 (21,63)
Woche 24	100 (67,1)	59,23 (20,34)	48 (63,2)	62,35 (21,22)
Woche 30	44 (29,5)	58,33 (22,43)	6 (7,9)	64,34 (19,11)
Woche 36	73 (49,0)	59,75 (20,37)	15 (19,7)	57,96 (19,80)
Woche 42	70 (47,0)	63,28 (19,62)	13 (17,1)	53,96 (20,59)
Woche 48	65 (43,6)	60,54 (18,48)	13 (17,1)	56,79 (20,45)
<b>Gesamtsummenskala<sup>(3)</sup></b>				
Baseline	135 (90,6)	58,16 (21,23)	67 (88,2)	57,72 (21,10)
Woche 6	119 (79,9)	58,52 (21,77)	62 (84,6)	56,38 (22,44)
Woche 12	111 (74,5)	58,64 (22,32)	63 (82,9)	58,96 (20,09)
Woche 18	95 (63,8)	59,39 (23,05)	54 (71,1)	57,85 (22,64)
Woche 24	100 (67,1)	59,57 (21,67)	48 (63,2)	61,33 (21,75)
Woche 30	44 (29,5)	58,17 (24,70)	5 (6,6)	74,29 (14,21)
Woche 36	73 (49,0)	62,56 (21,88)	15 (19,7)	60,65 (19,20)
Woche 42	70 (47,0)	63,94 (21,08)	13 (17,1)	55,83 (18,34)
Woche 48	64 (43,0)	63,60 (19,66)	13 (17,1)	54,95 (21,27)

Studie MEDALIST Erhebungszeitpunkt	Luspatercept + BSC <sup>(1)</sup> N = 149		Placebo + BSC <sup>(1)</sup> N = 76	
	n (%)	MW (SD)	n (%)	MW (SD)
<b>Behandlungseffekt<sup>(4)</sup></b>				
Baseline	144 (96,6)	55,40 (20,89)	75 (98,7)	53,40 (20,14)
Woche 6	125 (83,9)	55,81 (21,41)	69 (90,8)	54,32 (21,04)
Woche 12	123 (82,6)	55,32 (21,63)	67 (88,2)	55,01 (20,81)
Woche 18	105 (70,5)	56,27 (23,06)	59 (77,6)	54,47 (20,54)
Woche 24	107 (71,8)	56,67 (22,41)	53 (69,7)	58,05 (21,82)
Woche 30	46 (30,9)	55,96 (23,95)	8 (10,5)	69,41 (14,79)
Woche 36	77 (51,7)	58,99 (21,94)	15 (19,7)	58,21 (19,06)
Woche 42	75 (50,3)	59,73 (21,81)	14 (18,4)	52,05 (20,77)
Woche 48	72 (48,3)	57,59 (20,54)	13 (17,1)	53,60 (27,98)
Datenschnitt: 08.05.2018; Population: HRQoL-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  BSC = Best-Supportive-Care; HRQoL = <i>health-related quality of life</i> ; MW = Mittelwert; SD = Standardabweichung Quelle: [19]				

Die mittlere Änderung der krankheitsspezifischen Lebensqualität zu den einzelnen Erhebungszeitpunkten ist für die einzelnen Domänen und Summenscores des QoL-E über den gesamten Studienverlauf bis Woche 48 grafisch dargestellt. Durch die Selektion der Patienten mit einem Therapieerfolg zu Woche 25, ist die Interpretierbarkeit der Ergebnisse ab Woche 25 mit Unsicherheit behaftet. Darüber hinaus liegen die Rücklaufquoten nach Woche 25 < 70 %, sodass die Aussagekraft der Ergebnisse limitiert ist.

Insgesamt bestätigen die grafischen Darstellungen der mittleren Änderung der krankheitsspezifischen Lebensqualität die Ergebnisse der Hauptanalyse. Für keine der Domänen oder Summenscores zeigte sich ein über die Behandlungsphase bis Woche 24 konstant statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Abbildung 4-47 – Abbildung 4-55 und Tabelle 4-66). Insgesamt ist die Lebensqualität unter Luspatercept stabil.

**Zusammenfassung der Ergebnisse zum Endpunkt Lebensqualität – QoL-E**

Zusammenfassend wurde die Lebensqualität in beiden Behandlungsarmen und über den gesamten Studienverlauf hinweg konstant gut erhalten. Klinisch relevante Unterschiede zwischen den beiden Behandlungsarmen auf Ebene der einzelnen Domänen und Summenscores lagen nicht vor.

In der Zusammenschau der Ergebnisse ist ein Einzelitem des QoL-E im Kontext der Erkrankung und Bewertung des Therapieeffekts von Luspatercept besonders zu betonen. Hinsichtlich der Frage, ob die Transfusionsabhängigkeit das tägliche Leben beeinträchtigt (Item 14a), antwortete zu Woche 24 fast die Hälfte (44,5 % (49/110)) der mit Luspatercept behandelten Patienten mit „überhaupt nicht“. Im Vergleich zur Baseline, zu der nur 14,5 % der Patienten (22/149) im Luspatercept-Arm mit „überhaupt nicht“ antworteten, entspricht dies einem dreifach erhöhten Anteil an Patienten, die unter der Therapie mit Luspatercept keine tägliche Beeinträchtigung durch die Transfusionsabhängigkeit (mehr) wahrnehmen. Im Placebo-Arm blieb der Anteil mit 25,0 % (19/76) zur Baseline und 20,8 % (11/53) zu Woche 24 quasi unverändert. Dies entspricht einem RR von 2,15 (95 %-KI: [1,22; 3,78];  $p = 0,003$ ) im Vergleich zwischen Luspatercept und Placebo zu Woche 24 und unterstreicht den direkten Einfluss der Vermeidung bzw. Einsparung von Transfusionen auf die Wahrnehmung der Patienten hinsichtlich der Beeinträchtigung des täglichen Lebens durch Transfusionen [95].

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

### **Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

### **Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.1.9 Sicherheit

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-67: Operationalisierung des Endpunkts Sicherheit

Studie	Operationalisierung
MEDALIST	<p><b>Operationalisierung:</b></p> <p>Die Sicherheit wird anhand des Auftretens von UE erfasst, die im Studienverlauf seit Behandlungsbeginn aufgetreten sind oder sich verschlechtert haben. In den Auswertungen wurden alle UE berücksichtigt, auch unabhängig davon, ob ein Kausalzusammenhang zur Studienmedikation hergestellt werden konnte.</p> <p>UE werden gemäß MedDRA (Version 20.0) kodiert und gemäß NCI CTCAE (Version 4.03) nach Schweregrad klassifiziert. UE, dessen Schweregrad durch die CTCAE-Klassifizierung nicht definiert wird, werden analog anhand einer präspezifizierten Fünf-Punkte-Skala bewertet (Grad eins: mild; Grad zwei: moderat; Grad drei: schwer; Grad vier: lebensbedrohlich; Grad fünf: tödlich).</p> <p>Als SUE wurden solche Ereignisse berücksichtigt, die</p> <ul style="list-style-type: none"> <li>- tödlich verliefen,</li> <li>- lebensbedrohlich waren,</li> <li>- eine Hospitalisierung erforderlich machten oder zur Verlängerung eines bestehenden stationären Aufenthalts führten,</li> <li>- in bestehenden oder bedeutsamen Behinderungen/Unfähigkeiten resultierten,</li> <li>- angeborene Anomalien/Geburtsdefekte waren, oder</li> <li>- wichtige medizinische Ereignisse waren, die nicht unmittelbar tödlich oder lebensbedrohlich waren oder eine Hospitalisierung oder Behinderung verursachten, aber den Patienten diesen Risiken aussetzen oder eine medizinische oder chirurgische Behandlung zur Prävention dieser Ereignisse nötig machten.</li> </ul> <p>Dargestellte Auswertungen beziehen sich auf die Safety-Population, welche alle Patienten der ITT-Population umfasst, die mindestens eine Dosis Studienmedikation erhalten haben. Patienten wurden dem Behandlungsarm zugeordnet, dessen Medikation sie tatsächlich erhalten hatten.</p> <p>Auswertungen zu UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen waren definiert als Gesamtraten der UE der Kategorien jegliche UE, SUE und schwere UE (CTCAE <math>\geq 3</math>) ohne Berücksichtigung der Ereignisse Schwindelgefühl, Fatigue, Dyspnoe und Belastungsdyspnoe, welche jeweils auf Ebene der PT ausgeschlossen wurden, sowie Ereignissen, die eine stationäre Aufnahme erforderlich machten.</p> <p>Auswertungen zur Kategorie Therapieabbruch wegen UE werden auf Ebene der SOC und PT rein deskriptiv dargestellt.</p> <p>Als UE von besonderem Interesse waren die SMQ „Malignitäten“ und „Prämaligne Erkrankungen“ prädefiniert. Ergänzend werden Auswertungen zu den Ereignissen Neutropenie und Thrombozytopenie auf Ebene der PT als UE von besonderem Interesse berichtet. Neben den Gesamtraten werden auch Auswertungen zu den Inzidenzen von schweren und schwerwiegenden UE von besonderem Interesse dargestellt.</p> <p>Auswertungen der Zeit bis zur Progression in eine AML sowie Inzidenzen der Progression in ein Hochrisiko-MDS (erhoben als Ereignisse des PT „Anämie refraktär mit Blastenüberschuss“ oder „Myelodysplastisches Syndrom“) werden ergänzend berichtet.</p> <p>Unterschiede zwischen den Behandlungsarmen während der primären Behandlungsphase (Woche 1 – 24) wurden mittels CMH-Tests ermittelt. Dargestellt werden die Ergebnisse als OR, RR und ARR, jeweils inkl. zweiseitigem 95 %-KI.</p>

Studie	Operationalisierung
	<p>Auswertungen zu UE bis zum Datenschnitt vom 01.07.2019 sowie der Zeit bis zur Progression in eine AML werden ergänzend in Form von Ereigniszeitanalysen dargestellt. Patienten, bei denen bis zum Datenschnitt vom 01.07.2019 kein UE/Ereignis aufgetreten ist, wurden zensiert. Die mediane Zeit bis zum UE/Ereignis wurde mittels Kaplan-Meier-Methode geschätzt und inkl. zugehörigem 95 %-KI, welches über eine Log-Log-Transformation nach Brookmeyer und Crowley berechnet wurde, dargestellt. Unterschiede zwischen den Behandlungsarmen wurden mittels Log-Rank-Test analysiert. Das HR inklusive zweiseitigem 95 %-KI für Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, wurde anhand eines Cox-Proportional-Hazards-Modells mit Behandlung als Kovariable geschätzt. Für Auswertungen der Zeit bis zur Progression in eine AML wurden die Auswertungen stratifiziert nach Transfusionslast zur Baseline und IPSS-R-Score zur Baseline vorgenommen, unter Berücksichtigung aller Ereignisse bis zum Auswertungszeitpunkt.</p> <p>Auswertungen zur Progression in ein Hochrisiko-MDS werden aufgrund sehr geringer Ereigniszahlen rein deskriptiv berichtet.</p> <p><b>Darstellung der Ergebnisse:</b></p> <ul style="list-style-type: none"> <li>• <b>Hauptanalyse 1:</b> Gesamtraten UE; Woche 1 – 24 (siehe Tabelle 4-69) <ul style="list-style-type: none"> <li>○ jegliche UE</li> <li>○ SUE</li> <li>○ schwere UE (CTCAE <math>\geq</math> 3)</li> <li>○ Therapieabbrüche aufgrund von UE</li> </ul> </li> <li>• <b>Sensitivitätsanalyse 1:</b> Gesamtraten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen; Woche 1 – 24 (siehe Tabelle 4-71)</li> <li>• <b>Hauptanalyse 2:</b> Inzidenzen von UE auf Ebene der SOC und PT; Woche 1 – 24 <ul style="list-style-type: none"> <li>○ jegliche UE (siehe Tabelle 4-72)</li> <li>○ SUE (siehe Tabelle 4-73)</li> <li>○ schwere UE (CTCAE <math>\geq</math> 3) (siehe Tabelle 4-74)</li> <li>○ Therapieabbrüche aufgrund von UE (siehe Tabelle 4-75)</li> </ul> </li> <li>• <b>Zusatzanalyse 1:</b> Zeit bis zum ersten UE; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-76 f.)</li> <li>• <b>Zusatzanalyse 2:</b> Zeit bis zum ersten UE auf Ebene der SOC und PT; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-78 f.)</li> <li>• <b>Zusatzanalyse 3:</b> <ul style="list-style-type: none"> <li>○ Gesamtraten UE von besonderem Interesse; Woche 1 – 24 (siehe Tabelle 4-82)</li> <li>○ Zeit bis zum ersten UE von besonderem Interesse; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-83)</li> </ul> </li> <li>• <b>Zusatzanalyse 4:</b> Zeit bis zur Progression in eine AML; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-84)</li> <li>• <b>Zusatzanalyse 5:</b> Progression in ein Hochrisiko-MDS; Woche eins bis zum Datenschnitt vom 01.07.2019 (siehe Tabelle 4-85)</li> </ul>
	<p>ARR = absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran-Mantel-Haenszel; CTCAE = <i>Common Terminology Criteria for Adverse Events</i>; KI = Konfidenzintervall; IPSS-R = <i>International Prognostic Scoring System-revised</i>; MDS = Myelodysplastische Syndrome; MedDRA = <i>Medical Dictionary for Regulatory Activities</i>; NCI = <i>National Cancer Institute</i>; OR = Odds Ratio; PT = bevorzugter Begriff (<i>preferred term</i>); RR = Relatives Risiko; SMQ = Standardized MedDRA Query; SOC = Systemorganklasse (<i>system organ class</i>); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis</p>

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-68: Bewertung des Verzerrungspotenzials für den Endpunkt Sicherheit in MEDALIST

Studie MEDALIST	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Jegliche UE	niedrig	ja	ja	ja	ja	niedrig
SUE	niedrig	ja	ja	ja	ja	niedrig
Schwere UE (CTCAE $\geq 3$ )	niedrig	ja	ja	ja	ja	niedrig
Therapieab- brüche aufgrund von UE	niedrig	ja	ja	ja	ja	niedrig
CTCAE = <i>Common Terminology Criteria for Adverse Events</i> ; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis						

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Sicherheit wurde verblindet erhoben. In die Analyse gingen nur solche Patienten ein, die mindestens eine Dosis Studienmedikation erhalten haben (Safety-Population). Patienten werden in dem Behandlungsarm berücksichtigt, dessen Behandlung sie tatsächlich erhalten haben. Safety- und ITT-Population stimmen vollständig überein (siehe Tabelle 4-6), sodass das ITT-Prinzip adäquat umgesetzt wurde. Anhaltspunkte für eine ergebnisabhängige Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

Insgesamt ergibt sich somit ein niedriges Verzerrungspotential für den Endpunkt Sicherheit gemessen als jegliche UE, SUE, schwere UE (CTCAE  $\geq 3$ ) und Therapieabbrüche aufgrund von UE und eine hohe Aussagesicherheit der Ergebnisse.

Stellen Sie die Ergebnisse für den Endpunkt Sicherheit für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

### Hauptanalyse 1: Gesamtraten UE

Tabelle 4-69: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Gesamtraten UE; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
Endpunkt	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)				
Jegliche UE	145/153 (94,8)	70/76 (92,1)	1,55 [0,52; 4,65]	1,03 [0,95; 1,11]	2,67 [-4,35; 9,68]	0,429
SUE	40/153 (26,1)	16/76 (21,1)	1,33 [0,69; 2,57]	1,24 [0,75; 2,07]	5,09 [-6,42; 16,60]	0,400
Schwere UE (CTCAE ≥ 3)	54/153 (35,3)	27/76 (35,5)	0,99 [0,56; 1,76]	0,99 [0,69; 1,44]	-0,23 [-13,39; 12,93]	0,972
Therapieabbrüche aufgrund von UE	12/153 (7,8)	4/76 (5,3)	1,53 [0,48; 4,92]	1,49 [0,50; 4,47]	2,58 [-4,00; 9,16]	0,472

Datenschnitt: 08.05.2018; Population: Safety-Population. Stratifizierung: IPSS-R-Risikogruppe und durchschnittlicher Transfusionslast zur Baseline.  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet.  
<sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.

ARR = Absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran Mantel Haenszel; CTCAE = *Common Terminology Criteria for Adverse Events*; IPSS-R = *International Prognostic Scoring System-revised*; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis  
Quelle: [18]

Im Studienzeitraum der primären Behandlungsphase (Woche 1 – 24) zeigte sich über alle betrachteten UE-Kategorien hinweg (jegliche UE, SUE, schwere UE und Therapieabbrüche aufgrund von UE) kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-69).

Bei 94,8 % (145/153) der Patienten im Luspatercept-Arm und 92,1 % (70/76) der Patienten im Placebo-Arm trat im betrachteten Studienzeitraum ein UE auf. Der Unterschied zwischen den Behandlungsarmen war statistisch nicht signifikant ( $p = 0,429$ ).

26,1 % (40/153) und 21,1 % (16/76) der Patienten unter Luspatercept und Placebo erlitten ein SUE. Dies entspricht einem RR im Vergleich von Luspatercept zu Placebo von 1,24 (95 %-KI: [0,75; 2,07]) sowie einer ARR von 5,1 %-Punkten (95 %-KI: [-6,42; 16,60]). Der Unterschied zwischen den Behandlungsarmen war statistisch nicht signifikant ( $p = 0,400$ ).

Schwere UE, definiert als UE mit einem Schweregrad nach CTCAE  $\geq 3$ , traten bei 35,3 % (54/153) der Patienten im Luspatercept-Arm und 35,5 % (27/76) der Patienten im Placebo-Arm auf. Dies entspricht einem RR im Vergleich von Luspatercept zu Placebo von 0,99 (95 %-KI: [0,69; 1,44]) sowie einer ARR von -0,2 %-Punkten (95 %-KI: [-13,39; 12,93]). Der Unterschied zwischen den Behandlungsarmen war statistisch nicht signifikant ( $p = 0,972$ ).

Therapieabbrüche wegen UE traten in beiden Armen selten auf (7,8 % (12/153) der Patienten im Luspatercept-Arm und 5,3 % (4/76) der Patienten im Placebo-Arm). Dies entspricht einem RR im Vergleich von Luspatercept zu Placebo von 1,49 (95 %-KI: [0,50; 4,47]) sowie einer ARR von 2,6 %-Punkten (95 %-KI: [-4,00; 9,16]). Der Unterschied zwischen den Behandlungsarmen war statistisch nicht signifikant ( $p = 0,472$ ).

Die Auswertungen über die Inzidenzen von UE über die Zeit weisen zudem auf eine langfristig gute Sicherheit von Luspatercept hin. So traten UE vorwiegend während der ersten Behandlungszyklen auf und nahmen im weiteren Verlauf zunehmend ab (siehe Tabelle 4-70).

Dementsprechend ist das Sicherheitsprofil von Luspatercept insgesamt vergleichbar mit einer Placebo-Behandlung.

Tabelle 4-70: Ergebnisse zum Endpunkt Sicherheit aus MEDALIST – Gesamtraten UE über die Zeit bis zum Datenschnitt vom 01.07.2019

Studie MEDA- LIST	Luspatercept + BSC <sup>(1)</sup>					Placebo + BSC <sup>(1)</sup>				
	Behandlungs- phase (Woche 1 – 24)		Extensionsphase (ab Woche 25)			Behandlungs- phase (Woche 1 – 24)		Extensionsphase (ab Woche 25)		
Zyklus	1 – 4	5 – 8	1 – 4	5 – 8	$\geq 9$	1 – 4	5 – 8	1 – 4	5 – 8	$\geq 9$
N	153	145	100	89	78	76	73	26	16	9
jegliche UE	134 (87,6)	109 (75,2)	71 (71,0)	60 (67,4)	71 (91,0)	63 (82,9)	54 (74,0)	16 (61,5)	11 (68,8)	9 (100)
SUE	23 (15,0)	21 (14,5)	1 (1,0)	3 (3,4)	30 (38,5)	7 (9,2)	9 (12,3)	4 (15,4)	3 (18,8)	1 (11,1)
schwere UE	39 (25,5)	35 (24,1)	10 (10,0)	5 (5,6)	40 (51,3)	16 (21,1)	13 (17,8)	8 (30,8)	2 (12,5)	3 (33,3)
Therapie abbrüche aufgrund von UE	5 (3,3)	3 (2,1)	0	0	5 (6,4)	0	3 (4,1)	2 (7,7)	0	0

Datenschnitt: 01.07.2019; Population: Safety-Population  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
 BSC = Best-Supportive-Care; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis  
 Quelle: [20]

**Sensitivitätsanalyse: Gesamtraten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen**

Tabelle 4-71: Sensitivitätsanalyse: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Gesamtraten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
Endpunkt	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)				
Jegliche UE	141/153 (92,2)	66/76 (86,8)	1,78 [0,73; 4,33]	1,06 [0,96; 1,17]	5,31 [-3,40; 14,03]	0,200
SUE	9/153 (5,9)	4/76 (5,3)	1,12 [0,34; 3,78]	1,12 [0,36; 3,51]	0,62 [-5,63; 6,87]	0,849
Schwere UE (CTCAE ≥ 3)	45/153 (29,4)	21/76 (27,6)	1,09 [0,59; 2,01]	1,06 [0,69; 1,65]	1,78 [-10,60; 14,16]	0,780
Datenschnitt: 08.05.2018; Population: Safety Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.  ARR = Absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran Mantel Haenszel; CTCAE = <i>Common Terminology Criteria for Adverse Events</i> ; IPSS-R = <i>International Prognostic Scoring System-revised</i> ; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis Quelle: [18]						

Auch bei Betrachtung der Ergebnisse zu UE ohne Berücksichtigung von Symptomen der Anämie sowie Ereignissen, die in einer stationären Aufnahme der Patienten resultierten, die bereits über den EORTC QLQ-C30-Fragebogen bzw. die Gesamtraten an Hospitalisierungen als Endpunkte der Morbidität erhoben wurden, zeigt sich für alle drei UE-Kategorien jegliche UE, SUE und schwere UE kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-71).

**Hauptanalyse 2: Inzidenzen von UE auf Ebene der SOC und PT****Jegliche UE**

Tabelle 4-72: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Inzidenzen von UE auf Ebene der SOC und PT (jegliche UE ab Inzidenzen von mindestens 10 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
SOC PT	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)				
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	92/153 (60,1)	36/76 (47,4)	1,68 [0,96; 2,92]	1,27 [0,97; 1,66]	12,76 [-0,88; 26,41]	0,068
PT: Asthenie	28/153 (18,3)	6/76 (7,9)	2,61 [1,03; 6,62]	2,32 [1,00; 5,36]	10,41 [1,79; 19,03]	0,037
PT: Ermüdung	31/153 (20,3)	9/76 (11,8)	1,89 [0,85; 4,21]	1,71 [0,86; 3,41]	8,42 [-1,24; 18,08]	0,115
PT: Fieber	10/153 (6,5)	5/76 (6,6)	0,99 [0,33; 3,01]	0,99 [0,35; 2,80]	-0,04 [-6,86; 6,77]	0,990
PT: Ödem peripher	18/153 (11,8)	11/76 (14,5)	0,79 [0,35; 1,76]	0,81 [0,40; 1,63]	-2,71 [-12,12; 6,71]	0,562
SOC: Herzkrankungen	32/153 (20,9)	11/76 (14,5)	1,56 [0,74; 3,30]	1,45 [0,77; 2,71]	6,44 [-3,76; 16,64]	0,241
SOC: Erkrankun- gen der Haut und des Unterhaut- gewebes	22/153 (14,4)	15/76 (19,7)	0,68 [0,33; 1,41]	0,73 [0,40; 1,32]	-5,36 [-15,89; 5,18]	0,301
SOC: Gefäß- erkrankungen	28/153 (18,3)	9/76 (11,8)	1,67 [0,74; 3,74]	1,55 [0,77; 3,11]	6,46 [-3,04; 15,96]	0,212
PT: Hypertonie	10/153 (6,5)	5/76 (6,6)	0,99 [0,33; 3,01]	0,99 [0,35; 2,80]	-0,04 [-6,86; 6,77]	0,990
SOC: Unter- suchungen	26/153 (17,0)	10/76 (13,2)	1,35 [0,61; 2,97]	1,29 [0,66; 2,54]	3,84 [-5,82; 13,49]	0,454
SOC: Erkrankun- gen des Blutes und des Lymphsystems	22/153 (14,4)	12/76 (15,8)	0,90 [0,42; 1,92]	0,91 [0,48; 1,74]	-1,41 [-11,32; 8,50]	0,778
PT: Anämie	10/153 (6,5)	5/76 (6,6)	0,99 [0,33; 3,01]	0,99 [0,35; 2,80]	-0,04 [-6,86; 6,77]	0,990
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplika- tionen	19/153 (12,4)	12/76 (15,8)	0,76 [0,35; 1,65]	0,79 [0,40; 1,53]	-3,37 [-13,09; 6,35]	0,484
PT: Sturz	11/153 (7,2)	5/76 (6,6)	1,10 [0,37; 3,29]	1,09 [0,39; 3,03]	0,61 [-6,30; 7,53]	0,865

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			SOC PT	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>
SOC: Gastro- intestinale Erkrankungen	78/153 (51,0)	25/76 (32,9)	2,12 [1,20; 3,77]	1,55 [1,08; 2,21]	18,09 [4,88; 31,29]	0,010
PT: Diarrhö	24/153 (15,7)	6/76 (7,9)	2,17 [0,85; 5,56]	1,99 [0,85; 4,65]	7,79 [-0,57; 16,16]	0,101
PT: Obstipation	14/153 (9,2)	6/76 (7,9)	1,18 [0,43; 3,19]	1,16 [0,46; 2,90]	1,26 [-6,34; 8,85]	0,752
PT: Übelkeit	23/153 (15,0)	6/76 (7,9)	2,06 [0,80; 5,31]	1,90 [0,81; 4,48]	7,14 [-1,16; 15,43]	0,127
SOC: Psychia- trische Erkrankun- gen	21/153 (13,7)	7/76 (9,2)	1,57 [0,64; 3,87]	1,49 [0,66; 3,35]	4,51 [-3,97; 13,00]	0,327
SOC: Erkrankun- gen der Nieren und Harnwege	19/153 (12,4)	7/76 (9,2)	1,40 [0,56; 3,49]	1,35 [0,59; 3,07]	3,21 [-5,13; 11,55]	0,472
SOC: Infektionen und parasitäre Erkrankungen	57/153 (37,3)	26/76 (34,2)	1,14 [0,64; 2,03]	1,09 [0,75; 1,58]	3,04 [-10,09; 16,18]	0,652
PT: Infektion der oberen Atem- wege	10/153 (6,5)	2/76 (2,6)	2,59 [0,55; 12,12]	2,48 [0,56; 11,05]	3,90 [-1,41; 9,22]	0,213
SOC: Augen- erkrankungen	12/153 (7,8)	1/76 (1,3)	6,38 [0,81; 50,04]	5,96 [0,79; 45,00]	6,53 [1,56; 11,50]	0,045
SOC: Skelettmus- kulatur-, Bindegewebes- und Knochen- erkrankungen	52/153 (34,0)	27/76 (35,5)	0,93 [0,52; 1,66]	0,96 [0,66; 1,39]	-1,54 [-14,66; 11,58]	0,818
PT: Myalgie	10/153 (6,5)	5/76 (6,6)	0,99 [0,33; 3,01]	0,99 [0,35; 2,80]	-0,04 [-6,86; 6,77]	0,990
PT: Rücken- schmerzen	21/153 (13,7)	5/76 (6,6)	2,26 [0,82; 6,25]	2,09 [0,82; 5,32]	7,15 [-0,65; 14,94]	0,109
SOC: Erkrankun- gen des Nerven- systems	60/153 (39,2)	17/76 (22,4)	2,24 [1,19; 4,20]	1,75 [1,10; 2,79]	16,85 [4,70; 29,00]	0,011
PT: Kopf- schmerz	21/153 (13,7)	5/76 (6,6)	2,26 [0,82; 6,25]	2,09 [0,82; 5,32]	7,15 [-0,65; 14,94]	0,109
PT: Schwindel- gefühl	23/153 (15,0)	4/76 (5,3)	3,18 [1,06; 9,57]	2,86 [1,02; 7,96]	9,77 [2,20; 17,34]	0,031

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
SOC PT	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
SOC: Erkrankun- gen der Atem- wege, des Brust- raums und Mediastinums	50/153 (32,7)	21/76 (27,6)	1,27 [0,69; 2,33]	1,18 [0,77; 1,82]	5,05 [-7,45; 17,55]	0,438
PT: Dyspnoe	18/153 (11,8)	4/76 (5,3)	2,40 [0,78; 7,36]	2,24 [0,78; 6,37]	6,50 [-0,66; 13,66]	0,117
PT: Husten	20/153 (13,1)	8/76 (10,5)	1,28 [0,54; 3,05]	1,24 [0,57; 2,69]	2,55 [-6,18; 11,27]	0,581
SOC: Stoff- wechsel- und Ernährungs- störungen	42/153 (27,5)	16/76 (21,1)	1,42 [0,74; 2,73]	1,30 [0,79; 2,16]	6,40 [-5,18; 17,97]	0,296

Datenschnitt: 08.05.2018; Population: Safety-Population.  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> OR, RR und ARR wurden mittels Normal-Approximation berechnet.  
<sup>(3)</sup> Der p-Wert basiert auf einem CMH-Test.

ARR = Absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran Mantel Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; PT = bevorzugter Begriff (*preferred term*); RR = Relatives Risiko; SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis  
Quelle: [18]

Im Studienzeitraum der primären Behandlungsphase (Woche 1 – 24) zeigte sich für den Großteil der aufgetretenen UE auf Ebene der SOC und PT kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-72).

Für die SOC „Gastrointestinale Erkrankungen“ zeigte sich ein statistisch signifikanter Unterschied zuungunsten von Luspatercept (RR = 1,55; 95 %-KI: [1,08; 2,21]). Der Unterschied besteht ausschließlich auf Ebene der SOC; statistisch signifikante Unterschiede für spezifische UE auf Ebene der PT, die der SOC „Gastrointestinale Erkrankungen“ zugerechnet werden, konnten nicht beobachtet werden. Darüber hinaus wurden Symptome gastrointestinaler Nebenwirkungen mittels der Symptomskalen des EORTC QLQ-C30-Fragebogens erhoben. Hier zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen, sodass davon auszugehen ist, dass der Unterschied auf Ebene der SOC sich nicht in einer relevanten Verschlechterung der Symptomatik widerspiegelt (siehe Tabelle 4-41).

Die SOC „Erkrankungen des Nervensystems“ trat unter Luspatercept mit 39,2 % (60/153) der Patienten statistisch signifikant häufiger auf als unter Placebo mit 22,4 % (17/76) der Patienten. Dies entspricht einem RR im Vergleich von Luspatercept zu Placebo von 1,75 (95 %-KI: [1,10; 2,79]) sowie einer ARR von 16,9 %-Punkten (95 %-KI: [4,70; 29,00]). Dieser Unterschied spiegelt sich in dem PT „Schwindelgefühl“ wider, für den ebenfalls ein statistisch signifikantes

Ergebnis zuungunsten von Luspatercept vorlag (RR = 2,86; 95 %-KI: [1,02; 7,96]). Für alle anderen PT, die unter der SOC Erkrankungen des Nervensystems zusammengefasst werden, lagen keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen vor. Schwindelgefühl stellt als Symptom der Anämie ein erkrankungsbezogenes Ereignis dar, welches bereits über den EORTC QLQ-C30-Fragebogen erhoben wurde (siehe Tabelle 4-41). Hier zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen, sodass der hier festgestellte Unterschied auf Ebene des PT sich nicht in einer klinisch relevanten Verschlechterung der Symptomatik widerspiegelt. Darüber hinaus liegt die untere Grenze des 95 %-KI dieses UE nur knapp oberhalb der Signifikanzgrenze von 1. Dementsprechend ist das Ausmaß dieses Unterschieds als geringfügig einzustufen.

Neben „Schwindelgefühl“ besteht zudem ein statistisch signifikanter Unterschied zuungunsten von Luspatercept in Bezug auf das PT „Asthenie“ (RR = 2,32; 95 %-KI: [1,00; 5,36]). Auch in diesem Fall liegt die untere Grenze des 95 %-KI nur knapp oberhalb der Signifikanzschwelle von 1, so dass erneut von einem nicht klinisch relevanten Unterschied zwischen den Behandlungsarmen auszugehen ist.

Darüber hinaus zeigte sich ein weiteres statistisch signifikantes Ergebnis zuungunsten von Luspatercept hinsichtlich der SOC „Augenerkrankungen“ (RR = 5,96; 95 %-KI: [0,79; 45,00];  $p = 0,045$ ). Da die untere Grenze des 95 %-KI unterhalb der Signifikanzschwelle von 1 liegt, ist auch für diese SOC von keinem klinisch relevanten Unterschied zwischen den beiden Behandlungsarmen auszugehen.

Weitere statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen lagen nicht vor.

Insgesamt ist somit von einer vergleichbaren Sicherheit bezüglich der UE zwischen den beiden Behandlungsarmen auszugehen.

**Schwerwiegende UE**

Tabelle 4-73: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Inzidenzen von SUE auf Ebene der SOC und PT (SUE ab Inzidenzen von mindestens 5 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
SOC PT	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
SOC: Infektionen und parasitäre Erkrankungen	14/153 (9,2)	8/76 (10,5)	0,86 [0,34; 2,14]	0,87 [0,38; 1,98]	-1,38 [-9,65; 6,90]	0,740
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	8/153 (5,2)	3/76 (3,9)	1,34 [0,35; 5,21]	1,32 [0,36; 4,85]	1,28 [-4,34; 6,90]	0,670
Datenschnitt: 08.05.2018; Population: Safety Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR, RR und ARR wurden mittels Normal-Approximation berechnet. <sup>(3)</sup> Der p-Wert basiert auf einem CMH-Test.  ARR = Absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran Mantel Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; PT = bevorzugter Begriff ( <i>preferred term</i> ); RR = Relatives Risiko; SOC = Systemorganklasse ( <i>system organ class</i> ); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis Quelle: [18]						

Im Studienzeitraum der primären Behandlungsphase (Woche 1 – 24) traten überhaupt nur in den beiden SOC „Infektionen und parasitäre Erkrankungen“ und „Verletzungen, Vergiftung und durch Eingriffe bedingte Komplikationen“ UE mit einer Inzidenzrate von mindestens 5 % auf. Es zeigte sich analog zu den Ergebnissen der Gesamtraten für SUE auch auf Ebene der SOC und PT keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen (siehe Tabelle 4-73).

**Schwere UE (CTCAE  $\geq 3$ )**

Tabelle 4-74: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Inzidenzen von schweren UE (CTCAE  $\geq 3$ ) auf Ebene der SOC und PT (Schwere UE (CTCAE  $\geq 3$ ) ab Inzidenzen von mindestens 5 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
SOC PT	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
SOC: Erkrankungen des Blutes und des Lymphsystems	13/153 (8,5)	8/76 (10,5)	0,79 [0,31; 1,99]	0,81 [0,35; 1,86]	-2,03 [-10,22; 6,16]	0,617
PT: Anämie	9/153 (5,9)	4/76 (5,3)	1,12 [0,34; 3,78]	1,12 [0,36; 3,51]	0,62 [-5,63; 6,87]	0,849
PT: Neutropenie	5/153 (3,3)	5/76 (6,6)	0,48 [0,13; 1,71]	0,50 [0,15; 1,66]	-3,31 [-9,56; 2,93]	0,249
SOC: Infektionen und parasitäre Erkrankungen	14/153 (9,2)	7/76 (9,2)	0,99 [0,38; 2,57]	0,99 [0,42; 2,36]	-0,06 [-8,01; 7,89]	0,988
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	12/153 (7,8)	3/76 (3,9)	2,07 [0,57; 7,57]	1,99 [0,58; 6,83]	3,90 [-2,21; 10,00]	0,263
SOC: Stoffwechsel- und Ernährungsstörungen	12/153 (7,8)	3/76 (3,9)	2,07 [0,57; 7,57]	1,99 [0,58; 6,83]	3,90 [-2,21; 10,00]	0,263
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	8/153 (5,2)	3/76 (3,9)	1,34 [0,35; 5,21]	1,32 [0,36; 4,85]	1,28 [-4,34; 6,90]	0,670
SOC: Erkrankungen des Nervensystems	8/153 (5,2)	0/76 (0,0) <sup>(4)</sup>	8,94 [0,51; 156,94]	8,50 [0,50; 145,34]	5,23 [1,70; 8,76]	0,043
Datenschnitt: 08.05.2018; Population: Safety Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR, RR und ARR wurden mittels Normal-Approximation berechnet. <sup>(3)</sup> Der p-Wert basiert auf einem CMH-Test. <sup>(4)</sup> Nullzellenkorrektur zur Berechnung der Effektschätzer wurde durchgeführt.  ARR = Absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran Mantel Haenszel; CTCAE = <i>Common Terminology Criteria for Adverse Events</i> ; KI = Konfidenzintervall; OR = Odds Ratio; PT = bevorzugter Begriff ( <i>preferred term</i> ); RR = Relatives Risiko; SOC = Systemorganklasse ( <i>system organ class</i> ); UE = unerwünschtes Ereignis Quelle: [18]						

Im Studienzeitraum der primären Behandlungsphase (Woche 1 – 24) zeigte sich analog zu den Ergebnissen der Gesamtrate für schwere UE (CTCAE  $\geq 3$ ) auch auf Ebene der SOC und PT kein relevanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-74).

Lediglich für die SOC „Erkrankungen des Nervensystems“ lag ein statistisch signifikanter Unterschied zuungunsten von Luspatercept vor (RR = 8,50; 95 %-KI: [0,50; 145,34]; p = 0,043), der in seinem Ausmaß jedoch als geringfügig einzustufen ist. Insgesamt erlitten nur wenig Patienten ein Ereignis dieser SOC (5,2 % (8/153) der Patienten im Luspatercept-Arm und kein Patient im Placebo-Arm).

Auf Ebene der PT traten nur die Ereignisse „Anämie“ und „Neutropenie“ mit Ereignishäufigkeiten > 5 % auf. Alle anderen erhobenen PT traten nur sehr selten auf und erreichten nur nach Zusammenfassung zu SOC Ereignishäufigkeiten von > 5 %.

### **Therapieabbrüche aufgrund von UE**

Tabelle 4-75: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Inzidenzen von Therapieabbrüchen aufgrund von UE auf Ebene der SOC und PT; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>
SOC PT	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)
SOC: Erkrankungen des Blutes und des Lymphsystems	0/153 (0,0)	1/76 (1,3)
PT: Thrombocytose	0/153 (0,0)	1/76 (1,3)
SOC: Untersuchungen	1/153 (0,7)	0/76 (0,0)
PT: Alaninaminotransferase erhöht	1/153 (0,7)	0/76 (0,0)
PT: Aspartataminotransferase erhöht	1/153 (0,7)	0/76 (0,0)
SOC: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	1/153 (0,7)	0/76 (0,0)
PT: Myalgie	1/153 (0,7)	0/76 (0,0)
SOC: Erkrankungen der Atemwege, des Brustraums und Mediastinums	1/153 (0,7)	0/76 (0,0)
PT: Lungenfibrose	1/153 (0,7)	0/76 (0,0)
SOC: Gefäßerkrankungen	1/153 (0,7)	0/76 (0,0)
PT: Hämorrhagischer Schock	1/153 (0,7)	0/76 (0,0)

SOC: Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen)	2/153 (1,3)	2/76 (2,6)
PT: Myelodysplastische Syndrome	0/153 (0,0)	1/76 (1,3)
PT: Transformation zu einer akuten myeloischen Leukämie	2/153 (1,3)	1/76 (1,3)
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	2/153 (1,3)	1/76 (1,3)
PT: Ermüdung	2/153 (1,3)	0/76 (0,0)
PT: Generelle Verschlechterung des physischen Gesundheitszustandes	0/153 (0,0)	1/76 (1,3)
SOC: Erkrankungen des Nervensystems	3/153 (2,0)	0/76 (0,0)
PT: Erinnerungsvermögen eingeschränkt	1/153 (0,7)	0/76 (0,0)
PT: Kopfschmerz	1/153 (0,7)	0/76 (0,0)
PT: Morbus Parkinson	1/153 (0,7)	0/76 (0,0)
SOC: Infektionen und parasitäre Erkrankungen	2/153 (1,3)	0/76 (0,0)
PT: Sepsis	2/153 (1,3)	0/76 (0,0)
SOC: Erkrankungen der Nieren und Harnwege	2/153 (1,3)	0/76 (0,0)
PT: chronische Nierenerkrankung	1/153 (0,7)	0/76 (0,0)
PT: Nierenversagen	1/153 (0,7)	0/76 (0,0)
Datenschnitt: 08.05.2018; Population: Safety Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  BSC = Best-Supportive-Care; PT = bevorzugter Begriff ( <i>preferred term</i> ); SOC = Systemorganklasse ( <i>system organ class</i> ); UE = unerwünschtes Ereignis Quelle: [18]		

Im Studienzeitraum der primären Behandlungsphase (Woche 1 – 24) traten sowohl unter Luspatercept als auch unter Placebo insgesamt sehr wenig Therapieabbrüche auf. Die Gründe für den Therapieabbruch waren verschieden; es kann kein einzelnes PT oder keine einzelne SOC identifizieren werden, die häufiger als andere Grund für einen Therapieabbruch waren (siehe Tabelle 4-75). Zudem bestehen in keinem der PT und in keiner der SOC wesentliche Unterschiede zwischen den Behandlungsarmen.

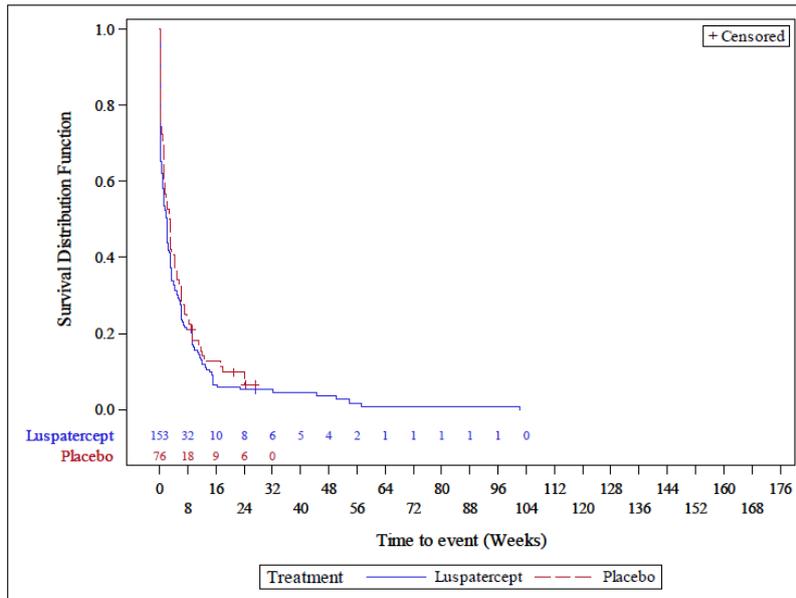
**Zusatzanalyse 1: Zeit bis zum ersten UE**

Tabelle 4-76: Zusatzanalyse 1: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten UE; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Jegliche UE	151/153 (98,7)	2,0 [1,0; 2,6]	70/76 (92,1)	2,9 [1,3; 4,4]	1,17 [0,88; 1,56]	0,292
SUE	64/153 (41,8)	115,7 [101,4.; 126,9]	23/76 (30,3)	N.E. [34,9; N.E.]	0,95 [0,58; 1,58]	0,858
Schwere UE (CTCAE ≥ 3)	84/153 (54,9)	64,4 [36,3; 112,3]	34/76 (44,7)	36,0 [26,0; N.E.]	0,85 [0,56; 1,28]	0,433
Therapieabbrüche aufgrund von UE	21/153 (13,7)	N.E. [N.E.; N.E.]	6/76 (7,9)	N.E. [N.E.; N.E.]	1,02 [0,39; 2,66]	0,972

Datenschnitt: 01.07.2019; Population: Safety Population.  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.  
<sup>(3)</sup> HR basiert auf einem Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.  
<sup>(4)</sup> Der p-Wert des HR basiert auf einem Log-Rank-Test.

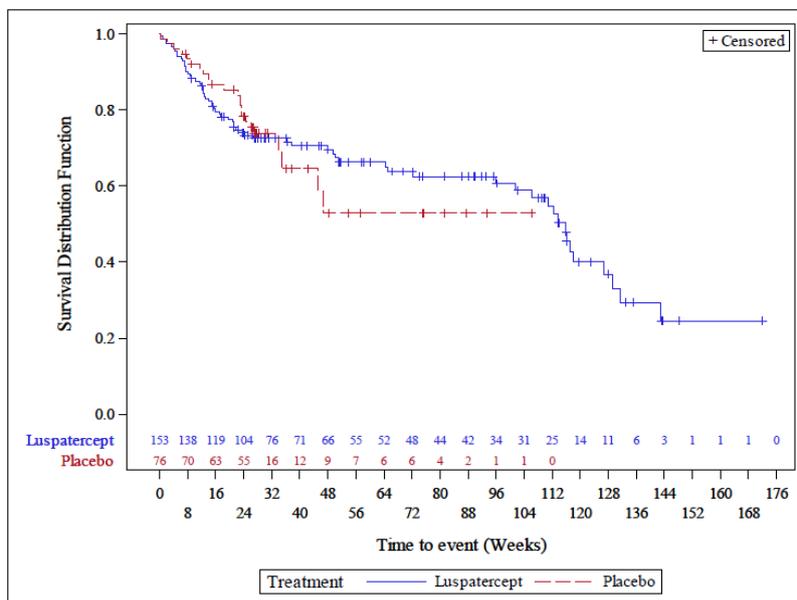
BSC = Best-Supportive-Care; CTCAE = *Common Terminology Criteria for Adverse Events*; HR = Hazard-Ratio; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis  
Quelle: [20]



Datenschnitt: 01.07.2019

Quelle: [20]

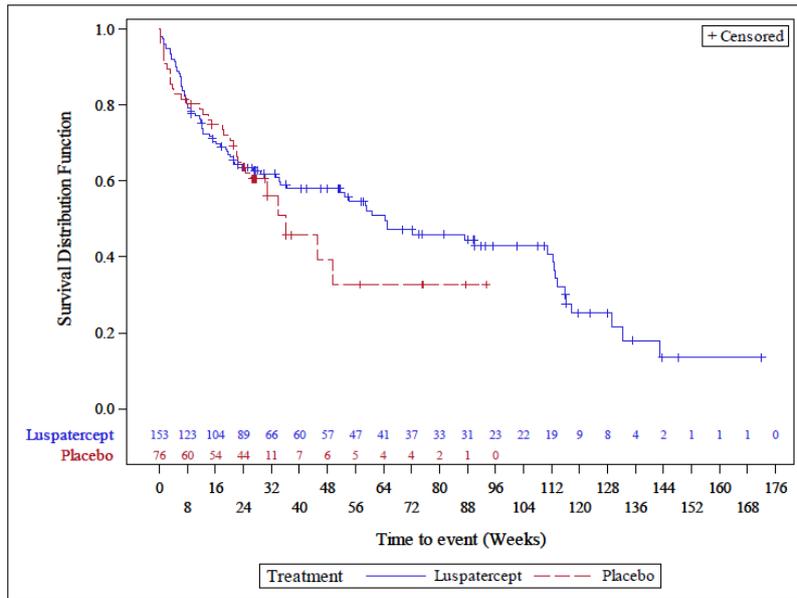
Abbildung 4-56: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten UE in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019

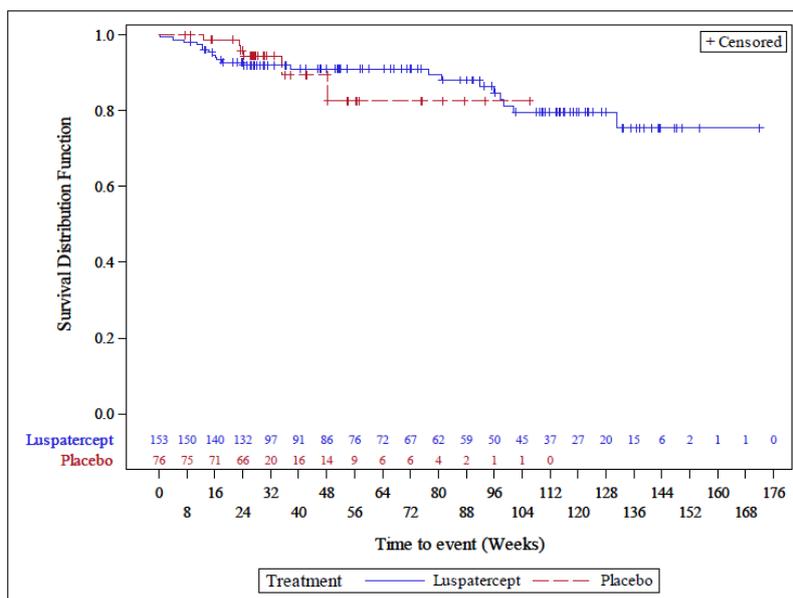
Quelle: [20]

Abbildung 4-57: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten SUE in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019  
 Quelle: [20]

Abbildung 4-58: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019  
 Quelle: [20]

Abbildung 4-59: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

Auswertungen über den Studienzeitraum bis zum Datenschnitt vom 01.07.2019 werden aufgrund der Unsicherheit hinsichtlich der Interpretierbarkeit infolge unterschiedlich langer Beobachtungszeiten nach Woche 24 ergänzend als Ereigniszeitanalysen dargestellt.

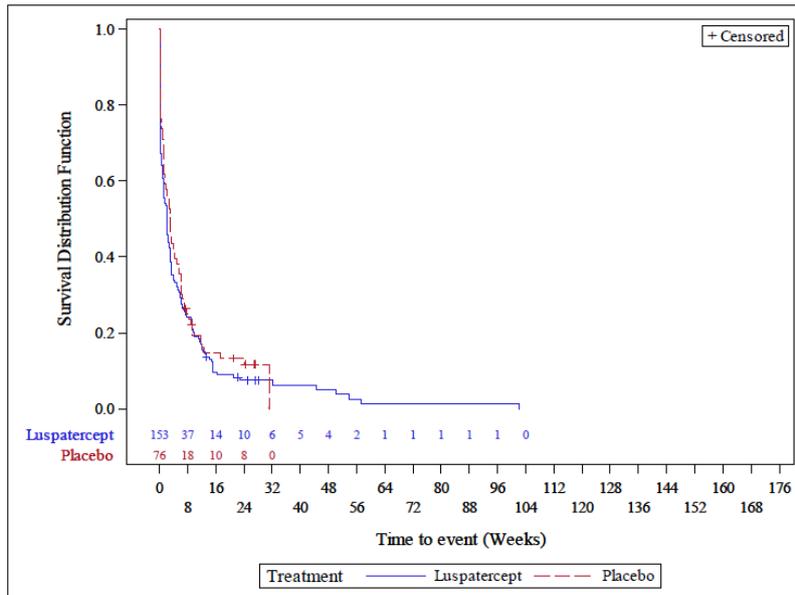
Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit mindestens einem UE in den ersten 24 Wochen der Behandlungsphase, zeigte sich auch in der Analyse der Zeit bis zum ersten UE für alle betrachteten Kategorien, d. h. jegliche UE, SUE, schwere UE und Therapieabbrüche aufgrund von UE, kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-76 & Abbildung 4-56 ff.).

Tabelle 4-77: Zusatzanalyse 1: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Jegliche UE	147/153 (96,1)	2,1 [1,1; 3,0]	67/76 (88,2)	3,1 [1,3; 5,0]	1,15 [0,86; 1,53]	0,371
SUE	19/153 (12,4)	N.E. [N.E.; N.E.]	6/76 (7,9)	N.E. [56,6; N.E.]	0,82 [0,30; 2,19]	0,685
Schwere UE (CTCAE ≥ 3)	69/153 (45,1)	90,1 [58,7; 115,7]	27/76 (35,5)	36,0 [33,9; N.E.]	0,88 [0,56; 1,40]	0,603

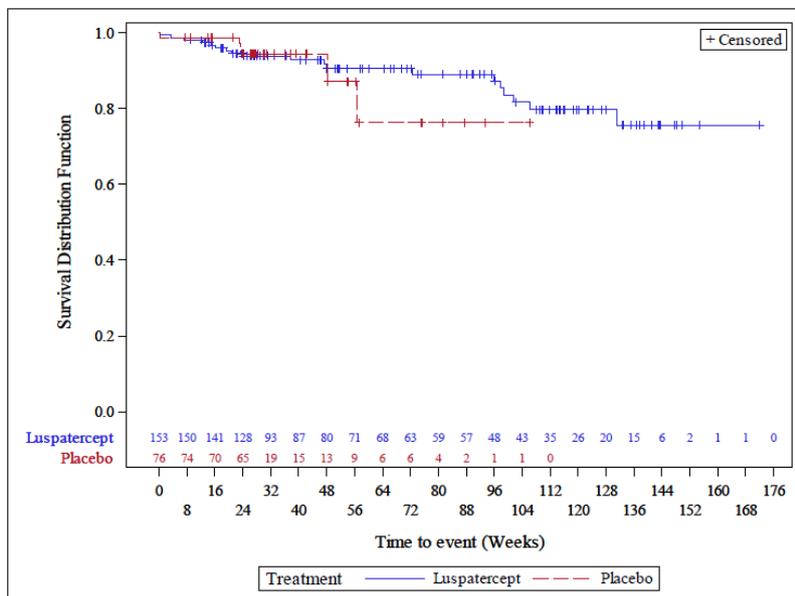
Datenschnitt: 01.07.2019; Population: Safety Population.  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.  
<sup>(3)</sup> HR basiert auf einem Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.  
<sup>(4)</sup> Der p-Wert des HR basiert auf einem Log-Rank-Test.

BSC = Best-Supportive-Care; CTCAE = *Common Terminology Criteria for Adverse Events*; HR = Hazard-Ratio; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis  
Quelle: [20]



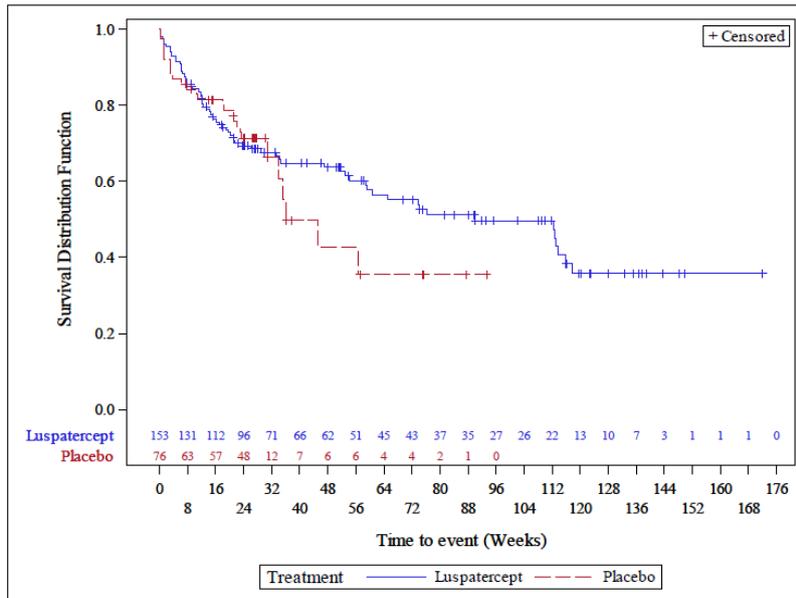
Datenschnitt: 01.07.2019; UE = unerwünschtes Ereignis  
Quelle: [20]

Abbildung 4-60: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten UE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis  
Quelle: [20]

Abbildung 4-61: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten SUE ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



Datenschnitt: 01.07.2019; CTCAE = *Common Terminology Criteria for Adverse Events*; UE = unerwünschtes Ereignis  
Quelle: [20]

Abbildung 4-62: Zusatzanalyse 1: Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) ohne Berücksichtigung von Symptomen der Anämie und Hospitalisierungen in MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

Auch bei Betrachtung der Ergebnisse zu UE ohne Berücksichtigung von Symptomen der Anämie sowie Ereignissen, die in einer stationären Aufnahme der Patienten resultierten, die bereits über den EORTC QLQ-C30-Fragebogen bzw. die Gesamtraten an Hospitalisierungen als Endpunkte der Morbidität erhoben wurden, zeigt sich für alle drei UE-Kategorien jegliche UE, SUE und schwere UE kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-77 & Abbildung 4-60 ff.).

**Zusatzanalyse 2: Zeit bis zum ersten UE auf Ebene der SOC und PT****Jegliche UE**

Tabelle 4-78: Zusatzanalyse 2: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten UE auf Ebene der SOC und PT (jegliche UE ab Inzidenzen von mindestens 10 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST SOC PT	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	123/153 (80,4)	12,1 [9,1; 15,1]	39/76 (51,3)	27,1 [18,1; N.E.]	1,57 [1,08; 2,27]	0,016
PT: Asthenie	39/153 (25,5)	N.E. [122,3; N.E.]	9/76 (11,8)	N.E. [56,1; N.E.]	1,70 [0,80; 3,58]	0,160
PT: Ermüdung	46/153 (30,1)	N.E. [128,1; N.E.]	11/76 (14,5)	N.E. [56,7; N.E.]	1,67 [0,85; 3,27]	0,133
PT: Fieber	20/153 (13,1)	N.E. [N.E.; N.E.]	7/76 (9,2)	N.E. [N.E.; N.E.]	0,76 [0,30; 1,93]	0,567
PT: Ödem peripher	37/153 (24,2)	N.E. [123,7; N.E.]	13/76 (17,1)	N.E. [N.E.; N.E.]	0,89 [0,46; 1,71]	0,720
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	46/153 (30,1)	125,3 [100,7; N.E.]	18/76 (23,7)	N.E. [41,0; N.E.]	0,67 [0,37; 1,21]	0,180
PT: Sturz	26/153 (17,0)	N.E. [138,4; N.E.]	10/76 (13,2)	N.E. [54,4; N.E.]	0,66 [0,30; 1,45]	0,297
SOC: Erkrankungen der Haut und des Unterhautgewebes	44/153 (28,8)	135,0 [111,0; N.E.]	16/76 (21,1)	N.E. [55,1; N.E.]	0,78 [0,43; 1,44]	0,431
SOC: Herzerkrankungen	47/153 (30,7)	N.E. [121,1; N.E.]	11/76 (14,5)	N.E. [N.E.; N.E.]	1,79 [0,92; 3,48]	0,085
PT: Palpitationen	11/153 (7,2)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [N.E.; N.E.]	1,03 [0,36; 2,99]	0,956
SOC: Untersuchungen	41/153 (26,8)	N.E. [132,1; N.E.]	12/76 (15,8)	N.E. [52,1; N.E.]	1,19 [0,61; 2,31]	0,616
PT: Aspartatamino-transferase erhöht	10/153 (6,5)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	3,45 [0,43; 27,82]	0,216
SOC: Gefäßerkrankungen	40/153 (26,1)	N.E. [N.E.; N.E.]	11/76 (14,5)	N.E. [N.E.; N.E.]	1,38 [0,69; 2,73]	0,353
PT: Hypertonie	16/153 (10,5)	N.E. [N.E.; N.E.]	6/76 (7,9)	N.E. [N.E.; N.E.]	0,92 [0,35; 2,42]	0,861

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
SOC: Erkrankungen des Blutes und des Lymphsystems	36/153 (23,5)	N.E. [123,7; N.E.]	14/76 (18,4)	N.E. [N.E.; N.E.]	0,84 [0,44; 1,61]	0,608
PT: Anämie	17/153 (11,1)	N.E. [N.E.; N.E.]	6/76 (7,9)	N.E. [N.E.; N.E.]	0,90 [0,34; 2,40]	0,832
SOC: Gastrointestinale Erkrankungen	101/153 (66,0)	21,6 [14,4; 41,6]	27/76 (35,5)	N.E. [30,7; N.E.]	1,76 [1,14; 2,70]	0,010
PT: Abdominalschmerz	10/153 (6,5)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	0,72 [0,21; 2,46]	0,594
PT: Diarrhö	43/153 (28,1)	N.E. [109,3; N.E.]	8/76 (10,5)	N.E. [59,4; N.E.]	1,74 [0,80; 3,81]	0,159
PT: Erbrechen	14/153 (9,2)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [76,3; N.E.]	0,90 [0,30; 2,65]	0,846
PT: Obstipation	21/153 (13,7)	N.E. [N.E.; N.E.]	7/76 (9,2)	N.E. [N.E.; N.E.]	1,11 [0,46; 2,69]	0,816
PT: Schmerzen Oberbauch	11/153 (7,2)	N.E. [N.E.; N.E.]	2/76 (2,6)	N.E. [N.E.; N.E.]	1,92 [0,41; 8,99]	0,400
PT: Übelkeit	34/153 (22,2)	N.E. [N.E.; N.E.]	6/76 (7,9)	N.E. [N.E.; N.E.]	2,41 [1,00; 5,80]	0,042
SOC: Psychiatrische Erkrankungen	37/153 (24,2)	N.E. [131,4; N.E.]	10/76 (13,2)	N.E. [N.E.; N.E.]	1,32 [0,64; 2,70]	0,453
PT: Schlaflosigkeit	12/153 (7,8)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	0,76 [0,23; 2,55]	0,658
PT: Verwirrheitszustand	10/153 (6,5)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
SOC: Erkrankungen der Nieren und Harnwege	34/153 (22,2)	N.E. [N.E.; N.E.]	11/76 (14,5)	85,7 [54,3; 85,7]	1,06 [0,53; 2,15]	0,864
SOC: Infektionen und parasitäre Erkrankungen	90/153 (58,8)	39,6 [28,9; 55,3]	31/76 (40,8)	49,4 [35,6; N.E.]	1,10 [0,73; 1,67]	0,644
PT: Bronchitis	19/153 (12,4)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	4,54 [0,59; 34,73]	0,112
PT: Grippe	11/153 (7,2)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
PT: Harnwegsinfektion	21/153 (13,7)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	1,52 [0,50; 4,59]	0,453
PT: Infektion der oberen Atemwege	19/153 (12,4)	N.E. [149,4; N.E.]	4/76 (5,3)	N.E. [54,4; N.E.]	1,28 [0,41; 3,97]	0,666

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
PT: virale Infektion der oberen Atemwege	17/153 (11,1)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	1,47 [0,48; 4,50]	0,501
SOC: Augenerkrankungen	22/153 (14,4)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [67,0; N.E.]	1,40 [0,51; 3,80]	0,508
SOC: Erkrankungen des Ohrs und des Labyrinths	20/153 (13,1)	N.E. [N.E.; N.E.]	3/76 (3,9)	N.E. [N.E.; N.E.]	1,88 [0,54; 6,58]	0,313
PT: Vertigo	12/153 (7,8)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
SOC: Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen)	18/153 (11,8)	N.E. [140,6; N.E.]	4/76 (5,3)	88,1 [88,1; N.E.]	1,04 [0,32; 3,35]	0,946
SOC: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	78/153 (51,0)	68,4 [32,1; 87,4]	29/76 (38,2)	51,7 [35,1; N.E.]	1,10 [0,71; 1,70]	0,670
PT: Arthralgie	12/153 (7,8)	N.E. [N.E.; N.E.]	9/76 (11,8)	N.E. [N.E.; N.E.]	0,32 [0,12; 0,82]	0,014
PT: Myalgie	13/153 (8,5)	N.E. [144,0; N.E.]	5/76 (6,6)	N.E. [N.E.; N.E.]	1,12 [0,39; 3,18]	0,838
PT: Rückenschmerzen	33/153 (21,6)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [N.E.; N.E.]	2,63 [1,01; 6,83]	0,039
SOC: Leber- und Gallenerkrankungen	13/153 (8,5)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [85,7; N.E.]	0,64 [0,21; 1,96]	0,427
SOC: Erkrankungen der Geschlechtsorgane und der Brustdrüse	14/153 (9,2)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
SOC: Erkrankungen des Nervensystems	81/153 (52,9)	56,6 [33,1; 106,7]	21/76 (27,6)	N.E. [47,0; N.E.]	1,67 [1,02; 2,74]	0,038
PT: Kopfschmerz	27/153 (17,6)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [N.E.; N.E.]	2,29 [0,87; 6,03]	0,084
PT: Schwindelgefühl	34/153 (22,2)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	3,44 [1,20; 9,80]	0,014
PT: Synkope	10/153 (6,5)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [55,1; N.E.]	2,19 [0,26; 18,47]	0,462
SOC: Erkrankungen der Atemwege, des Brustraums und Mediastinums	72/153 (47,1)	80,9 [53,6; 115,7]	28/76 (36,8)	41,7 [30,9; N.E.]	0,95 [0,60; 1,50]	0,831

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
PT: Dyspnoe	29/153 (19,0)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [N.E.; N.E.]	2,14 [0,81; 5,65]	0,116
PT: Epistaxis	13/153 (8,5)	150,6 [150,6; N.E.]	3/76 (3,9)	N.E. [N.E.; N.E.]	0,71 [0,18; 2,80]	0,621
PT: Husten	34/153 (22,2)	N.E. [117,3; N.E.]	10/76 (13,2)	N.E. [N.E.; N.E.]	1,09 [0,52; 2,29]	0,809
SOC: Stoffwechsel- und Ernährungsstörungen	72/153 (47,1)	101,4 [60,9; 116,1]	21/76 (27,6)	84,9 [51,0; 94,3]	1,12 [0,67; 1,86]	0,662
PT: Appetit vermindert	14/153 (9,2)	N.E. [N.E.; N.E.]	3/76 (3,9)	N.E. [N.E.; N.E.]	1,48 [0,41; 5,39]	0,546
PT: Hyperglykämie	10/153 (6,5)	N.E. [N.E.; N.E.]	3/76 (3,9)	N.E. [N.E.; N.E.]	1,13 [0,30; 4,25]	0,857
PT: Hyperurikämie	10/153 (6,5)	N.E. [N.E.; N.E.]	2/76 (2,6)	94,3 [N.E.; N.E.]	1,47 [0,30; 7,20]	0,631

Datenschnitt: 01.07.2019; Population: Safety Population.

<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).

<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.

<sup>(3)</sup> HR basiert auf einem Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.

<sup>(4)</sup> Der p-Wert des HR basiert auf einem Log-Rank-Test.

BSC = Best-Supportive-Care; HR = Hazard-Ratio; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar; UE = unerwünschtes Ereignis  
Quelle: [20]

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit mindestens einem UE auf Ebene der SOC und PT in den ersten 24 Wochen der Behandlungsphase, zeigte sich auch in der Analyse der Zeit bis zum ersten UE bis zum Datenschnitt vom 01.07.2019 für den Großteil der SOC und PT kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-78).

Statistisch signifikante Unterschiede zuungunsten von Luspatercept lagen lediglich für die SOC „Gastrointestinale Erkrankungen“ (HR = 1,76; 95 %-KI: [1,14; 2,70]; p = 0,010), „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ (HR = 1,57; 95 %-KI: [1,08; 2,27]; p = 0,016) und „Erkrankungen des Nervensystems“ (HR = 1,67; 95 %-KI: [1,02; 2,74]; p = 0,038) vor. Wie bereits ausgeführt, spiegelt sich der Unterschied der gastrointestinalen Erkrankungen nicht in einer Verschlechterung der Symptomatik wider (siehe Tabelle 4-41). Darüber hinaus ist das Ausmaß des Unterschieds, insbesondere für die SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ und „Erkrankungen des Nervensystems“ von geringfügigem Ausmaß.

Auf Ebene der PT lagen nur für die PT „Übelkeit“ (HR = 2,41; 95 %-KI: [1,00; 5,80]; p = 0,042), „Rückenschmerzen“ (HR = 2,63; 95 %-KI: [1,01; 6,83]; p = 0,039) und „Schwindelgefühl“ (HR = 3,44; 95 %-KI: [1,20; 9,80]; p = 0,014) statistisch signifikante Ergebnisse zuungunsten von Luspatercept vor. Auch hier spiegelt sich der Unterschied für die PT „Übelkeit“ und „Schwindelgefühl“ nicht in einer Verschlechterung der Symptomatik, die bereits über den EORTC QLQ-C30-Fragebogen erhoben wurde, wider (siehe Tabelle 4-41). Darüber hinaus ist das Ausmaß des Unterschieds für die PT „Übelkeit“ und „Rückenschmerzen“ von geringfügigem Ausmaß.

Ein statistisch signifikanter Vorteil für Luspatercept zeigte sich für den PT „Arthralgie“ (HR = 0,32; 95 %-KI: [0,12; 0,82]; p = 0,014).

Zugehörige Kaplan-Meier-Kurven sind im Anhang 4-G dargestellt.

**Schwerwiegende UE**

Tabelle 4-79: Zusatzanalyse 2: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten SUE auf Ebene der SOC und PT (SUE ab Inzidenzen von mindestens 5 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST SOC PT	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
SOC: Infektionen und parasitäre Erkrankungen	24/153 (15,7)	N.E. [N.E.; N.E.]	10/76 (13,2)	N.E. [49,4; N.E.]	0,73 [0,33; 1,60]	0,425
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	15/153 (9,8)	N.E. [143,0; N.E.]	7/76 (9,2)	N.E. [N.E.; N.E.]	0,54 [0,20; 1,48]	0,226
SOC: Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen)	13/153 (8,5)	N.E. [N.E.; N.E.]	2/76 (2,6)	N.E. [N.E.; N.E.]	1,60 [0,33; 7,70]	0,551
SOC: Herzerkrankungen	12/153 (7,8)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	3,53 [0,44; 28,41]	0,206
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	8/153 (5,2)	N.E. [N.E.; N.E.]	3/76 (3,9)	N.E. [N.E.; N.E.]	0,61 [0,14; 2,72]	0,518
SOC: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	10/153 (6,5)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
SOC: Erkrankungen des Gastrointestinaltrakts	9/153 (5,9)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.

Datenschnitt: 01.07.2019; Population: Safety Population.  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.  
<sup>(3)</sup> HR basiert auf einem Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.  
<sup>(4)</sup> Der p-Wert des HR basiert auf einem Log-Rank-Test.

BSC = Best-Supportive-Care; HR = Hazard-Ratio; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar;  
SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis  
Quelle: [20]

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit mindestens einem SUE auf Ebene der SOC und PT in den ersten 24 Wochen der Behandlungsphase, zeigte sich auch in der Analyse der Zeit bis zum ersten SUE bis zum Datenschnitt vom 01.07.2019 über alle SOC und PT hinweg kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-79).

Zugehörige Kaplan-Meier-Kurven sind im Anhang 4-G dargestellt.

### **Schwere UE (CTCAE $\geq 3$ )**

Tabelle 4-80: Zusatzanalyse 2: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) auf Ebene der SOC und PT (schwere UE ab Inzidenzen von mindestens 5 % bzw. mindestens 1 % und mindestens 10 Ereignissen in mindestens 1 Behandlungsarm); Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST SOC PT	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
SOC: Erkrankungen des Blutes und des Lymphsystems	23/153 (15,0)	N.E. [N.E.; N.E.]	11/76 (14,5)	N.E. [N.E.; N.E.]	0,62 [0,29; 1,34]	0,222
PT: Anämie	15/153 (9,8)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [N.E.; N.E.]	0,97 [0,34; 2,80]	0,959
PT: Neutropenie	7/153 (4,6)	N.E. [N.E.; N.E.]	6/76 (7,9)	N.E. [N.E.; N.E.]	0,41 [0,13; 1,29]	0,118
SOC: Infektionen und parasitäre Erkrankungen	24/153 (15,7)	N.E. [N.E.; N.E.]	9/76 (11,8)	N.E. [49,4; N.E.]	0,80 [0,35; 1,80]	0,584
SOC: Stoffwechsel- und Ernährungs- störungen	24/153 (15,7)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	1,80 [0,60; 5,40]	0,285
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	17/153 (11,1)	N.E. [N.E.; N.E.]	5/76 (6,6)	N.E. [56,6; N.E.]	1,16 [0,41; 3,28]	0,779
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	15/153 (9,8)	N.E. [143,0; N.E.]	6/76 (7,9)	N.E. [N.E.; N.E.]	0,62 [0,22; 1,78]	0,372
PT: Sturz	10/153 (6,5)	N.E. [N.E.; N.E.]	2/76 (2,6)	N.E. [N.E.; N.E.]	1,75 [0,37; 8,39]	0,477
SOC: Erkrankungen des Nervensystems	13/153 (8,5)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [55,1; N.E.]	0,93 [0,28; 3,03]	0,900

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
SOC PT	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
SOC: Erkrankungen der Atemwege, des Brustraums und Mediastinums	12/153 (7,8)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	0,76 [0,22; 2,59]	0,658
SOC: Herzerkrankungen	13/153 (8,5)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	4,03 [0,51; 31,99]	0,154
SOC: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	9/153 (5,9)	N.E. [N.E.; N.E.]	4/76 (5,3)	N.E. [N.E.; N.E.]	0,69 [0,20; 2,42]	0,563
SOC: Erkrankungen des Gastrointestinaltrakts	10/153 (6,5)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.

Datenschnitt: 01.07.2019; Population: Safety Population.  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.  
<sup>(3)</sup> HR basiert auf einem Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.  
<sup>(4)</sup> Der p-Wert des HR basiert auf einem Log-Rank-Test.

BSC = Best-Supportive-Care; CTCAE = *Common Terminology Criteria for Adverse Events*; HR = Hazard-Ratio; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar; UE = unerwünschtes Ereignis  
Quelle: [20]

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit mindestens einem schweren UE auf Ebene der SOC und PT in den ersten 24 Wochen der Behandlungsphase, zeigte sich auch in der Analyse der Zeit bis zum ersten schweren UE bis zum Datenschnitt vom 01.07.2019 über alle SOC und PT hinweg kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-80).

Auf Ebene der PT erreichten nur die Ereignisse „Neutropenie“, „Anämie“ und „Sturz“ Ereignishäufigkeiten von mindestens 5 %. Statistisch signifikante Unterschiede lagen auch hier nicht vor.

Zugehörige Kaplan-Meier-Kurven sind im Anhang 4-G dargestellt.

**Therapieabbrüche aufgrund von UE**

Tabelle 4-81: Zusatzanalyse 2: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum Therapieabbruch aufgrund von UE auf Ebene der SOC und PT; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST SOC PT	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>
SOC: Infektionen und parasitäre Erkrankungen	2/153 (1,3)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Sepsis	2/153 (1,3)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
SOC: Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	1/153 (0,7)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
PT: Arthralgie	0/153 (0,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
PT: Myalgie	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
SOC: Erkrankungen der Nieren und Harnwege	2/153 (1,3)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: chronische Nierenerkrankung	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Nierenversagen	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
SOC: Erkrankungen der Atemwege, des Brustraums und Mediastinums	1/153 (0,7)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
PT: Lungenfibrose	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: respiratorische Insuffizienz	0/153 (0,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
SOC: Erkrankungen des Blutes und des Lymphsystems	0/153 (0,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
PT: Thrombozytose	0/153 (0,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
SOC: Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Verkehrsunfall	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
SOC: Untersuchungen	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Alaninaminotransferase erhöht	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]

Studie MEDALIST SOC PT	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>
PT: Aspartataminotransferase erhöht	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
SOC: Gefäßerkrankungen	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Hämorrhagischer Schock	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
SOC: Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen)	8/153 (5,2)	N.E. [N.E.; N.E.]	2/76 (2,6)	N.E. [N.E.; N.E.]
PT: Anämie refraktär mit Blastenschub	3/153 (2,0)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Kolonadenom	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Myelodysplastisches Syndrom	2/153 (1,3)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
PT: Transformation zu einer akuten myeloischen Leukämie	2/153 (1,3)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
SOC: Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	3/153 (2,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
PT: Asthenie	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Ermüdung	2/153 (1,3)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Generelle Verschlechterung des physischen Gesundheitszustandes	0/153 (0,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]
SOC: Erkrankungen des Nervensystems	4/153 (2,6)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Erinnerungsvermögen eingeschränkt	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Hirnblutung	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Kopfschmerz	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]
PT: Morbus Parkinson	1/153 (0,7)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>	
SOC PT	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>
Datenschnitt: 01.07.2019; Population: Safety Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.  BSC = Best-Supportive-Care; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar; UE = unerwünschtes Ereignis Quelle: [20]				

Konsistent zu den Ergebnissen der Hauptanalyse, dem Anteil an Patienten mit einem Therapieabbruch aufgrund von UE auf Ebene der SOC und PT in den ersten 24 Wochen der Behandlungsphase, zeigten sich auch in der Analyse der Zeit bis zum Therapieabbruch aufgrund von UE bis zum Datenschnitt vom 01.07.2019 über alle SOC und PT hinweg nur vereinzelt Ereignisse in beiden Behandlungsarmen (siehe Tabelle 4-81). Insgesamt traten auch bis zum Datenschnitt vom 01.07.2019 nur wenig Therapieabbrüche aufgrund von UE auf, einzelne Ereignisse, die vornehmlich einen Therapieabbruch aufgrund eines UE auslösen, konnten nicht identifiziert werden.

Zugehörige Kaplan-Meier-Kurven sind im Anhang 4-G dargestellt.

### Zusatzanalyse 3: UE von besonderem Interesse

Tabelle 4-82: Zusatzanalyse 3: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – UE von besonderem Interesse; Woche 1 – 24

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
UE von besonderem Interesse Schweregrad	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>SMQ Malignitäten</b>						
Jegliche UE	4/153 (2,6)	1/76 (1,3)	2,01 [0,22; 18,33]	1,99 [0,23; 17,47]	1,30 [-2,30; 4,90]	0,528
SUE	4/153 (2,6)	1/76 (1,3)	2,01 [0,22; 18,33]	1,99 [0,23; 17,47]	1,30 [-2,30; 4,90]	0,528
Schwere UE (CTCAE ≥ 3)	2/153 (1,3)	1/76 (1,3)	0,99 [0,09; 11,13]	0,99 [0,09; 10,78]	-0,01 [-3,14; 3,12]	0,996

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
			OR [95 %-KI] <sup>(2)</sup>	RR [95 %- KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
UE von besonderem Interesse Schweregrad	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)				
<b>SMQ Prämaligne Erkrankungen</b>						
Jegliche UE	1/153 (0,7)	3/76 (3,9)	0,16 [0,02; 1,57]	0,17 [0,02; 1,57]	-3,29 [-7,85; 1,27]	0,074
SUE	1/153 (0,7)	1/76 (1,3)	0,49 [0,03; 8,00]	0,50 [0,03; 7,83]	-0,66 [-3,52; 2,20]	0,613
Schwere UE (CTCAE ≥ 3)	0/153 <sup>(4)</sup> (0,0)	0/76 <sup>(4)</sup> (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
<b>PT Neutropenie</b>						
Jegliche UE	6/153 (3,9)	6/76 (7,9)	0,48 [0,15; 1,53]	0,50 [0,17; 1,49]	-3,97 [-10,77; 2,82]	0,205
SUE	0/153 <sup>(4)</sup> (0,0)	0/76 <sup>(4)</sup> (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
Schwere UE (CTCAE ≥ 3)	5/153 (3,3)	5/76 (6,6)	0,48 [0,13; 1,71]	0,50 [0,15; 1,66]	-3,31 [-9,56; 2,93]	0,249
<b>PT Thrombozytopenie</b>						
Jegliche UE	2/153 (1,3)	1/76 (1,3)	0,99 [0,09; 11,13]	0,99 [0,09; 10,78]	-0,01 [-3,14; 3,12]	0,996
SUE	0/153 <sup>(4)</sup> (0,0)	0/76 <sup>(4)</sup> (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
Schwere UE (CTCAE ≥ 3)	0/153 <sup>(4)</sup> (0,0)	0/76 <sup>(4)</sup> (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
<p>Datenschnitt: 08.05.2018; Population: Safety Population. Stratifizierung: IPSS-R-Risikoklasse und durchschnittliche Transfusionslast zur Baseline.</p> <p><sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).</p> <p><sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet.</p> <p><sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.</p> <p><sup>(4)</sup> Nullzellenkorrektur zur Berechnung unstratifizierter Effektschätzer wurde herangezogen.</p> <p>ARR = Absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran Mantel Haenszel; CTCAE = <i>Common Terminology Criteria for Adverse Events</i>; IPSS-R = <i>International Prognostic Scoring System-revised</i>; KI = Konfidenzintervall; MedDRA = <i>Medical Dictionary for Regulatory Activities</i>; N.E. = nicht erreicht oder nicht berechenbar; OR = Odds Ratio; PT = bevorzugter Begriff (<i>preferred term</i>); RR = Relatives Risiko; SMQ = <i>Standardized MedDRA Query</i>; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis</p> <p>Quelle: [18]</p>						

Tabelle 4-83: Zusatzanalyse 3: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zum ersten UE von besonderem Interesse; Woche eins bis zum Datenschnitt vom 01.07.2019

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
<b>SMQ Malignitäten</b>						
Jegliche UE	6/153 (3,9)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	2,25 [0,26; 19,56]	0,450
SUE	6/153 (3,9)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	2,25 [0,26; 19,56]	0,450
Schwere UE (CTCAE ≥ 3)	2/153 (1,3)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	1,02 [0,09; 11,28]	0,985
<b>SMQ Prämaligne Erkrankungen</b>						
Jegliche UE	7/153 (4,6)	N.E. [N.E.; N.E.]	3/76 (3,9)	N.E. [N.E.; N.E.]	0,36 [0,07; 1,78]	0,194
SUE	7/153 (4,6)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	0,97 [0,10; 9,70]	0,983
Schwere UE (CTCAE ≥ 3)	5/153 (3,3)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
<b>PT Neutropenie</b>						
Jegliche UE	9/153 (5,9)	N.E. [N.E.; N.E.]	7/76 (9,2)	N.E. [N.E.; N.E.]	0,46 [0,16; 1,29]	0,132
SUE	0/153 (0,0)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
Schwere UE (CTCAE ≥ 3)	7/153 (4,6)	N.E. [N.E.; N.E.]	6/76 (7,9)	N.E. [N.E.; N.E.]	0,41 [0,13; 1,29]	0,118
<b>PT Thrombozytopenie</b>						
Jegliche UE	3/153 (2,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	1,12 [0,11; 11,47]	0,923
SUE	0/153 (0,0)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
Schwere UE (CTCAE ≥ 3)	0/153 (0,0)	N.E. [N.E.; N.E.]	0/76 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
Endpunkt	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Datenschnitt: 01.07.2019; Population: Safety Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet. <sup>(3)</sup> HR basiert auf einem Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable. <sup>(4)</sup> Der p-Wert des HR basiert auf einem Log-Rank-Test.  BSC = Best-Supportive-Care; CTCAE = <i>Common Terminology Criteria for Adverse Events</i> ; HR = Hazard-Ratio; KI = Konfidenzintervall; MedDRA = <i>Medical Dictionary for Regulatory Activities</i> ; N.E. = nicht erreicht oder nicht berechenbar; PT = bevorzugter Begriff ( <i>preferred term</i> ); SMQ = <i>Standardized MedDRA Query</i> ; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis Quelle: [20]						

Die Ereignisse der beiden SMQ „Malignitäten“ und „Prämaligne Erkrankungen“ sowie der PT „Neutropenie“ und „Thrombozytopenie“ wurden als UE von besonderem Interesse ausgewertet.

Insgesamt traten im Studienzeitraum der primären Behandlungsphase bis Woche 24 nur bei sehr wenigen Patienten UE von besonderem Interesse auf. 2,6 % (4/153) der Patienten im Luspatercept-Arm und 1,3 % (1/76) der Patienten im Placebo-Arm hatten ein Ereignis der SMQ „Malignitäten“ und 0,7 % (1/153) bzw. 3,9 % (3/76) der Patienten hatten ein Ereignis der SMQ „Prämaligne Erkrankungen“. 3,9 % (6/153) der Patienten im Luspatercept- und 7,9 % (6/67) im Placebo-Arm hatten das UE „Neutropenie“, wobei keines der Ereignisse schwerwiegend war. 1,3 % (2/153) der Patienten unter Luspatercept und 1,3 % (1/76) der Patienten unter Placebo hatten eine Thrombozytopenie, die jedoch nicht schwer (CTCAE < 3) oder schwerwiegend war (siehe Tabelle 4-82). Statistisch signifikante Unterschiede lagen über alle betrachteten UE von besonderem Interesse unabhängig vom Schweregrad hinweg nicht vor.

Bis zum Datenschnitt vom 01.07.2019 kamen lediglich vereinzelt Ereignisse im Luspatercept-Arm und im Placebo-Arm hinzu. Unter Luspatercept hatte zwei Patienten mehr ein Ereignis der SMQ „Malignitäten“, drei Patienten mehr ein Ereignis der SMQ „Prämaligne Erkrankungen“ und des PT „Neutropenie“ und ein Patient mehr ein Ereignis des PT „Thrombozytopenie“. Im Placebo-Arm hatte ein weiterer das Ereignis des PT „Neutropenie“ (siehe Tabelle 4-83).

Insgesamt zeigte sich somit hinsichtlich der UE von besonderem Interesse ein gutes Sicherheitsprofil von Luspatercept.

Zugehörige Kaplan-Meier-Kurven sind im Anhang 4-G dargestellt.

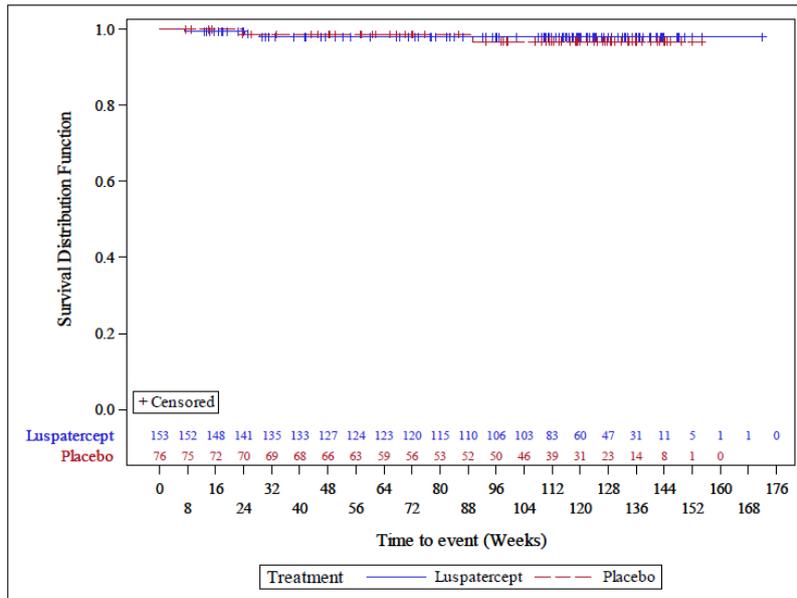
**Zusatzanalyse 4: Zeit bis zur Progression in eine AML**

Tabelle 4-84: Zusatzanalyse 4: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Zeit bis zur Progression in eine AML

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>		Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>	
	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	Patienten mit Ereignis n/N (%)	Median (Wochen) [95 %-KI] <sup>(2)</sup>	HR [95 %-KI] <sup>(3)</sup>	p-Wert <sup>(4)</sup>
Woche 1 bis zum Datenschnitt vom 08.05.2018	3/153 (2,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	1,57 [0,16; 15,19]	0,693
Woche 1 bis zum Datenschnitt vom 07.01.2019	3/153 (2,0)	N.E. [N.E.; N.E.]	1/76 (1,3)	N.E. [N.E.; N.E.]	1,57 [0,16; 15,19]	0,693
Woche 1 bis zum Datenschnitt vom 01.07.2019	3/153 (2,0)	N.E. [N.E.; N.E.]	2/76 (2,6)	N.E. [N.E.; N.E.]	0,75 [0,12; 4,51]	0,753

Population: ITT-Population; Stratifizierung: IPSS-R-Risikogruppe und durchschnittliche Transfusionslast zur Baseline  
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).  
<sup>(2)</sup> KI des Medians wurde über eine Log-Log-Transformation (nach Brookmeyer und Crowley) berechnet.  
<sup>(3)</sup> HR basiert auf einem stratifizierten Cox-Proportional-Hazards-Modell mit Behandlung als Kovariable.  
<sup>(4)</sup> Der p-Wert des HR basiert auf einem stratifizierten Log-Rank-Test.

AML = akute myeloische Leukämie; BSC = Best-Supportive-Care; HR = Hazard-Ratio; ITT = Intention-to-Treat; IPSS-R = *International Prognostic Scoring System-revised*; KI = Konfidenzintervall; N.E. = nicht erreicht oder nicht berechenbar  
Quelle: [1, 18, 20, 94]



Datenschnitt: 01.07.2019; + = Zensierung

Quelle: [20]

Abbildung 4-63: Zusatzanalyse 4: Kaplan-Meier-Kurve für die Zeit bis zur Progression in eine AML aus MEDALIST; Analysezeitraum bis zum Datenschnitt vom 01.07.2019

Unter Berücksichtigung des gesamten verfügbaren Beobachtungszeitraums (bis zum Datenschnitt vom 01.07.2019) hatten unter Luspatercept nur 2,0 % (3/153) und unter Placebo nur 1,3 % (1/67) der Patienten eine Progression in eine AML. Dementsprechend wurde die mediane Zeit bis zur Progression in eine AML nicht erreicht und es lag kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsarmen vor (HR = 1,57; 95 %-KI: [0,16; 15,19]; p = 0,693) (siehe Tabelle 4-84 & Abbildung 4-63).

**Zusatzanalyse 5: Progression in ein Hochrisiko-MDS**

Tabelle 4-85: Zusatzanalyse 5: Ergebnisse für den Endpunkt Sicherheit aus MEDALIST – Progression in ein Hochrisiko-MDS

Studie MEDALIST	Luspatercept + BSC <sup>(1)</sup>		Placebo + BSC <sup>(1)</sup>	
	N	Patienten mit Ereignis n (%)	N	Patienten mit Ereignis n (%)
Woche 1 bis zum Datenschnitt vom 08.05.2018	153	1 (0,7)	76	1 (1,3)
Woche 1 bis zum Datenschnitt vom 07.01.2019	153	3 (2,0)	76	1 (1,3)
Woche 1 bis zum Datenschnitt vom 01.07.2019	153	6 (3,9)	76	1 (1,3)

Population: ITT-Population; Progression in ein Hochrisiko-MDS war definiert als Ereignisse des PT „Anämie refraktär mit Blastenüberschuss“ oder „Myelodysplastisches Syndrom“  
<sup>(1)</sup>BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).

BSC = Best-Supportive-Care; ITT = Intention-to-Treat; PT = bevorzugter Begriff (*preferred term*)  
Quelle: [1, 20]

Im gesamten Studienzeitraum traten bis zum Datenschnitt vom 01.07.2019 nur vereinzelt Ereignisse einer Progression in ein Hochrisiko-MDS statt (siehe Tabelle 4-85).

**Zusammenfassung der Ergebnisse zum Endpunkt Sicherheit**

Insgesamt weisen die Ergebnisse zum Endpunkt Sicherheit auf ein sehr gutes Sicherheits- und Verträglichkeitsprofil von Luspatercept hin. Da Luspatercept ein aktives, zielgerichtetes Arzneimittel ist, wäre im Vergleich zu Placebo potentiell auch mit einer erhöhten Rate an Nebenwirkungen zu rechnen gewesen. Da dies nicht der Fall ist, also keine relevanten Unterschiede hinsichtlich des Auftretens von UE beobachtet werden konnten, die den Einsatz von Luspatercept einschränken würden, belegt umso mehr die äußerst gute Verträglichkeit von Luspatercept.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

### **Zusammenfassung der Ergebnisse mithilfe von Meta-Analysen**

Nicht zutreffend. Es liegen Ergebnisse einer RCT vor, sodass keine Meta-Analysen durchgeführt wurden.

### **Übertragbarkeit auf den deutschen Versorgungskontext**

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext ist in Abschnitt 4.3.1.2.1 beschrieben.

Für den Endpunkt Sicherheit liegt keine abweichende Bewertung vor. Insgesamt ist somit von einer guten Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext auszugehen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.1.3.2 Subgruppenanalysen – RCT

Für die Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.<sup>17</sup>

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

- Subgruppenanalysen sind nur für die Merkmale (z.B. Alter) durchzuführen, bei denen die resultierenden Subgruppen jeweils mindestens 10 Patienten umfassen.
- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.
- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ( $p < 0,05$ ) dargestellt werden.
- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.
- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.
- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmodifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.
- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.
- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

---

<sup>16</sup> unbesetzt

*Beschreiben Sie die Ergebnisse von Subgruppenanalysen. Stellen Sie dabei zunächst tabellarisch dar, zu welchen der in Abschnitt 4.2.5.5 genannten Effektmodifikatoren Subgruppenanalysen zu den relevanten Endpunkten vorliegen, und ob diese a priori geplant und im Studienprotokoll festgelegt waren oder posthoc durchgeführt wurden.*

*Orientieren Sie sich an der beispielhaften Angabe in der ersten Tabellenzeile.*

Im Folgenden werden die Ergebnisse von Subgruppenanalysen für sämtliche patientenrelevanten Endpunkte dargestellt, die zur Ableitung des Zusatznutzens herangezogen werden. Die Auswahl der relevanten Subgruppen richtet sich einerseits nach den Vorgaben des G-BA und andererseits danach, welche Subgruppen für den primären Endpunkt (Transfusionsfreiheit) und den hauptsächlich sekundären Endpunkt (HI-E) in der Studie MEDALIST prädefiniert waren (siehe auch Abschnitt 4.2.5.5).

Abweichend von den im Studienprotokoll prädefinierten Subgruppenmerkmalen wurde auf eine Darstellung der Subgruppen „weiß“, „schwarz oder afroamerikanisch“, „andere“ und „nicht berichtet“ des Subgruppenmerkmals Ethnische Zugehörigkeit verzichtet. Hintergrund ist, dass in den Subgruppen „schwarz oder afroamerikanisch“ (ein Patient im Luspatercept-Arm) und „andere“ (jeweils ein Patient in beiden Behandlungsarmen) insgesamt nur sehr wenig Patienten enthalten waren, die eine sinnvolle Interpretation nicht möglich machen. Eine sinnvolle medizinische Interpretation von potentiellen Effektmodifikationen zwischen der Subgruppe „weiß“ und „nicht berichtet“ ist ebenfalls nicht möglich, da diese beiden Subgruppen nicht sinnvoll voneinander abgegrenzt werden können.

Auf eine Analyse nach Länder- oder Zentrumseffekten wurde verzichtet, da aussagekräftige Ergebnisse aufgrund von geringen Patientenzahlen pro Studienzentrum bzw. pro Land nicht zu erwarten waren und somit nicht geeignet sind, um Schlussfolgerungen für diese Subgruppen abzuleiten. Stattdessen wurden post-hoc definierte Subgruppenanalysen zu den übergeordneten, geografischen Regionen durchgeführt. In MEDALIST eingeschlossene Patienten waren weitestgehend gleichmäßig auf über 60 Studienzentren in 11 Ländern verteilt, sodass zudem nicht von einem Ungleichgewicht aufgrund von einzelnen, sehr großen Zentren oder Ländern auszugehen ist. Das Subgruppenmerkmal „Nordamerika“ fasst die in Kanada und den USA behandelten Patienten zusammen, während die Länder Belgien, Deutschland, Frankreich, Großbritannien, Italien, Niederlande, Schweden, Spanien und die Türkei unter dem Subgruppenmerkmal Europa zusammengefasst werden.

Eine Übersicht der durchgeführten Subgruppenanalysen auf Endpunktebene findet sich in Tabelle 4-86.

Tabelle 4-86: Matrix der durchgeführten Subgruppenanalysen

Subgruppenmerkmal Studie MEDALIST	Mortalität (OS)	Morbidität				Lebensqualität		Sicherheit <sup>(1)</sup>
		Transfusionsfreiheit	Reduktion der Transfusionslast (HI-E)	Symptomatik (EORTC QLQ-C30)	Hospitalisierungen	EORTC QLQ-C30	QoL-E	
Alter	○	●	●	○	○	○	○	○
Geschlecht	○	●	●	○	○	○	○	○
WHO-Klassifizierung der MDS	○	●	●	○	○	○	○	○
Zeit seit initialer MDS-Diagnose	○	●	●	○	○	○	○	○
ECOG-PS	○	●	●	○	○	○	○	○
Transfusionslast zur Baseline I <sup>(2)</sup>	○	●	●	○	○	○	○	○
Transfusionslast zur Baseline II <sup>(3)</sup>	○	●	●	○	○	○	○	○
IPSS-R-Risikogruppe	○	●	●	○	○	○	○	○
Dauer vorherige ESA-Behandlung	○	●	●	○	○	○	○	○
Nierenfunktionsstatus (Kreatinin-Clearance)	○	●	●	○	○	○	○	○
SF3B1-Mutationsstatus	○	●	●	○	○	○	○	○
Serum-EPO-Spiegel	○	●	●	○	○	○	○	○
Thrombozytenzahl	○	●	●	○	○	○	○	○
Region	○	○	○	○	○	○	○	○

●: A priori geplante Subgruppenanalyse. ○: Posthoc durchgeführte Subgruppenanalyse. n. d.: Subgruppenanalyse nicht durchgeführt.

<sup>(1)</sup> Sicherheit umfasst Auswertungen zu den Kategorien jegliche UE, SUE, schwere UE und Therapieabbrüche aufgrund von UE.

<sup>(2)</sup> Kategorien des Subgruppenmerkmals: < 6 EK-Einheiten pro acht Wochen vs. ≥ 6 EK-Einheiten pro acht Wochen.

<sup>(3)</sup> Kategorien des Subgruppenmerkmals: < 4 EK-Einheiten pro acht Wochen vs. 4 bis < 6 EK-Einheiten pro acht Wochen vs. ≥ 6 EK-Einheiten pro acht Wochen.

ECOG = Eastern Cooperative Oncology Group; EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module; EPO = Erythropoetin; ESA = Erythropoese-stimulierende Agenzien (erythropoiesis stimulating agents); HI-E = erythroides Ansprechen (hematologic improvement-erythroid); IPSS-R = International Prognostic Scoring System-revised; MDS = Myelodysplastische Syndrome; OS = Gesamtüberleben (overall survival); PS = Performance Status; QoL-E = Quality of Life-E; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis; WHO = World Health Organization

*Stellen Sie anschließend in Tabelle 4-87 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante ( $p < 0,05$ ) Interaktionsterme.*

Die Herangehensweise für die Prüfung und Interpretation der Subgruppenanalyse erfolgte wie in Abschnitt 4.2.5.5 beschrieben.

Subgruppen mit einem Interaktions-p-Wert  $\geq 0,05$  werden als homogen hinsichtlich des betrachteten Behandlungseffekts beurteilt. Die Ergebnisse werden ergänzend in Anhang 4-G dargestellt, die Bewertung des Effekts erfolgt ausschließlich basierend auf den Ergebnissen der Gesamtpopulation.

Bei Vorliegen eines statistisch signifikanten Interaktionstests (Interaktions-p-Wert  $< 0,05$ ) werden die Ergebnisse der jeweiligen Subgruppe detailliert dargestellt. Bei der Bewertung der Ergebnisse sind insbesondere die Limitationen hinsichtlich der Möglichkeit von falsch positiven Ergebnissen aufgrund des multiplen Testens zu berücksichtigen. Dementsprechend kann es bereits alleinig durch die angenommene Irrtumswahrscheinlichkeit, also per Definition bei 5 % der durchgeführten Tests, zu einem falsch positiven Ergebnis kommen.

Bei den im vorliegenden Fall durchgeführten 602 Subgruppenanalysen (43 Hauptanalysen [zu 8 Endpunkten] mit jeweils 14 Subgruppen) muss aus statistischer Sicht bereits mit rund 30 falsch positiven Ergebnissen gerechnet werden. Die Ergebnisse der Subgruppenanalysen sind dementsprechend kritisch unter Berücksichtigung dieser Irrtumswahrscheinlichkeit zu prüfen, ob eine erklärende klinische Rationale und/oder die Konsistenz der Ergebnisse über mehrere Endpunkte oder Studien vorliegt. Im vorliegenden Dossier wurden insgesamt 19 statistisch signifikante Interaktions-p-Werte berichtet. Bei Annahme einer Unabhängigkeit der Testungen untereinander und einer Abwesenheit von tatsächlichen Effektmodifikationen wäre eine solche Anzahl positiver Tests mit einer Wahrscheinlichkeit von 98,9 % zu erwarten gewesen. Demgegenüber liegt die Wahrscheinlichkeit, dass unter gleichen Bedingungen kein statistisch signifikanter Interaktionstest beobachtet würde, bei  $< 0,001$  %.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Im Folgenden werden die Ergebnisse der Interaktionstests für alle untersuchten Subgruppen Endpunkt dargestellt.

Tabelle 4-87: Ergebnisse des Interaktionstests der Subgruppenanalysen je Endpunkt für MEDALIST

Endpunkt	p-Wert aus Interaktionstest													
	Alter	Geschlecht	Region	WHO-Klassifizierung der MDS	Zeit seit initialer MDS-Diagnose	ECOG Performance-Status	Transfusionslast zur Baseline I <sup>(1)</sup>	Transfusionslast zur Baseline II <sup>(2)</sup>	IPSS-R Risikogruppe	Thrombozytenzahl	Nierenfunktionsstatus (Kreatinin-Clearance)	Serum-EPO-Spiegel	SF3B1-Mutationsstatus	Dauer vorherige ESA-Behandlung
<i>Mortalität</i>														
Gesamtüberleben	0,584	0,544	0,989	0,993	0,997	0,790	0,229	0,815	0,457	0,947	0,998	0,943	0,174	0,298
<i>Morbidität</i>														
<i>Transfusionsfreiheit</i>														
Transfusionsfreiheit ≥ 8 Wochen	0,837	0,194	0,742	0,972	0,595	0,709	0,964	0,303	0,800	0,470	0,878	0,653	0,350	0,753
Transfusionsfreiheit ≥ 12 Wochen	0,551	0,689	0,480	0,669	0,353	0,835	0,578	0,780	0,703	0,428	0,727	0,691	0,529	0,860
Transfusionsfreiheit ≥ 16 Wochen	0,632	0,992	0,369	0,364	0,109	0,590	0,383	0,747	0,212	0,341	0,573	0,875	0,990	0,809
Transfusionsfreiheit ≥ 24 Wochen	0,816	0,918	0,715	0,215	0,505	0,765	0,081	0,327	0,398	0,247	0,335	0,974	N.E.	0,590
<i>Reduktion der Transfusionslast</i>														
HI-E (IWG 2006)	0,465	0,445	0,875	0,858	0,146	0,566	<b>0,042</b>	0,136	0,339	0,084	0,816	0,988	0,339	0,998
HI-E I <sup>(3)</sup>	0,503	0,693	0,960	0,911	0,572	0,712	0,216	0,216	0,437	0,161	0,684	0,983	0,499	0,879
HI-E II <sup>(4)</sup>	0,993	0,828	0,648	N.E.	0,075	0,609	N.E.	N.E.	0,125	0,796	N.E.	0,910	0,905	0,589

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Endpunkt	p-Wert aus Interaktionstest													
	Alter	Geschlecht	Region	WHO-Klassifizierung der MDS	Zeit seit initialer MDS-Diagnose	ECOG Performance-Status	Transfusionslast zur Baseline I <sup>(1)</sup>	Transfusionslast zur Baseline II <sup>(2)</sup>	IPSS-R Risikogruppe	Thrombozytenzahl	Nierenfunktionsstatus (Kreatinin-Clearance)	Serum-EPO-Spiegel	SF3B1-Mutationsstatus	Dauer vorherige ESA-Behandlung
<i>Symptomatik (EORTC QLQ-C30)</i>														
Fatigue	<b>0,026</b>	0,405	0,585	0,961	0,540	0,202	0,240	0,082	0,191	0,506	0,622	0,322	0,727	0,399
Übelkeit und Erbrechen	0,958	0,517	0,843	0,966	0,953	0,337	0,594	0,685	0,631	0,995	0,657	0,683	0,950	0,959
Schmerz	0,943	0,377	0,104	1,000	0,328	0,946	0,987	0,607	0,598	0,833	0,987	0,485	0,996	0,534
Dyspnoe	0,532	0,669	0,397	0,949	0,829	0,588	0,399	0,596	0,681	0,567	0,810	0,776	0,661	0,262
Schlaflosigkeit	0,884	0,995	0,745	0,895	0,489	0,580	0,517	0,284	0,830	0,579	0,949	0,388	0,412	0,989
Appetitminderung	0,581	0,466	0,217	0,999	0,249	0,606	0,169	0,339	0,193	0,545	0,976	0,484	0,954	0,259
Obstipation	0,676	0,749	0,765	0,999	0,204	0,678	0,709	0,343	0,935	0,493	0,561	0,949	0,402	0,879
Diarrhö	0,967	0,457	0,454	1,000	0,916	0,954	0,783	0,923	0,646	1,000	0,748	0,958	0,950	0,958
Finanzielle Schwierigkeiten	0,188	0,770	0,905	0,999	0,870	0,523	0,794	0,551	0,334	0,891	0,117	0,908	0,401	0,717
<i>Hospitalisierungen</i>														
Hospitalisierungen jeglicher Ursache	0,278	0,917	0,173	0,697	0,121	0,445	0,521	0,139	0,743	0,775	0,079	0,132	0,424	0,129
<i>Lebensqualität</i>														
<i>Gesundheitsbezogene Lebensqualität (EORTC QLQ-C30)</i>														
Körperliche Funktion	0,108	0,846	0,817	0,947	0,506	0,979	0,811	0,825	0,465	0,833	0,218	0,313	0,353	0,865
Rollenfunktion	0,728	0,801	0,291	0,946	0,220	0,788	0,933	0,588	0,061	0,672	0,471	0,233	0,949	0,822

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Endpunkt	p-Wert aus Interaktionstest													
	Alter	Geschlecht	Region	WHO-Klassifizierung der MDS	Zeit seit initialer MDS-Diagnose	ECOG Performance-Status	Transfusionslast zur Baseline I <sup>(1)</sup>	Transfusionslast zur Baseline II <sup>(2)</sup>	IPSS-R Risikogruppe	Thrombozytenzahl	Nierenfunktionsstatus (Kreatinin-Clearance)	Serum-EPO-Spiegel	SF3B1-Mutationsstatus	Dauer vorherige ESA-Behandlung
Kognitive Funktion	0,102	0,694	0,779	0,297	0,951	0,236	0,275	0,519	0,350	0,452	0,513	0,656	0,949	0,254
Emotionale Funktion	0,965	0,710	0,323	0,999	0,823	0,484	0,327	0,352	0,357	0,842	0,290	0,666	0,224	0,998
Soziale Funktion	0,290	0,733	0,705	0,953	0,559	0,537	0,426	0,450	0,271	0,982	0,589	0,233	0,951	0,220
Gesundheitlicher Status insgesamt	0,618	0,058	0,479	0,956	0,163	0,602	0,201	0,194	0,172	0,066	0,658	0,773	0,673	0,213
<i>Krankheitsspezifische Lebensqualität (QoL-E)</i>														
Körperliches Wohlbefinden	<b>0,003</b>	0,198	<b>0,015</b>	0,539	0,945	0,669	0,306	0,496	0,721	0,332	0,684	0,813	0,364	0,392
Funktionales Wohlbefinden	<b>0,022</b>	0,613	0,928	0,936	0,359	0,145	0,091	0,169	0,505	0,891	0,848	0,797	0,883	0,997
Soziales/familiäres Wohlbefinden	0,115	0,438	0,907	0,614	0,938	0,371	0,654	0,844	0,291	0,411	0,223	0,983	<b>0,049</b>	0,251
sexuelles Wohlbefinden	0,073	0,389	0,703	0,448	0,449	0,918	0,413	0,682	0,363	0,334	0,829	0,923	0,524	0,320
Fatigue	0,119	0,579	0,789	0,445	0,720	0,642	0,323	0,596	0,987	0,438	0,748	0,396	0,702	0,954
MDS-bezogene Störungen	0,207	0,931	0,737	0,857	0,888	0,584	0,389	0,691	0,296	0,353	0,140	0,183	0,959	0,622
Allgemeines Wohlbefinden <sup>(5)</sup>	<b>0,008</b>	0,967	0,581	0,450	0,746	0,374	0,073	0,172	0,862	0,567	0,702	0,681	0,175	0,880
Behandlungseffekt <sup>(6)</sup>	<b>0,006</b>	0,346	0,517	0,967	0,892	0,235	0,179	0,351	0,590	0,912	0,193	0,544	0,868	0,977
Gesamtsummenskala <sup>(7)</sup>	<b>0,030</b>	0,944	0,590	0,870	0,940	0,630	0,614	0,816	0,370	0,431	0,331	0,513	0,295	0,956

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Endpunkt	p-Wert aus Interaktionstest													
	Alter	Geschlecht	Region	WHO-Klassifizierung der MDS	Zeit seit initialer MDS-Diagnose	ECOG Performance-Status	Transfusionslast zur Baseline I <sup>(1)</sup>	Transfusionslast zur Baseline II <sup>(2)</sup>	IPSS-R Risikogruppe	Thrombozytenzahl	Nierenfunktionsstatus (Kreatinin-Clearance)	Serum-EPO-Spiegel	SF3B1-Mutationsstatus	Dauer vorherige ESA-Behandlung
<b>Sicherheit</b>														
jegliche UE	0,639	0,722	0,654	0,301	0,924	0,758	0,225	0,057	0,984	0,200	0,962	<b>0,015</b>	0,459	<b>0,047</b>
Gastrointestinale Erkrankungen (SOC)	0,936	0,385	0,733	0,475	0,422	0,646	0,227	0,394	0,189	0,742	<b>0,010</b>	0,308	0,457	0,313
Erkrankungen des Nervensystems (SOC)	<b>0,034</b>	0,428	0,646	0,218	<b>0,005</b>	0,092	0,278	0,526	0,571	0,066	0,795	0,504	0,493	0,794
Augenerkrankungen (SOC)	0,967	0,212	0,357	0,395	0,812	0,392	0,177	0,486	0,320	0,760	0,798	0,462	N.E.	0,870
Asthenie (PT)	0,903	0,686	0,598	0,910	0,763	0,782	0,206	0,387	0,915	0,657	<b>0,002</b>	0,264	0,454	<b>0,032</b>
Schwindelgefühl (PT)	0,643	0,241	0,531	N.E.	<b>0,043</b>	0,230	0,130	0,346	0,264	0,602	0,987	0,501	N.E.	0,784
SUE	0,934	0,248	0,573	0,743	0,213	0,299	0,452	0,226	0,329	0,822	<b>0,039</b>	0,347	0,299	0,284
schwere UE (CTCAE ≥ 3)	0,887	0,934	0,723	0,450	0,739	0,348	0,868	0,056	0,756	0,494	0,092	0,997	0,202	0,547
Erkrankungen des Nervensystems (SOC)	0,854	0,933	0,528	N.E.	0,995	0,990	0,834	0,959	0,642	0,482	0,691	0,984	0,629	0,988
Therapieabbrüche aufgrund von UE	0,245	0,184	0,824	0,889	0,203	0,302	0,079	0,102	0,747	<b>0,017</b>	0,459	0,413	0,412	0,477

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Endpunkt	p-Wert aus Interaktionstest												
	Alter	Geschlecht	Region	WHO-Klassifizierung der MDS	Zeit seit initialer MDS-Diagnose	ECOG Performance-Status	Transfusionslast zur Baseline I <sup>(1)</sup>	Transfusionslast zur Baseline II <sup>(2)</sup>	IPSS-R Risikogruppe	Thrombozytenzahl	Nierenfunktionsstatus (Kreatinin-Clearance)	Serum-EPO-Spiegel	SF3B1-Mutationsstatus
<p>Datenschnitt: 08.05.2018; <b>fett</b> = statistisch signifikanter Interaktions-p-Wert</p> <p><sup>(1)</sup> Kategorien des Subgruppenmerkmals: &lt; 6 EK-Einheiten pro acht Wochen vs. ≥ 6 EK-Einheiten pro acht Wochen</p> <p><sup>(2)</sup> Kategorien des Subgruppenmerkmals: &lt; 4 EK-Einheiten pro acht Wochen vs. 4 bis &lt; 6 EK-Einheiten pro acht Wochen vs. ≥ 6 EK-Einheiten pro acht Wochen</p> <p><sup>(3)</sup> HI-E I = Reduktion der Transfusionslast um ≥ 4 EK-Einheiten pro acht Wochen für Patienten mit hoher Transfusionslast zur Baseline (≥ 4 EK-Einheiten pro acht Wochen)</p> <p><sup>(4)</sup> HI-E II = Erhöhung des Hb-Werts um ≥ 1,5 g/dl über acht Wochen für Patienten mit niedriger Transfusionslast zur Baseline (&lt; 4 EK-Einheiten pro acht Wochen)</p> <p><sup>(5)</sup> Summe aller Domänen mit Ausnahme der MDS-bezogenen Störungen</p> <p><sup>(6)</sup> Summe der physischen, funktionellen und MDS-spezifischen Dimension</p> <p><sup>(7)</sup> Summe aller Domänen</p> <p>CTCAE = <i>Common Terminology Criteria for Adverse Events</i>; ECOG = <i>Eastern Cooperative Oncology Group</i>; EORTC QLQ-C30 = <i>European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module</i>; EPO = Erythropoetin; ESA = Erythropoese-stimulierende Agenzien (<i>erythropoiesis stimulating agents</i>); HI-E = erythroides Ansprechen (<i>hematologic improvement-erythroid</i>); IPSS-R = <i>International Prognostic Scoring System-revised</i>; N.E. = nicht erreicht oder nicht berechnet; PT = bevorzugter Begriff (<i>preferred term</i>); QoL-E = <i>Quality of Life-E</i>; SOC = Systemorganklasse (<i>system organ class</i>); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis; WHO = World Health Organization</p> <p>Quelle: [18]</p>													

Bezogen auf die Analysen zu den Endpunktkategorien Mortalität, Morbidität und Lebensqualität zeigten sich insgesamt neun positive Interaktionstests (siehe Tabelle 4-87):

- Reduktion der Transfusionslast (gemessen als HI-E):
  - Subgruppenmerkmal „Transfusionslast zur Baseline I“:  $p = 0,042$
- Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens – Fatigue:
  - Subgruppenmerkmal „Alter“:  $p = 0,026$
- krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens – körperliches Wohlbefinden:
  - Subgruppenmerkmal „Alter“:  $p = 0,003$
  - Subgruppenmerkmal „Region“:  $p = 0,015$
- krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens – funktionales Wohlbefinden:
  - Subgruppenmerkmal „Alter“:  $p = 0,022$
- krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens – soziales/familiäres Wohlbefinden:
  - Subgruppenmerkmal „SF3B1-Mutationsstatus“:  $p = 0,049$
- krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens – allgemeines Wohlbefinden:
  - Subgruppenmerkmal „Alter“:  $p = 0,008$
- krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens – Behandlungseffekt:
  - Subgruppenmerkmal „Alter“:  $p = 0,006$
- krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens – Gesamtsommenskala:
  - Subgruppenmerkmal „Alter“:  $p = 0,030$

Für die Endpunkte Gesamtmortalität, Transfusionsfreiheit, Reduktion der Transfusionslast gemessen als HI-E für Patienten mit hoher (HI-E I) und niedriger (HI-E II) Transfusionslast zur Baseline, Hospitalisierungen und gesundheitsbezogene Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens traten keine statistisch signifikanten Interaktions-p-Werte auf.

Bezogen auf die Analysen zur Gesamtrate von UE (jegliche UE, SUE, schwere UE, Therapieabbrüche aufgrund von UE) zeigten sich insgesamt vier positive Interaktionstests (siehe Tabelle 4-87):

- jegliche UE:
  - Subgruppenmerkmal „Dauer einer vorherigen ESA-Behandlung“:  $p = 0,047$
  - Subgruppenmerkmal „Serum-EPO-Spiegel“:  $p = 0,015$
- SUE:
  - Subgruppenmerkmal „Nierenfunktionsstatus, gemessen als Kreatinin-Clearance“:  $p = 0,039$
- Therapieabbrüche aufgrund von UE:
  - Subgruppenmerkmal „Thrombozytenzahl“:  $p = 0,017$

Für schwere UE (CTCAE  $\geq 3$ ) zeigten sich keine statistisch signifikanten Interaktions-p-Werte.

UE von besonderem Interesse traten nur sehr vereinzelt auf ( $< 10$  Ereignisse pro Behandlungsarm, siehe Tabelle 4-82), sodass aufgrund einer limitierten Interpretierbarkeit der Ergebnisse auf eine Darstellung von Subgruppenanalysen verzichtet wurde.

Entsprechend der aufgetretenen statistisch signifikanten Ergebnisse auf Ebene der SOC und PT bezogen auf die Gesamtpopulation wurden die Interaktionstests auf Ebene der SOC und PT für diese Ereignisse berechnet und dargestellt. Da nur für die Kategorie jegliche UE statistisch signifikante Effekte auf Ebene der SOC und PT aufgetreten sind, werden auch nur für diese die Interaktions-p-Werte dargestellt (siehe Tabelle 4-87).

Bezogen auf die Analysen zu Inzidenzen von jeglichen UE auf Ebene der SOC und PT, für die sich ein statistisch signifikantes Ergebnis in der Gesamtpopulation zeigte, lagen insgesamt sechs positive Interaktionstests:

- SOC „Gastrointestinale Erkrankungen“:
  - Subgruppenmerkmal „Nierenfunktionsstatus, gemessen als Kreatinin-Clearance“:  $p = 0,010$
- SOC „Erkrankungen des Nervensystems“:
  - Subgruppenmerkmal „Alter“:  $p = 0,034$
  - Subgruppenmerkmal „Zeit seit initialer MDS-Diagnose“:  $p = 0,005$

- PT „Asthenie“:
  - Subgruppenmerkmal „Dauer einer vorherigen ESA-Behandlung“:  $p = 0,032$
  - Subgruppenmerkmal „Nierenfunktionsstatus, gemessen als Kreatinin-Clearance“:  $p = 0,002$
- PT „Schwindelgefühl“:
  - Subgruppenmerkmal „Zeit seit initialer MDS-Diagnose“:  $p = 0,043$

*Stellen Sie schließlich alle Subgruppenergebnisse dar.*

*Sofern eine Effektmodifikation für mehr als ein Subgruppenmerkmal vorliegt, kann eine Untersuchung auf eine Wechselwirkung höherer Ordnung sinnvoll sein. Dies gilt insbesondere dann, wenn diese Effektmodifikation konsistent über mehrere Endpunkte besteht. Zur Interpretation der Ergebnisse sollte dann für diese Endpunkte zusätzlich eine Subgruppenanalyse durchgeführt werden, die die Merkmale mit Effektmodifikation kombiniert. Beispiel: Für die Endpunkte Mortalität, gesundheitsbezogene Lebensqualität und schwere unerwünschte Ereignisse liegt sowohl für das Merkmal Geschlecht (mit den Ausprägungen „weiblich“ und „männlich“) als auch für das Merkmal Schweregrad (mit den Ausprägungen „niedrig“ und „hoch“) eine Effektmodifikation vor. Die zusätzliche Subgruppenanalyse erfolgt dann für die 3 genannten Endpunkte für das kombinierte Merkmal Geschlecht/Schweregrad mit den 4 Ausprägungen weiblich/niedrig, weiblich/hoch, männlich/niedrig und männlich/hoch.*

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.*

*Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.*

Nachfolgend sind die Subgruppenanalysen, für die ein statistisch signifikanter Interaktionstest vorliegt, detailliert dargestellt und diskutiert.

**Subgruppenanalysen für den Endpunkt Morbidität – Reduktion der Transfusionslast**

Tabelle 4-88: Ergebnisse für die Subgruppenanalysen zum Endpunkt Morbidität – Reduktion der Transfusionslast aus MEDALIST; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Morbidität – Reduktion der Transfusionslast (gemessen als HI-E gemäß IWG 2006)						
Subgruppenmerkmal	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Transfusionslast zur Baseline I: p-Wert aus Interaktionstest 0,042</b>						
< 6 EK-Einheiten pro acht Wochen	45/87 (51,7)	2/43 (4,7)	21,96 [5,00; 96,52]	11,12 [2,83; 43,70]	47,07 [34,83; 59,32]	<0,001
≥ 6 EK-Einheiten pro acht Wochen	36/66 (54,5)	7/33 (21,2)	4,46 [1,70; 11,70]	2,57 [1,29; 5,14]	33,33 [14,93; 51,74]	0,002
Datenschnitt: 08.05.2018; Population: ITT-Population; <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(3)</sup> Der p-Wert basiert auf einem CMH-Test.  ARR = Absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran Mantel Haenszel; ; HI-E = erythroides Ansprechen ( <i>hematologic improvement-erythroid</i> ); ITT = Intention-to-Treat; IWG = <i>International Working Group</i> ; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko Quelle: [18]						

Für den Endpunkt Reduktion der Transfusionslast, gemessen als Anteil Patienten mit einem HI-E, zeigte sich für das Subgruppenmerkmal „Transfusionslast zur Baseline I“ ein statistisch signifikanter Interaktionstest ( $p = 0,042$ ). Eine statistisch signifikante Interaktion lag nur bei Kategorisierung der Patienten in die zwei Subgruppenmerkmale (< 6 vs. ≥ 6 EK-Einheiten pro acht Wochen), nicht jedoch in die drei Subgruppenmerkmale (< 4 vs. 4 bis < 6 vs. ≥ 6 EK-Einheiten pro acht Wochen) vor (siehe Tabelle 4-87). Die Effektschätzer zeigen über beide Subgruppen hinweg in die gleiche Richtung und ein identisches Ausmaß des Zusatznutzens an (siehe Tabelle 4-88). Auch traten keine positiven Interaktionstest hinsichtlich des Subgruppenmerkmals „Transfusionslast zur Baseline I“ über andere Endpunkte hinweg auf. Demnach handelt es sich selbst bei Vorliegen einer tatsächlichen Effektmodifikation nicht um eine bewertungsrelevante Interaktion und es findet keine abweichende Ableitung des Zusatznutzens statt.

### Subgruppenanalysen für den Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens

Tabelle 4-89: Ergebnisse für die Subgruppenanalysen zum Endpunkt Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens aus MEDALIST – Fatigue; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Morbidität – Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens - Fatigue						
Subgruppenmerkmal	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Alter: p-Wert aus Interaktionstest 0,028</b>						
≤ 64 Jahre	11/28 (39,3)	2/16 (12,5)	4,53 [0,86; 23,93]	3,14 [0,79; 12,44]	26,8 [2,5; 51,1]	0,075
65 – 74 Jahre	10/70 (14,3)	11/29 (37,9)	0,27 [0,10; 0,75]	0,38 [0,18; 0,79]	-23,6 [-43,1; -4,2]	0,011
≥ 75 Jahre	11/51 (21,6)	11/31 (35,5)	0,50 [0,19; 1,35]	0,61 [0,30; 1,23]	-13,9 [-34,2; 6,4]	0,171
Datenschnitt: 08.05.2018; Population: HRQoL-Population; <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(3)</sup> Der p-Wert basiert auf einem CMH-Test.  ARR = Absolute Risikodifferenz; BSC = Best-Supportive-Care; CMH = Cochran Mantel Haenszel; EORTC QLQ-C30 = European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module; HRQoL = gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> ); KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko Quelle: [19]						

Für das Symptom Fatigue des Endpunkts Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens zeigte sich für das Subgruppenmerkmal „Alter“ ein statistisch signifikanter Interaktionstest ( $p = 0,028$ ). Zwischen den Subgruppen liegen sowohl qualitative (entgegengesetzte Richtungen der Effektschätzer) als auch quantitative (statistische Signifikanz der Effekte) Unterschiede vor (siehe Tabelle 4-89). In der Subgruppe  $\leq 64$  Jahre sind nur wenig Patienten mit Ereignis enthalten (Luspatercept-Arm: 11; Placebo-Arm: 2), sodass der hier vorliegende qualitative Unterschied im Vergleich zu den beiden anderen Subgruppen hinsichtlich einer Interpretation stark limitiert ist. Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, ist nicht bekannt.

Somit ist davon auszugehen, dass ein falsch positives Ergebnis infolge des multiplen Testens vorliegt. Von einem relevanten Unterschied hinsichtlich des Subgruppenmerkmals „Alter“ ist nicht auszugehen und es findet dementsprechend keine abweichende Ableitung des Zusatznutzens statt.

### Subgruppenanalysen für den Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens

Tabelle 4-90: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – körperliches Wohlbefinden; Woche 1 – 24

Studie MEDALIST									
Endpunkt: Lebensqualität gemessen anhand des QoL-E-Fragebogens – körperliches Wohlbefinden									
Subgruppenmerkmal	Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC		
	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	LS-MWD [95 %-KI]	p-Wert <sup>(2)</sup>	SMD [95 %-KI] <sup>(3)</sup>
<b>Alter: p-Wert aus Interaktionstest 0,003</b>									
≤ 64 Jahre	28	57,14 (21,09)	55,77 (21,57)	16	56,25 (20,92)	52,27 (26,70)	2,79 [-6,99; 12,57]	0,575	0,13 [-0,47; 0,73]
65 – 74 Jahre	70	55,25 (20,94)	56,10 (21,71)	29	51,29 (19,00)	57,95 (19,88)	-12,76 [-20,58; -4,94]	0,002	-0,62 [-1,06; -0,18]
≥ 75 Jahre	51	49,49 (19,76)	51,52 (24,56)	31	49,19 (27,18)	51,25 (26,87)	-0,40 [-8,41; 7,61]	0,922	-0,02 [-0,46; 0,43]
<b>Region: p-Wert aus Interaktionstest 0,015</b>									
Nordamerika	31	57,66 (17,58)	64,22 (21,84)	19	56,58 (19,26)	53,33 (17,34)	3,69 [-5,46; 12,84]	0,428	0,20 [-0,36; 0,76]
Europa	118	52,61 (21,36)	50,77 (22,04)	57	49,78 (23,92)	54,61 (26,22)	-7,64 [-14,17; -1,10]	0,022	-0,34 [-0,66; -0,02]
Datenschnitt: 08.05.2018; Population: HRQoL-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktions-term als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte <sup>(3)</sup> SMD in Form von Hedges' g inkl. 95 %-KI  BSC = Best-Supportive-Care; HRQoL = gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> ); IPSS-R = <i>International Prognostic Scoring System-revised</i> ; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MMRM = gemischtes Modell für wiederholte Messungen ( <i>mixed model for repeated measurements</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; QoL-E = <i>Quality of Life-E</i> ; SD = Standardabweichung; SMD = standardisierte Mittelwertdifferenz Quelle: [19]									

Für die Domäne körperliches Wohlbefinden des Endpunkts krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens zeigten sich für die Subgruppenmerkmale „Alter“ und „Region“ statistisch signifikante Interaktionstests ( $p = 0,003$  und  $p = 0,015$ ). Innerhalb der Merkmale vorliegende qualitative Unterschiede zwischen den Subgruppen (entgegengesetzte Richtungen der Effektschätzer) sind in ihrem Ausmaß nicht klinisch relevant und führen somit nicht zu einer unterschiedlichen Bewertung der Ergebnisse der einzelnen Subgruppen (siehe Tabelle 4-90).

Tabelle 4-91: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – funktionales Wohlbefinden; Woche 1 – 24

Studie MEDALIST									
Endpunkt: Lebensqualität gemessen anhand des QoL-E-Fragebogens – funktionales Wohlbefinden									
Subgruppenmerkmal	Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC		
	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	LS-MWD [95 %-KI]	p-Wert <sup>(2)</sup>	SMD [95 %-KI] <sup>(3)</sup>
<b>Alter: p-Wert aus Interaktionstest 0,022</b>									
≤ 64 Jahre	28	58,33 (31,66)	62,82 (31,10)	16	50,00 (31,69)	54,55 (35,99)	11,20 [-4,64; 27,04]	0,165	0,35 [-0,26; 0,95]
65 – 74 Jahre	70	55,72 (33,25)	50,65 (32,92)	29	47,89 (28,96)	59,60 (32,08)	-14,26 [-26,41; -2,10]	0,022	-0,44 [-0,88; -0,01]
≥ 75 Jahre	51	52,61 (35,84)	52,42 (33,33)	31	54,48 (29,59)	66,11 (31,73)	-7,41 [-19,88; 5,07]	0,243	-0,22 [-0,67; 0,23]
Datenschnitt: 08.05.2018; Population: HRQoL-Population									
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).									
<sup>(2)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktions-term als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.									
<sup>(3)</sup> SMD in Form von Hedges' g inkl. 95 %-KI.									
BSC = Best-Supportive-Care; HRQoL = gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> ); IPSS-R = <i>International Prognostic Scoring System-revised</i> ; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MMRM = gemischtes Modell für wiederholte Messungen ( <i>mixed model for repeated measurements</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; QoL-E = <i>Quality of Life-E</i> ; SD = Standardabweichung; SMD = standardisierte Mittelwertdifferenz Quelle: [19]									

Für die Domäne funktionales Wohlbefinden des Endpunkts krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens zeigte sich für das Subgruppenmerkmal „Alter“ ein statistisch signifikanter Interaktionstest ( $p = 0,022$ ). Die Effektschätzer sind teils entgegengerichtet, klinisch relevante Unterschiede, die zu unterschiedlichen Bewertungen der Ergebnisse zwischen den Subgruppen führen würden, liegen jedoch nicht vor (siehe Tabelle 4-91).

Tabelle 4-92: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – soziales/familiäres Wohlbefinden; Woche 1 – 24

<b>Studie MEDALIST</b>									
<b>Endpunkt: Lebensqualität gemessen anhand des QoL-E-Fragebogens – soziales/familiäres Wohlbefinden</b>									
Subgruppenmerkmal	Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC		
	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	LS-MWD [95 %-KI]	p-Wert <sup>(2)</sup>	SMD [95 %-KI] <sup>(3)</sup>
<b>SF3B1-Mutationsstatus: p-Wert aus Interaktionstest 0,049</b>									
mutiert	138	47,83 (37,14)	48,92 (34,95)	65	46,98 (38,35)	54,88 (38,00)	-5,99 [-15,07; 3,10]	0,195	-0,16 [-0,47; 0,15]
nicht mutiert	11	62,50 (44,49)	38,89 (33,33)	10	44,44 (32,54)	37,50 (30,62)	-26,35 [-46,51; -6,20]	0,011	-0,64 [-1,52; 0,24]
Datenschnitt: 08.05.2018; Population: HRQoL-Population									
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).									
<sup>(2)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktions-term als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.									
<sup>(3)</sup> SMD in Form von Hedges' g inkl. 95 %-KI.									
BSC = Best-Supportive-Care; HRQoL = gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> ); IPSS-R = <i>International Prognostic Scoring System-revised</i> ; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MMRM = gemischtes Modell für wiederholte Messungen ( <i>mixed model for repeated measurements</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; QoL-E = <i>Quality of Life-E</i> ; SD = Standardabweichung; SMD = standardisierte Mittelwertdifferenz Quelle: [19]									

Für die Domäne soziales/familiäres Wohlbefinden des Endpunkts krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens zeigte sich für das Subgruppenmerkmal „SF3B1-Mutationsstatus“ ein statistisch signifikanter Interaktionstest ( $p = 0,049$ ). Qualitative Unterschiede zwischen den beiden Subgruppen liegen nicht vor (Effektschätzer sind gleichgerichtet). Auch wenn nur für die Subgruppe ohne SF3B1-Mutation ein statistisch signifikanter Unterschied vorliegt, ist dieser nicht klinisch relevant, sodass sich zwischen den Subgruppen keine unterschiedlichen Aussagen zum Zusatznutzen ableiten lassen (siehe Tabelle

4-92). Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, ist nicht bekannt. Positive Interaktionstests hinsichtlich des Subgruppenmerkmals „SF3B1-Mutationsstatus“ lagen darüber hinaus bei keinem anderen betrachteten Endpunkt vor. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals SF3B1-Mutationsstatus auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

Tabelle 4-93: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – allgemeines Wohlbefinden; Woche 1 – 24

Studie MEDALIST									
Endpunkt: Lebensqualität gemessen anhand des QoL-E-Fragebogens – allgemeines Wohlbefinden									
Subgruppenmerkmal	Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC		
	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	LS-MWD [95 %-KI]	p-Wert <sup>(2)</sup>	SMD [95 %-KI] <sup>(3)</sup>
<b>Alter: p-Wert aus Interaktionstest 0,008</b>									
≤ 64 Jahre	28	61,94 (21,97)	63,27 (20,74)	16	57,50 (22,86)	58,49 (22,23)	2,96 [-5,73; 11,64]	0,503	0,13 [-0,52; 0,78]
65 – 74 Jahre	70	58,82 (19,26)	56,82 (18,94)	29	55,23 (19,14)	59,20 (21,03)	-12,18 [-18,73; -5,64]	< 0,001	-0,63 [-1,09; -0,17]
≥ 75 Jahre	51	58,70 (23,68)	59,22 (21,68)	31	59,32 (21,74)	68,00 (20,84)	-4,13 [-10,83; 2,57]	0,226	-0,18 [-0,64; 0,29]
Datenschnitt: 08.05.2018; Population: HRQoL-Population									
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).									
<sup>(2)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktions-term als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.									
<sup>(3)</sup> SMD in Form von Hedges' g inkl. 95 %-KI.									
BSC = Best-Supportive-Care; HRQoL = gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> ); IPSS-R = <i>International Prognostic Scoring System-revised</i> ; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MMRM = gemischtes Modell für wiederholte Messungen ( <i>mixed model for repeated measurements</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; QoL-E = <i>Quality of Life-E</i> ; SD = Standardabweichung; SMD = standardisierte Mittelwertdifferenz Quelle: [19]									

Für den Summenscore allgemeines Wohlbefinden des Endpunkts krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens zeigte sich für das Subgruppenmerkmal „Alter“ ein statistisch signifikanter Interaktionstest ( $p = 0,008$ ). Die Effektschätzer sind teils entgegengerichtet, klinisch relevante Unterschiede zwischen den Subgruppen, die eine unterschiedliche Bewertung der Ergebnisse zulassen würden, liegen jedoch nicht vor (siehe

Tabelle 4-93). Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, ist nicht bekannt. Demnach handelt es sich selbst bei Vorliegen einer tatsächlichen Effektmodifikation nicht um eine bewertungsrelevante Interaktion und es findet keine abweichende Ableitung des Zusatznutzens statt.

Tabelle 4-94: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Behandlungseffekt; Woche 1 – 24

Studie MEDALIST									
Endpunkt: Lebensqualität gemessen anhand des QoL-E-Fragebogens – Behandlungseffekt									
Subgruppenmerkmal	Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC		
	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	LS-MWD [95 %-KI]	p-Wert <sup>(2)</sup>	SMD [95 %-KI] <sup>(3)</sup>
<i>Alter: p-Wert aus Interaktionstest 0,006</i>									
≤ 64 Jahre	28	58,94 (21,19)	60,29 (22,81)	16	54,41 (20,38)	55,09 (22,47)	4,56 [-4,82; 13,94]	0,339	0,21 [-0,39; 0,82]
65 – 74 Jahre	70	56,47 (19,79)	54,58 (21,99)	29	52,05 (18,10)	57,83 (22,02)	-11,72 [-18,94; -4,51]	0,002	-0,60 [-1,04; -0,16]
≥ 75 Jahre	51	51,98 (22,14)	56,57 (22,89)	31	54,15 (22,39)	59,91 (22,20)	-1,69 [-9,15; 5,77]	0,656	-0,08 [-0,53; 0,37]
Datenschnitt: 08.05.2018; Population: HRQoL-Population									
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).									
<sup>(2)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktions-term als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte.									
<sup>(3)</sup> SMD in Form von Hedges' g inkl. 95 %-KI.									
BSC = Best-Supportive-Care; HRQoL = gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> ); IPSS-R = <i>International Prognostic Scoring System-revised</i> ; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MMRM = gemischtes Modell für wiederholte Messungen ( <i>mixed model for repeated measurements</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; QoL-E = <i>Quality of Life-E</i> ; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz									
Quelle: [19]									

Für den Summenscore Behandlungseffekt des Endpunkts krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens zeigte sich für das Subgruppenmerkmal „Alter“ ein statistisch signifikanter Interaktionstest ( $p = 0,006$ ). Die Effektschätzer sind teils entgegengerichtet, klinisch relevante Unterschiede zwischen den Subgruppen, die eine unterschiedliche Bewertung der Ergebnisse der einzelnen Subgruppen zulassen würden, liegen jedoch nicht vor (siehe Tabelle 4-94). Eine medizinische Rationale, die unterschiedliche Effekte vermuten

lassen würde, ist nicht bekannt. Demnach handelt es sich selbst bei Vorliegen einer tatsächlichen Effektmodifikation nicht um eine bewertungsrelevante Interaktion und es findet keine abweichende Ableitung des Zusatznutzens statt.

Tabelle 4-95: Ergebnisse für die Subgruppenanalysen zum Endpunkt Lebensqualität gemessen anhand des QoL-E-Fragebogens aus MEDALIST – Gesamtsummenscore; Woche 1 – 24

Studie MEDALIST									
Endpunkt: Lebensqualität gemessen anhand des QoL-E-Fragebogens – Gesamtsummenscore									
Subgruppenmerkmal	Luspatercept + BSC <sup>(1)</sup>			Placebo + BSC <sup>(1)</sup>			Luspatercept + BSC vs. Placebo + BSC		
	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	N	MW (SD) zu Studienbeginn	MW (SD) zu Woche 24	LS-MWD [95 %-KI]	p-Wert <sup>(2)</sup>	SMD [95 %-KI] <sup>(3)</sup>
<i>Alter: p-Wert aus Interaktionstest 0,030</i>									
≤ 64 Jahre	28	59,74 (23,73)	63,23 (21,75)	16	56,98 (23,42)	59,48 (24,07)	1,50 [-7,42; 10,41]	0,741	0,06 [-0,58; 0,71]
65 – 74 Jahre	70	58,89 (19,04)	56,15 (21,50)	29	56,83 (19,01)	58,23 (22,39)	-10,44 [-17,24; -3,64]	0,003	-0,54 [-1,00; -0,08]
≥ 75 Jahre	51	56,35 (23,00)	60,93 (21,86)	31	58,96 (22,50)	65,81 (20,15)	-1,79 [-8,83; 5,25]	0,616	-0,08 [-0,55; 0,29]
Datenschnitt: 08.05.2018; Population: HRQoL-Population <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> MMRM-Modell mit Änderungen zur Baseline als abhängige Variable, Baseline-Werte, Behandlungsarm, Stratifizierungsfaktoren (Baseline-Transfusionslast und Baseline-IPSS-R-Score) und Zeit als Kovariate, ein Baseline-Wert-Zeit-Interaktionsterm und Behandlungsarm-Zeit-Interaktions-term als fester Effekt und Achsenabschnitt und Zeit als zufällige Effekte. <sup>(3)</sup> SMD in Form von Hedges' g inkl. 95 %-KI.  BSC = Best-Supportive-Care; HRQoL = gesundheitsbezogene Lebensqualität ( <i>health-related quality of life</i> ); IPSS-R = <i>International Prognostic Scoring System-revised</i> ; LS = Methode der kleinsten Quadrate ( <i>least square</i> ); MMRM = gemischtes Modell für wiederholte Messungen ( <i>mixed model for repeated measurements</i> ); MW = Mittelwert; MWD = Mittelwertdifferenz; QoL-E = <i>Quality of Life-E</i> ; SD = Standardabweichung; SE = Standardfehler; SMD = standardisierte Mittelwertdifferenz Quelle: [19]									

Für den Gesamtsummenscore des Endpunkts krankheitsspezifische Lebensqualität gemessen anhand des QoL-E-Fragebogens zeigte sich für das Subgruppenmerkmal „Alter“ ein statistisch signifikanter Interaktionstest ( $p = 0,030$ ). Die Effektschätzer sind teils entgegengerichtet, klinisch relevante Unterschiede zwischen den Subgruppen, die eine unterschiedliche Bewertung der Ergebnisse der einzelnen Subgruppen zulassen würden, liegen jedoch nicht vor (siehe Tabelle 4-95). Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, ist nicht bekannt. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich

des Subgruppenmerkmals Alter auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

### Subgruppenanalysen für den Endpunkt Sicherheit – jegliche UE

Tabelle 4-96: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE aus MEDALIST; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Sicherheit – jegliche UE						
Endpunkt	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Serum-EPO-Spiegel: p-Wert aus Interaktionstest 0,015</b>						
< 100 U/l	46/51 (90,2)	31/31 (100,0)	0,13 [0,01; 2,51]	0,90 [0,82; 0,99]	-9,80 [-17,97; -1,64]	0,074
100 – < 200 U/l	37/37 (100,0)	18/19 (94,7)	6,08 [0,24; 156,63]	1,06 [0,95; 1,17]	5,26 [-4,78; 15,30]	0,163
200 – 500 U/l	41/43 (95,3)	12/15 (80,0)	5,12 [0,77; 34,31]	1,19 [0,92; 1,55]	15,35 [-5,85; 36,55]	0,071
> 500 U/l	20/21 (95,2)	9/11 (81,8)	4,44 [0,36; 55,58]	1,16 [0,87; 1,56]	13,42 [-11,13; 37,97]	0,223
<b>Längste Dauer einer vorherigen ESA-Behandlung: p-Wert aus Interaktionstest 0,047</b>						
< 6 Monate	48/48 (100,0)	18/20 (90,0)	13,11 [0,60; 286,14]	1,11 [0,96; 1,29]	10,00 [-3,15; 23,15]	0,027
6 – 12 Monate	30/34 (88,2)	9/9 (100,0)	0,36 [0,02; 7,25]	0,88 [0,78; 1,00]	-11,76 [-22,59; -0,93]	0,286
> 12 – 24 Monate	34/35 (97,1)	19/21 (90,5)	3,58 [0,30; 42,11]	1,07 [0,92; 1,25]	6,67 [-7,05; 20,38]	0,288
> 24 Monate	28/31 (90,3)	19/20 (95,0)	0,49 [0,05; 5,08]	0,95 [0,82; 1,11]	-4,68 [-18,80; 9,45]	0,548
Datenschnitt: 08.05.2018; Population: Safety-Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.  ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; EPO = Erythropoetin; ESA = Erythropoese-stimulierende Agenzien ( <i>erythropoiesis-stimulating agents</i> ); KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; UE = unerwünschtes Ereignis Quelle: [18]						

Für den Endpunkt jegliche UE zeigte sich für die Subgruppenmerkmale „längste Dauer einer vorherigen ESA-Behandlung“ und „Serum-EPO-Spiegel“ ein statistisch signifikanter Interaktionstest ( $p = 0,047$  bzw.  $p = 0,015$ ).

Für das Subgruppenmerkmal „längste Dauer einer vorherigen ESA-Behandlung“ zeigte sich mit Ausnahme der Subgruppe < 6 Monate für alle Subgruppen ein quantitativ nicht voneinander abweichendes Ergebnis. Für die Subgruppe < 6 Monate lag ein statistisch signifikanter Unterschied zuungunsten von Luspatercept vor, der jedoch in seinem Ausmaß als geringfügig einzustufen ist (RR = 1,11; 95 %-KI: [0,96; 1,29]; siehe Tabelle 4-96). Darüber hinaus werden in der Kategorie jegliche UE auch solche Ereignisse subsummiert, die als Laborwerte einen Surrogatcharakter haben und somit nicht per se als patientenrelevant einzustufen sind. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals „längste Dauer einer vorherigen ESA-Behandlung“ auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

Für das Subgruppenmerkmal „Serum-EPO-Spiegel“ zeigte sich für alle Subgruppen ein quantitativ nicht voneinander abweichendes Ergebnis, sodass eine unterschiedliche Bewertung der Ergebnisse nicht vorgenommen wird (siehe Tabelle 4-96). Darüber hinaus werden in der Kategorie jegliche UE auch solche Ereignisse subsummiert, die als Laborwerte einen Surrogatcharakter haben und somit nicht per se als patientenrelevant einzustufen sind. Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, ist nicht bekannt. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals „Serum-EPO-Spiegel“ auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

**Subgruppenanalysen für den Endpunkt Sicherheit – SUE**

Tabelle 4-97: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – SUE aus MEDALIST; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Sicherheit – SUE						
Endpunkt	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Nierenfunktionsstatus (Kreatinin-Clearance): p-Wert aus Interaktionstest 0,039</b>						
40 – < 60 ml/min	12/33 (36,4)	2/12 (16,7)	2,86 [0,53; 15,26]	2,18 [0,57; 8,36]	19,70 [-7,02; 46,42]	0,212
≥ 60 ml/min	26/113 (23,0)	13/62 (21,0)	1,13 [0,53; 2,39]	1,10 [0,61; 1,98]	2,04 [-10,72; 14,80]	0,757
< 40 ml/min	2/7 (28,6)	1/1 (100,0)	0,15 [0,00; 5,18]	0,29 [0,09; 0,92]	-71,43 [-100,00; -37,96]	0,197
Datenschnitt: 08.05.2018; Population: Safety-Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.  ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis Quelle: [18]						

Für den Endpunkt SUE zeigte sich für das Subgruppenmerkmal „Nierenfunktionsstatus, gemessen als Kreatinin-Clearance“, ein statistisch signifikanter Interaktionstest ( $p = 0,039$ ).

Für alle Subgruppen zeigte sich ein quantitativ nicht voneinander abweichendes Ereignis, sodass eine unterschiedliche Bewertung der Ergebnisse nicht vorgenommen wird (siehe Tabelle 4-97). Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, ist nicht bekannt. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals „Nierenfunktionsstatus“ auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

## Subgruppenanalysen für den Endpunkt Sicherheit – Therapieabbrüche aufgrund von UE

Tabelle 4-98: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – Therapieabbrüche aufgrund von UE aus MEDALIST; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Sicherheit – Therapieabbrüche aufgrund von UE						
Endpunkt	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Thrombozytenzahl: p-Wert aus Interaktionstest 0,017</b>						
< 100×10 <sup>9</sup> /l	0/8 (0,0)	0/6 (0,0)	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E. [N.E.; N.E.]	N.E.
100 – 400×10 <sup>9</sup> /l	12/128 (9,4)	1/61 (1,6)	6,21 [0,79; 48,88]	5,72 [0,76; 42,98]	7,74 [1,76; 13,71]	0,050
> 400×10 <sup>9</sup> /l	0/17 (0,0)	3/9 (33,3)	0,05 [0,00; 1,17]	0,08 [0,00; 1,39]	-33,33 [-64,13; -2,54]	0,013
Datenschnitt: 08.05.2018; Population: Safety-Population.						
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).						
<sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet.						
<sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.						
ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; UE = unerwünschtes Ereignis						
Quelle: [18]						

Für den Endpunkt Therapieabbrüche aufgrund von UE zeigten sich für das Subgruppenmerkmal „Thrombozytenzahl“ ein statistisch signifikanter Interaktionstest ( $p = 0,017$ ). Für das betrachtete Subgruppenmerkmale traten in den jeweiligen Subgruppen nur sehr vereinzelt Ereignisse auf, sodass die Interpretierbarkeit der Ergebnisse limitiert ist (siehe Tabelle 4-98). Darüber hinaus liegt für das Subgruppenmerkmal „Thrombozytenzahl“ keine medizinische Rationale vor, welche unterschiedliche Effekte vermuten lassen würde. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals „Thrombozytenzahl“ auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

## Subgruppenanalysen für den Endpunkt Sicherheit – jegliche UE auf Ebene der SOC und PT

Tabelle 4-99: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE: SOC Gastrointestinale Erkrankungen aus MEDALIST; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Sicherheit – jegliche UE: SOC Gastrointestinale Erkrankungen						
Endpunkt	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Nierenfunktionsstatus (Kreatinin-Clearance): p-Wert aus Interaktionstest 0,010</b>						
40 – < 60 ml/min	16/33 (48,5)	3/12 (25,0)	2,82 [0,65; 12,33]	1,94 [0,68; 5,49]	23,48 [-6,36; 53,33]	0,163
≥ 60 ml/min	58/113 (51,3)	21/62 (33,9)	2,06 [1,08; 3,91]	1,52 [1,02; 2,24]	17,46 [2,50; 32,41]	0,027
< 40 ml/min	4/7 (57,1)	1/1 (100,0)	0,43 [0,01; 14,08]	0,57 [0,30; 1,09]	-42,86 [-79,52; -6,20]	0,439
Datenschnitt: 08.05.2018; Population: Safety-Population.						
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).						
<sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet.						
<sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.						
ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; SOC = Systemorganklasse ( <i>system organ class</i> ); UE = unerwünschtes Ereignis						
Quelle: [18]						

Für die SOC „Gastrointestinale Erkrankungen“ des Endpunkts jegliche UE zeigte sich für das Subgruppenmerkmal „Nierenfunktionsstatus, gemessen als Kreatinin-Clearance“, ein statistisch signifikanter Interaktionstest ( $p = 0,010$ ). In der Subgruppe „< 40 ml/min“ traten nur sehr wenig (< 10) Ereignisse auf, sodass die in dieser Subgruppe beobachtete qualitative Abweichung von den anderen beiden Subgruppen nur begrenzt interpretierbar ist. Zwischen den anderen beiden Subgruppen lag ein quantitativer, im Ausmaß jedoch als geringfügig einzu-stufender Unterschied vor, sodass keine abweichende Bewertung der Ergebnisse vorgenommen wird (siehe Tabelle 4-99). Darüber hinaus liegt keine medizinische Rationale vor, die unterschiedliche Effekte vermuten lassen würde. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals „Nierenfunktionsstatus“ auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

Tabelle 4-100: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE: SOC Erkrankungen des Nervensystems aus MEDALIST; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Sicherheit – jegliche UE: SOC Erkrankungen des Nervensystems						
Endpunkt	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Alter: p-Wert aus Interaktionstest 0,034</b>						
≤ 64 Jahre	6/29 (20,7)	6/16 (37,5)	0,43 [0,11; 1,68]	0,55 [0,21; 1,43]	-16,81 [-44,74; 11,12]	0,227
65 – 74 Jahre	33/72 (45,8)	5/29 (17,2)	4,06 [1,39; 11,83]	2,66 [1,15; 6,13]	28,59 [10,66; 46,52]	0,008
≥ 75 Jahre	21/52 (40,4)	6/31 (19,4)	2,82 [0,99; 8,06]	2,09 [0,95; 4,60]	21,03 [1,76; 40,30]	0,049
<b>Zeit seit initialer MDS-Diagnose: p-Wert aus Interaktionstest 0,005</b>						
≤ 2 Jahre	17/40 (42,5)	3/19 (15,8)	3,94 [0,99; 15,72]	2,69 [0,90; 8,08]	26,71 [4,27; 49,15]	0,045
> 2 – 5 Jahre	20/62 (32,3)	2/34 (5,9)	7,62 [1,66; 35,00]	5,48 [1,36; 22,06]	26,38 [12,31; 40,45]	0,003
> 5 Jahre	23/51 (45,1)	12/23 (52,2)	0,75 [0,28; 2,02]	0,86 [0,53; 1,42]	-7,08 [-31,64; 17,49]	0,575
Datenschnitt: 08.05.2018; Population: Safety-Population.						
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelatherapie (siehe auch Tabelle 4-13).						
<sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet.						
<sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.						
ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; RR = Relatives Risiko; SOC = Systemorganklasse ( <i>system organ class</i> ); UE = unerwünschtes Ereignis						
Quelle: [18]						

Für die SOC „Erkrankungen des Nervensystems“ des Endpunkts jegliche UE zeigten sich für die Subgruppenmerkmale „Alter“ und „Zeit seit initialer MDS-Diagnose“ statistisch signifikante Interaktionstests ( $p = 0,034$  und  $p = 0,005$ ).

Für das Subgruppenmerkmal „Alter“ zeigten sich teils qualitative und quantitative Unterschiede zwischen den Subgruppen, von denen lediglich für die Subgruppe „65 – 74 Jahre“ eine abweichende Bewertung der Ergebnisse gerechtfertigt wäre (siehe Tabelle 4-100). Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, liegt jedoch nicht vor. Zwar liegen hinsichtlich des Subgruppenmerkmals „Alter“ auch bei anderen Endpunkten positive Interaktionstests vor, jedoch führend diese auch bei den anderen Endpunkten nicht zu einer abweichenden Ableitung des Zusatznutzens. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals „Alter“ auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

Für das Subgruppenmerkmal „Zeit seit initialer MDS-Diagnose“ zeigten sich teils qualitative und quantitative Unterschiede zwischen den Subgruppen, von denen lediglich für die Subgruppe „> 2 – 5 Jahre“ eine abweichende Bewertung der Ergebnisse gerechtfertigt wäre (siehe Tabelle 4-100). Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, ist nicht bekannt. Für das Subgruppenmerkmal „Zeit seit initialer MDS-Diagnose“ liegen nur bei zwei betrachteten Endpunkten positive Interaktionstests vor. Insgesamt ist somit von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals Zeit seit initialer MDS-Diagnose auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

Tabelle 4-101: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE: PT Asthenie aus MEDALIST; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Sicherheit – jegliche UE: PT Asthenie						
Endpunkt	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Dauer einer vorherigen ESA-Behandlung: p-Wert aus Interaktionstest 0,032</b>						
< 6 Monate	6/48 (12,5)	0/20 (0,0)	6,27 [0,34; 116,77]	5,57 [0,33; 94,47]	12,50 [3,14; 21,86]	0,100
6 – 12 Monate	8/34 (23,5)	4/9 (44,4)	0,38 [0,08; 1,79]	0,53 [0,20; 1,37]	-20,92 [-56,37; 14,54]	0,219
> 12 – 24 Monate	9/35 (25,7)	1/21 (4,8)	6,92 [0,81; 59,23]	5,40 [0,74; 39,66]	20,95 [3,85; 38,06]	0,050
> 24 Monate	5/31 (16,1)	1/20 (5,0)	3,65 [0,39; 33,88]	3,23 [0,41; 25,62]	11,13 [-4,96; 27,22]	0,233
<b>Nierenfunktionsstatus (Kreatinin-Clearance): p-Wert aus Interaktionstest 0,002</b>						
< 40 ml/min	5/33 (15,2)	0/12 (0,0)	4,82 [0,25; 94,08]	4,21 [0,25; 70,82]	15,15 [2,92; 27,38]	0,157
40 – < 60 ml/min	20/113 (17,7)	5/62 (8,1)	2,45 [0,87; 6,89]	2,19 [0,87; 5,56]	9,63 [-0,14; 19,40]	0,082
≥ 60 ml/min	3/7 (42,9)	1/1 (100,0)	0,26 [0,01; 8,52]	0,43 [0,18; 1,01]	-57,14 [-93,80; -20,48]	0,317
Datenschnitt: 08.05.2018; Population: Safety-Population.						
<sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13).						
<sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet.						
<sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.						
ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; PT = bevorzugter Begriff ( <i>preferred term</i> ); RR = Relatives Risiko; UE = unerwünschtes Ereignis						
Quelle: [18]						

Für den PT „Asthenie“ des Endpunkts jegliche UE zeigten sich für die Subgruppenmerkmale „Dauer einer vorherigen ESA-Behandlung“ und „Nierenfunktionsstatus, gemessen als Kreatinin-Clearance“, statistisch signifikante Interaktionstests ( $p = 0,032$  und  $p = 0,002$ ).

Für das Subgruppenmerkmal „Dauer einer vorherigen ESA-Behandlung“ traten in allen Subgruppen nur vereinzelt Ereignisse auf, sodass eine sinnvolle Interpretation der Ergebnisse starken Limitationen unterliegt (siehe Tabelle 4-101). Qualitative Unterschiede zwischen den Subgruppen liegen nur für eine Subgruppe vor, dieser Effekt führt jedoch nicht zu einer quantitativ unterschiedlichen Bewertung der Ergebnisse.

Für das Subgruppenmerkmal „Nierenfunktionsstatus, gemessen als Kreatinin-Clearance“, zeigten sich nur für die Subgruppe „< 40 ml/min“ entgegengesetzte Richtungen der Effektschätzer. Da diese Subgruppe jedoch nur sehr geringe Patienten- und Ereigniszahlen aufweist, ist die Interpretierbarkeit der Ergebnisse stark limitiert. Das Ausmaß des Effekts lässt zudem keine abweichende Bewertung des Effekts im Vergleich zu den anderen Subgruppen zu. Darüber hinaus liegt keine medizinische Rationale vor, die unterschiedliche Effekte vermuten lassen würde. Dementsprechend ist von keinem relevanten Unterschied hinsichtlich des Subgruppenmerkmals „Nierenfunktionsstatus“ auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

Tabelle 4-102: Ergebnisse für die Subgruppenanalysen zum Endpunkt Sicherheit – jegliche UE: PT Schwindelgefühl aus MEDALIST; Woche 1 – 24

Studie MEDALIST						
Endpunkt: Sicherheit – jegliche UE: PT Schwindelgefühl						
Endpunkt	Luspatercept + BSC <sup>(1)</sup>	Placebo + BSC <sup>(1)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(1)</sup>			
	Patienten mit Ereignis n/N (%)	Patienten mit Ereignis n/N (%)	OR [95 %-KI] <sup>(2)</sup>	RR [95 %-KI] <sup>(2)</sup>	ARR [95 %-KI] <sup>(2)</sup>	p-Wert <sup>(3)</sup>
<b>Zeit seit initialer MDS-Diagnose: p-Wert aus Interaktionstest 0,043</b>						
≤ 2 Jahre	9/40 (22,5)	0/19 (0,0)	11,76 [0,65; 213,62]	9,27 [0,57; 151,37]	22,50 [9,56; 35,44]	0,026
> 2 – 5 Jahre	8/62 (12,9)	0/34 (0,0)	10,76 [0,60; 192,45]	9,44 [0,56; 158,78]	12,90 [4,56; 21,25]	0,030
> 5 Jahre	6/51 (11,8)	4/23 (17,4)	0,63 [0,16; 2,50]	0,68 [0,21; 2,17]	-5,63 [-23,46; 12,21]	0,515
Datenschnitt: 08.05.2018; Population: Safety-Population. <sup>(1)</sup> BSC umfasste insbesondere die regelmäßige, bedarfsgerechte Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit einer Eisenchelattherapie (siehe auch Tabelle 4-13). <sup>(2)</sup> OR, RR und ARR wurden unstratifiziert mittels Normal-Approximation berechnet. <sup>(3)</sup> Der p-Wert basiert auf einem stratifizierten CMH-Test.  ARR = Absolute Risikodifferenz; CMH = Cochran Mantel Haenszel; KI = Konfidenzintervall; OR = Odds Ratio; PT = bevorzugter Begriff ( <i>preferred term</i> ); RR = Relatives Risiko; UE = unerwünschtes Ereignis Quelle: [18]						

Für den PT „Schwindelgefühl“ des Endpunkts jegliche UE zeigte sich für das Subgruppenmerkmal „Zeit seit initialer MDS-Diagnose“ ein statistisch signifikanter Interaktionstest ( $p = 0,043$ ). Insgesamt traten in den Subgruppen nur sehr vereinzelt Ereignisse auf, sodass die Interpretierbarkeit der Ergebnisse limitiert ist (siehe Tabelle 4-102). Auch liegen keine Effekte vor, die eine unterschiedliche Bewertung der Ergebnisse zwischen den Subgruppen erlauben würden. Eine medizinische Rationale, die unterschiedliche Effekte vermuten lassen würde, ist nicht bekannt. Für das Subgruppenmerkmal „Zeit seit initialer MDS-Diagnose“ liegen nur bei zwei betrachteten Endpunkten positive Interaktionstests vor. Insgesamt ist von keinem

relevanten Unterschied hinsichtlich des Subgruppenmerkmals „Zeit seit initialer MDS-Diagnose“ auszugehen und es findet keine abweichende Ableitung des Zusatznutzens statt.

### **Zusammenfassung der Ergebnisse zu Subgruppenanalysen**

In MEDALIST wurden eine Reihe an potentiellen Effektmodifikationen untersucht. Insgesamt traten nur vereinzelt positive Interaktionstests (Interaktions-p-Wert < 0,05) auf, sodass keine abweichende Bewertung für bestimmte Subgruppen sachgerecht und sinnvoll wäre.

#### **4.3.1.4 Liste der eingeschlossenen Studien - RCT**

*Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).*

##### **Studie MEDALIST**

- Publikationen [79]
- Studienberichte [1]
- Studienregistereinträge [74-78]
- Post-hoc-Auswertungen durch Celgene für das vorliegende Dossier
  - Abschnitt 4.3.1.2
  - Abschnitt 4.3.1.3
  - Anhang 4-G
  - Anhang 4-H

## 4.3.2 Weitere Unterlagen

### 4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

#### 4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der Internetseite des G-BA*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

#### 4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

*Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche identifiziert wurden und bewerten Sie darüber hinaus deren Ähnlichkeit. Begründen Sie darauf basierend den Ein- bzw. Ausschluss von Studien für die von Ihnen durchgeführten indirekten Vergleiche. Bewerten Sie das Verzerrungspotenzial der für indirekte Vergleiche herangezogenen Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.*

Nicht zutreffend.

#### 4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-103: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

##### 4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in drei Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-104: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Intervention	<Vergleichs- therapie 1>	<Vergleichs- therapie 2>	<Vergleichs- therapie 3>
1	<Studie 1>	•		•	•
2	<Studie 2>	•		•	
	<Studie 3>	•		•	
1	<Studie 4>		•	•	•
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-105: Operationalisierung von &lt;Endpunkt xxx&gt;

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-106: Bewertung des Verzerrungspotenzials für &lt;Endpunkt xxx&gt; in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

*Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

Tabelle 4-107: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)
<Studie 1>	

Nicht zutreffend.

*Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:*

- *Homogenität der Ergebnisse: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere inkonsistente Ergebnisse.*

*Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt für den ein indirekter Vergleich vorgenommen wird fortlaufend in einem eigenen Abschnitt dar.

#### 4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.*

Nicht zutreffend.

#### 4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

*Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).*

Nicht zutreffend.

#### 4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

##### 4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

**4.3.2.2 Charakteristika der nicht randomisierten vergleichenden Studien**

*Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.*

*Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.*

*Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.*

Tabelle 4-108: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

*Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.*

Nicht zutreffend.

#### 4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus nicht randomisierten vergleichenden Studien beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-109: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

#### 4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-110: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-111: Verzerrungsaspekte für &lt;Endpunkt xxx&gt; – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

*Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.*

Nicht zutreffend.

*Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

#### **4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien**

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.*

Nicht zutreffend.

#### 4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien

*Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).*

Nicht zutreffend.

#### 4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

##### 4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie***

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

#### 4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht zutreffend.

#### 4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus weiteren Untersuchungen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-112: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

#### 4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-113: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht zutreffend.

*Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

#### **4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen**

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

#### **4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen**

*Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).*

Nicht zutreffend.

#### 4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

##### 4.4.1 Beurteilung der Aussagekraft der Nachweise

*Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.*

Zur Darstellung des Zusatznutzens von Luspatercept zur Behandlung von erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, wurde im vorliegenden Dossier die Phase-III-Studie MEDALIST herangezogen, auf deren Basis die Zulassung durch die Europäische Kommission erteilt wurde.

MEDALIST ist eine doppelblinde randomisiert kontrolliert durchgeführte multizentrische Studie, in der die Sicherheit und Wirksamkeit von Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, untersucht wurde. In beiden Behandlungsarmen war im Rahmen von BSC eine bedarfsgerechte Transfusionstherapie mit EK in Kombination mit Eisenchelatoren erlaubt, sofern angezeigt. Somit wurde eine adäquate supportive Therapie, die auch den derzeitigen Behandlungsstandard in der vorliegenden Indikation darstellt, sichergestellt. MEDALIST wurde doppelblind durchgeführt. Die Aufrechterhaltung der Verblindung erfolgte unter anderem durch eine zur Luspatercept-Therapie analoge Placebo-Gabe.

Gemäß § 5 Absatz 6 des 5. Kapitels der Verfo entspricht die Studie MEDALIST als RCT der Evidenzstufe 1b. Darüber hinaus wurde das Verzerrungspotential sowohl auf Studien- als auch auf Endpunktebene als niedrig eingestuft.

Die folgenden patientenrelevanten Endpunkte wurden dargestellt und zur Ableitung eines Zusatznutzens herangezogen (siehe auch Abschnitt 4.2.5.2):

- **Mortalität** – Gesamtüberleben
- **Morbidität**
  - Transfusionsfreiheit
  - Reduktion der Transfusionslast, gemessen als HI-E
  - Symptomatik, gemessen anhand des EORTC QLQ-C30-Fragebogens
  - Hospitalisierungen

- **Lebensqualität**

- gesundheitsbezogene Lebensqualität gemessen anhand des generischen Fragebogens EORTC QLQ-C30
- krankheitsspezifische Lebensqualität gemessen anhand des krankheitsspezifischen Fragebogens QoL-E

- **Sicherheit**

- (jegliche UE)
- SUE
- schwere UE (CTCAE  $\geq$  3)
- Therapieabbrüche aufgrund von UE

Insgesamt ergibt sich somit auf Basis der hohen Evidenzstufe der Studie MEDALIST, des niedrigen Verzerrungspotentials und der Betrachtung patientenrelevanter Endpunkte **eine hohe Aussagesicherheit** der Ergebnisse. Dementsprechend wird die Evidenzlage als **Hinweis** für einen Zusatznutzen gewertet.

#### **4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß**

*Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie worin der Zusatznutzen besteht.*

*Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.*

*Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):*

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

*Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.*

Bei Luspatercept handelt es sich um ein Arzneimittel zur Behandlung eines seltenen Leidens gemäß § 35a Absatz 1 Satz 11 SGB V, sodass bereits durch die Zulassung der medizinische Zusatznutzen als belegt gilt.

Zur Ableitung des Ausmaßes des Zusatznutzens wurden die Ergebnisse der pivotalen Zulassungsstudie MEDALIST herangezogen, die in Tabelle 4-114 zusammengefasst sind. Die demographischen und krankheitsspezifischen Charakteristika der eingeschlossenen Patienten sowie das Studiendesign stellen eine gute Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext sicher.

Im vorliegenden Anwendungsgebiet der Niedrigrisiko-MDS tritt bei ca. 70 – 80 % der Patienten als Erstmanifestation eine Anämie infolge eines Erythrozyten-Reifungs-Defekts auf, welche auch als Kernsymptom angesehen werden kann [2-4]. Bei den meisten Patienten führt die Anämie zu einer erheblichen Krankheitslast, Verringerung der Leistungsfähigkeit und Einschränkungen in der Lebensqualität. Im Gegensatz zu fortgeschrittenen Stadien

(Hochrisiko-MDS), bei denen die Verlängerung des Gesamtüberlebens und Verzögerung der Progression in eine AML im Vordergrund steht, stehen bei Patienten mit Niedrigrisiko-MDS, die vergleichsweise lange leben, andere Therapieziele im Vordergrund. Das primäre Therapieziel der Niedrigrisiko-MDS ist die Behandlung der Zytopenie(n), insbesondere der Anämie, um die damit assoziierte Morbidität und Symptomlast zu minimieren und die Leistungsfähigkeit, Lebensqualität und Autonomie der Patienten zu erhalten oder zu verbessern [2, 5, 6]. Für den größten Anteil an Patienten mit Niedrigrisiko-MDS stehen zu diesem Zweck weder potentiell kurative noch zielgerichtete Therapieoptionen zur Verfügung. Für diese Patienten kommt lediglich eine supportive Therapie in Frage.

Zentraler Bestandteil der Therapie einer Anämie – für Patienten im Anwendungsgebiet von Luspatercept sogar die einzige zugelassene Behandlungsoption – ist die supportive Behandlung mit EK-Transfusionen, welche in Abhängigkeit des klinischen Zustands (dynamisch bestimmt durch Symptome, Alter, Erkrankungsstadium, individuelle Patientenbiologie, Komorbiditäten) des Patienten verabreicht werden [2].

Durch die Verabreichung von fremden, funktionstüchtigen Erythrozyten kann der Mangel an körpereigenen, reifen Erythrozyten temporär kompensiert, der Hb-Wert transient erhöht und die Symptomatik der Anämie vorübergehend gemildert werden. Die Transfusionstherapie mit EK vermag jedoch nicht eine anhaltende, stabile Korrektur des Hb-Spiegels und somit andauernde Vermeidung der Anämie-Symptomatik zu erreichen. Darüber hinaus birgt sie das Risiko teils schwerwiegender Folgekomplikationen, wie die Eisenüberladung, geht mit erheblichen psychosozialen und zeitlichen Belastungen für den Patienten einher und beeinflusst die Überlebensprognose negativ [7, 8, 16]. Die Transfusionstherapie mit EK stellt daher nur eine unzureichende therapeutische Maßnahme dar.

Unter Berücksichtigung der mangelnden Therapiealternativen und bedeutsamen Limitationen und Risiken einer Transfusionstherapie besteht für die Patienten im vorliegenden Anwendungsgebiet ein großer, ungedeckter therapeutischer Bedarf nach einem Arzneimittel, das das Erreichen einer Transfusionsfreiheit ermöglicht und zu einer besseren Kontrolle der Erkrankung, insbesondere der Anämie, führt.

Luspatercept ist der erste Vertreter der neuen Wirkstoffklasse der ERA, der den der Anämie zugrundeliegenden Erythrozyten-Reifungs-Defekt gezielt korrigiert und damit die Anämie verbessert und eine Vermeidung von EK-Transfusionen und transfusionsbedingten Komplikationen ermöglicht. Die Behandlung mit Luspatercept bei erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, bewirkt gegenüber Placebo, jeweils in Kombination mit BSC, eine **Verbesserung der Anämie und Aufrechterhaltung der Lebensqualität sowie Autonomie** bei gleichzeitiger **Vermeidung von Transfusionen** und somit auch einer **Minimierung transfusionsassoziierter Risiken und Folgekomplikationen**. Die hohe Wirksamkeit bei gleichzeitig guter Verträglichkeit wird den Therapiezielen im Niedrigrisiko-MDS gerecht.

Der Zusatznutzen von Luspatercept beruht insbesondere auf

- *einer schnell einsetzenden und langfristigen, klinisch relevanten Vermeidung von EK-Transfusionen:*
  - konsistent über alle betrachteten klinisch relevanten Zeiträume zeigte sich unter Luspatercept deutlich häufiger und schneller eine vollständige Vermeidung von EK-Transfusionen. Dies ist zum einen auf das Fehlen von Anämie-Symptomen zurückzuführen, die Anlass zu einer erneuten Transfusion darstellen würden, und entspricht zum anderen einer Vermeidung von transfusionsassoziierten Folgekomplikationen:
    - Transfusionsfreiheit über mindestens acht Wochen:
      - RR = 2,88 (95 %-KI: [1,56; 5,31])
      - HR = 3,92 (95 %-KI: [2,12; 7,25])
    - Transfusionsfreiheit über mindestens 12 Wochen:
      - RR = 3,56 (95 %-KI: [1,59; 7,99])
      - HR = 3,56 (95 %-KI: [1,75; 7,23])
    - Transfusionsfreiheit über mindestens 16 Wochen:
      - RR = 4,80 (95 %-KI: [1,51; 15,26])
      - HR = 4,25 (95 %-KI: [1,81; 9,98])
    - Transfusionsfreiheit über mindestens 24 Wochen:
      - RR = 9,93 (95 %-KI: [1,36; 72,64])
      - HR = 5,78 (95 %-KI: [2,06; 16,17])
  - darüber hinaus zeigte sich ein starker Vorteil hinsichtlich der Langfristigkeit der Vermeidung von Transfusionen, gemessen als längste Dauer eines transfusionsfreien Zeitraums: HR = 0,46 (95 %-KI: [0,33; 0,63])

- *einer langfristigen Verbesserung der Erythropoese:*
  - die Verbesserung des klinischen Zustands des Patienten findet Ausdruck in einer großen, statistisch signifikanten und klinisch relevanten Stabilisierung des Hb-Werts, die über den gesamten Studienverlauf anhielt und die effektive Verstärkung der Erythropoese und Vermeidung der Anämie widerspiegelt:
    - LS-MWD = 0,54 (95 %-KI: [0,34; 0,74])
    - SMD = 0,61 (95 %-KI: [0,33; 0,89])
- *einer deutlichen Überlegenheit hinsichtlich des Erreichens eines erythroiden Ansprechens (HI-E):*
  - auch das klinisch relevante Therapieziel HI-E, das die Verbesserung der MDS-bedingten erythroiden Zytopenie (die Anämie) unter Berücksichtigung von sowohl klinisch relevanten Anhebung des Hb-Spiegels als auch Senkung der Transfusionslast abbildet, wurde unter Luspatercept sowohl häufiger als auch schneller erreicht:
    - RR = 4,47 (95 %-KI: [2,38; 8,41])
    - HR = 4,37 (95 %-KI: [2,48; 7,70])
  - darüber hinaus zeigte sich unter Luspatercept auch eine deutliche, klinisch relevante zeitliche Entlastung durch Verringerung der Häufigkeit notwendiger Transfusionstermine – im Mittel entspricht diese Entlastung einer Verringerung der Transfusionstermine von fast 19 Tagen auf knapp 13 Tage pro Jahr, also etwa einer Woche weniger Fehltagen bei der Arbeit bzw. aufzubringende Zeit für die Therapie:
    - LS-MWD = -1,78 (95 %-KI: [-2,46; -1,11])
    - SMD = -0,62 (95 %-KI: [-0,91; -0,32])
- *einer über den gesamten Studienzeitraum stabil aufrecht erhaltenen, guten Lebensqualität und somit Vermeidung der weiteren Beeinträchtigung der Lebensqualität durch den Krankheitsverlauf:*
  - die über den EORTC QLQ-C30 erhobene gesundheitsbezogene Lebensqualität blieb auch unter Luspatercept bei Minimierung der supportiven Maßnahmen, insbesondere der Transfusionstherapie, konstant gut
  - der Effekt der Vermeidung von Transfusionen auf den Alltag der Patienten zeigte sich insbesondere im signifikant höheren Anteil an Patienten unter Luspatercept, die zu Woche 24 keine Beeinträchtigung des Alltags durch die Transfusionsabhängigkeit wahrnahmen; der Anteil im Luspatercept-Arm betrug 44,5 % im Vergleich zu 20,8 % im Placebo-Arm (p = 0,003)

- *einem gleichzeitig günstigen Sicherheits- und Verträglichkeitsprofil, das zu keiner Einschränkung in der Behandlung führt.*

Tabelle 4-114: Zusammenfassung der Ergebnisse zum medizinischen Zusatznutzen von Luspatercept

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
<b>Mortalität</b>			
<b>Gesamtüberleben</b> (bis zum Datenschnitt vom 01.07.2019)	L: 27 (17,6)/153; mOS = N.E. P: 18 (23,7)/76; mOS = N.E.	HR = 0,75 [0,41; 1,38]	0,357
<b>Morbidität</b>			
<b>Transfusionsfreiheit</b> (Woche 1 – 24)			
≥ 8 Wochen	L: 58 (37,9)/153 P: 10 (13,2)/76	RR = 2,88 [1,56; 5,31]	< 0,001
≥ 12 Wochen	L: 43 (28,1)/153 P: 6 (7,9)/76	RR = 3,56 [1,59; 7,99]	< 0,001
≥ 16 Wochen	L: 29 (19,0)/153 P: 3 (3,9)/76	RR = 4,80 [1,51; 15,26]	0,002
≥ 24 Wochen	L: 20 (13,1)/153 P: 1 (1,3)/76	RR = 9,93 [1,36; 72,64]	0,003
<i>Regulation der Erythropoese</i> <i>Hb-Wert [g/dl]/16 Wochen</i>	L: 152 (99,3)/153 P: 75 (98,7)/76	LS-MWD = 0,54 [0,34; 0,74]	< 0,001
<b>Reduktion der Transfusionslast</b> (HI-E modifiziert gemäß IWG 2006 <sup>(4)</sup> ; Woche 1 – 24)			
HI-E (gesamt)	L: 81 (52,9)/153 P: 9 (11,8)/76	RR = 4,47 [2,38; 8,41]	< 0,001
HI-E I <sup>(5)</sup>	L: 52 (48,6)/107 P: 8 (14,3)/56	RR = 3,40 [1,74; 6,65]	< 0,001
HI-E II <sup>(6)</sup>	L: 29 (63,0)/46 P: 1 (5,0)/20	RR = 12,61 [1,84; 86,27]	< 0,001

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
<i>Reduktion der Eisenüberladung (Woche 1 – 24)</i>			
<i>sFerritin</i>	L: 153 (100)/153 P: 76 (100)/76	<i>LS-MWD = -169,05 [-303,60; -34,50]</i>	<i>0,014</i>
<i>Bedarf an Eisenchelatoren</i>	L: 128 (83,7)/153 P: 68 (89,5)/76	<i>LS-MWD = -41,0 [-120,3; 38,2]</i>	<i>0,309</i>
<b>Symptomatik</b> (EORTC QLQ-C30; Woche 1 – 24)			
Fatigue	L: 32 (21,5)/149 P: 24 (31,6)/76	RR = 0,68 [0,43; 1,07]	0,099
Übelkeit und Erbrechen	L: 17 (11,4)/149 P: 5 (6,6)/76	RR = 1,73 [0,67; 4,52]	0,254
Schmerz	L: 25 (16,8)/149 P: 14 (18,4)/76	RR = 0,91 [0,50; 1,65]	0,758
Dyspnoe	L: 24 (16,1)/149 P: 16 (21,1)/76	RR = 0,77 [0,43; 1,35]	0,360
Schlaflosigkeit	L: 28 (18,8)/149 P: 18 (23,7)/76	RR = 0,79 [0,47; 1,34]	0,390
Appetitminderung	L: 22 (14,8)/149 P: 9 (11,8)/76	RR = 1,25 [0,60; 2,57]	0,548
Obstipation	L: 31 (20,8)/149 P: 13 (17,1)/76	RR = 1,22 [0,68; 2,18]	0,509
Durchfall	L: 11 (7,4)/149 P: 6 (7,9)/76	RR = 0,94 [0,36; 2,43]	0,891
Finanzielle Schwierigkeiten	L: 13 (8,7)/149 P: 9 (11,8)/76	RR = 0,74 [0,33; 1,65]	0,458
<b>Hospitalisierungen</b> (Woche 1 – 24)	L: 34 (22,2)/153 P: 17 (22,4)/76	RR = 0,99 [0,60; 1,64]	0,977
<b>Lebensqualität</b>			
<b>EORTC QLQ-C30</b> (Woche 1 – 24)			
Körperliche Funktion	L: 25 (16,8)/149 P: 18 (23,7)/76	RR = 0,71 [0,41; 1,21]	0,214
Rollenfunktion	L: 30 (20,1)/149 P: 18 (23,7)/76	RR = 0,85 [0,51; 1,42]	0,539
Emotionale Funktion	L: 18 (12,1)/149 P: 11 (14,5)/76	RR = 0,83 [0,42; 1,68]	0,613

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
Kognitive Funktion	L: 29 (19,5)/149 P: 14 (18,4)/76	RR = 1,06 [0,59; 1,88]	0,851
Soziale Funktion	L: 26 (17,4)/149 P: 16 (21,1)/76	RR = 0,83 [0,47; 1,45]	0,512
Gesundheitlicher Status insgesamt	L: 31 (20,8)/149 P: 12 (15,8)/76	RR = 1,32 [0,72; 2,42]	0,367
<b>QoL-E</b> (Woche 1 – 24)			
Körperliches Wohlbefinden	L: 110 (73,8)/149 P: 53 (69,7)/76	SMD = -0,23 [-0,51; 0,05]	0,125
Funktionales Wohlbefinden	L: 108 (72,5)/149 P: 53 (69,7)/76	SMD = -0,19 [-0,47; 0,09]	0,184
Soziales/familiäres Wohlbefinden	L: 102 (68,5)/149 P: 48 (63,2)/76	SMD = -0,21 [-0,51; 0,08]	0,070
Sexuelles Wohlbefinden	L: 105 (70,5)/149 P: 51 (67,1)/76	SMD = 0,00 [-0,29; 0,28]	0,976
Fatigue	L: 110 (73,8)/149 P: 53 (69,7)/76	SMD = -0,34 [-0,62; -0,07]	0,015
MDS-bezogene Störungen	L: 109 (73,2)/149 P: 53 (69,7)/76	SMD = -0,09 [-0,36; 0,19]	0,492
Allgemeines Wohlbefinden	L: 100 (67,1)/149 P: 48 (63,2)/76	SMD = -0,29 [-0,58; 0,00]	0,015
Gesamtsummenskala	L: 100 (67,1)/149 P: 48 (63,2)/76	SMD = -0,24 [-0,53; 0,06]	0,062
Behandlungseffekt	L: 107 (71,8)/149 P: 53 (69,7)/76	SMD = -0,22 [-0,50; 0,06]	0,099
<b>Sicherheit</b>			
Jegliche UE (Woche 1 – 24)	L: 145 (94,8)/153 P: 70 (92,1)/76	RR = 1,03 [0,95; 1,11]	0,429
SOC: Gastrointestinale Erkrankungen	L: 78 (51,0)/153 P: 25 (32,9)/76	RR = 1,55 [1,08; 2,21]	0,010
SOC: Erkrankungen des Nervensystems	L: 60 (39,2)/153 P: 17 (22,4)/76	RR = 1,75 [1,10; 2,79]	0,011
SOC: Augenerkrankungen	L: 12 (7,8)/153 P: 1 (1,3)/76	RR = 5,96 [0,79; 45,00]	0,045
PT: Schwindelgefühl	L: 23 (15,0)/153 P: 4 (5,3)/76	RR = 2,86 [1,02; 7,96]	0,031

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Studie MEDALIST Endpunkt	Anzahl Patienten  n <sup>(1)</sup> (%) / N <sup>(2)</sup>	Luspatercept + BSC vs. Placebo + BSC <sup>(3)</sup>	
		Effektschätzer [95 %-KI]	p-Wert
<i>PT: Asthenie</i>	L: 28 (18,3)/153 P: 6 (7,9)/76	RR = 2,32 [1,00; 5,36]	0,037
<b>SUE</b> (Woche 1 – 24)	L: 40 (26,1)/153 P: 16 (21,1)/76	RR = 1,24 [0,75; 2,07]	0,400
<b>Schwere UE</b> (CTCAE ≥ 3; Woche 1 – 24)	L: 54 (35,3)/153 P: 27 (35,5)/76	RR = 0,99 [0,69; 1,44]	0,972
SOC: Erkrankungen des Nervensystems	L: 8 (5,2)/153 P: 0 (0,0)/76 <sup>(7)</sup>	RR = 8,50 [0,50; 145,34]	0,043
<b>Therapieabbrüche aufgrund von UE</b> (Woche 1 – 24)	L: 12 (7,8)/153 P: 4 (5,3)/76	RR = 1,49 [0,50; 4,47]	0,472

Datenschnitt vom 08.05.2018 (außer anders angegeben).  
Ergänzende, nicht zur Ableitung des Zusatznutzen herangezogene Analysen (*kursiv dargestellt*)

<sup>(1)</sup> Anzahl Studienteilnehmer mit einem Ereignis.  
<sup>(2)</sup> Anzahl analysierter Studienteilnehmer. Für die Endpunkte Gesamtüberleben, Transfusionsfreiheit, Reduktion der Transfusionslast, Reduktion der Eisenüberladung und Hospitalisierungen wurde die ITT-Population herangezogen, die alle randomisierten Studienteilnehmer entsprechend ihrer Zuteilung umfasst. Symptomatik und Lebensqualität wurden in der HRQoL-Population analysiert, welche alle Studienteilnehmer umfasst, die zur Baseline und zu mindestens einem Zeitpunkt im Studienverlauf eine Erhebung aufwiesen. Die Sicherheitsendpunkte wurden anhand der Safety-Population analysiert, welche alle randomisierten Studienteilnehmer umfasst, die mindestens eine Dosis der Studienmedikation erhalten haben. Studienteilnehmer wurden dem Behandlungsarm zugeordnet, deren Behandlung sie auch tatsächlich erhalten haben.  
<sup>(3)</sup> Unter BSC wird insbesondere eine bedarfsgerechte Behandlung mit EK-Transfusionen und Eisenchelatoren verstanden (siehe auch Tabelle 4-13).  
<sup>(4)</sup> In MEDALIST wurde ein HI-E für Patienten mit Baseline-Transfusionslast < 4 EK-Einheiten pro acht Wochen nur erreicht, wenn gleichzeitige zur Erhöhung des Hb-Werts auch eine Transfusionsfreiheit in den betrachteten acht Wochen vorlag.  
<sup>(5)</sup> HI-E I = erythroides Ansprechen für Patienten mit einer Transfusionslast zur Baseline von ≥ 4 EK-Einheiten pro acht Wochen, definiert als Reduktion der Transfusionslast um ≥ 4 EK-Einheiten pro acht Wochen.  
<sup>(6)</sup> HI-E II = erythroides Ansprechen für Patienten mit einer Transfusionslast zur Baseline von < 4 EK-Einheiten pro acht Wochen, definiert als Erhöhung des Hb-Werts um ≥ 1,5 g/dl bei gleichzeitiger Transfusionsfreiheit über acht Wochen.  
<sup>(7)</sup> Nullzellenkorrektur zur Berechnung der Effektschätzer wurde durchgeführt.

AML = akute myeloische Leukämie; BSC = Best-Supportive-Care; CTCAE = *Common Terminology Criteria for Adverse Events*; EK = Erythrozytenkonzentrat; EORTC QLQ-C30 = *European Organization for Research and Treatment of Cancer Quality-of-Life Questionnaire Core Module*; Hb = Hämoglobin; HI-E = erythroides Ansprechen (*hematologic improvement-erythroid*); HR = Hazard Ratio; HRQoL = gesundheitsbezogene Lebensqualität (*health-related quality of life*); ITT = Intention-to-Treat; IWG = *International Working Group*; KI = Konfidenzintervall; L = Luspatercept; N.E. = nicht erreicht oder nicht berechnet; P = Placebo; PT = bevorzugter Begriff (*preferred term*); QoL-E = *Quality of Life-E*; RR = relatives Risiko; sFerritin = Serum-Ferritin; SMD = standardisierte Mittelwertdifferenz; SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis  
Quelle: [1, 18-20]

Die Auswertungen der Endpunkte erfolgten primär zu Woche 24. Bis Woche 24 wurden alle randomisierten Patienten entsprechend ihrer jeweiligen Zuordnung behandelt. Zu Woche 25 fand eine Selektion von Patienten statt, die nach Ermessen des Prüfarztes in den ersten 24 Wochen von der Behandlung profitiert hatten. Somit liegt ab Woche 25 eine Imbalance der behandelten Patienten zwischen den beiden Behandlungsarmen vor (Weiterbehandlung nach Woche 25 zum Datenschnitt vom 08.05.2018: Luspatercept vs. Placebo: 51 % vs. 15,8 %), was eine Interpretierbarkeit der Daten limitiert. Die Aussagen zum medizinischen Zusatznutzen basieren daher auf den Ergebnissen der ersten 24 Wochen, die bereits zum Datenschnitt vom 08.05.2018 vollständig vorlagen. Davon ausgenommen sind die Ergebnisse zum Gesamtüberleben, welche für den gesamten verfügbaren Studienzeitraum bis zum aktuellsten Datenschnitt vom 01.07.2019 berichtet werden. Die Ergebnisse bis zum letzten verfügbaren Datenschnitt (01.07.2019) werden der Vollständigkeit halber als Zusatzanalysen ergänzend berichtet. Relevante neue Erkenntnisse zur Beurteilung des Zusatznutzens von Luspatercept basierend auf den Datenschnitten aus dem Jahr 2019 liegen nicht vor und wären auch nicht zu erwarten gewesen.

Bezogen auf die einzelnen Endpunktkategorien ist das Ausmaß des Zusatznutzens für Luspatercept zur Behandlung von erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von Niedrigrisiko-MDS mit Ringsideroblasten, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, maßgeblich wie folgt begründet:

## **Mortalität**

### ***Gesamtüberleben***

Zwischen den beiden Behandlungsarmen lag hinsichtlich des Gesamtüberlebens kein statistisch signifikanter Unterschied vor. Aufgrund der geringen Ereigniszahlen (27 Patienten (17,6 %) im Luspatercept-Arm und 18 Patienten (23,7 %) im Placebo-Arm) wurde das mediane Gesamtüberleben in beiden Behandlungsarmen bislang nicht erreicht (siehe Tabelle 4-21; Datenschnitt vom 01.07.2019).

### ***Zusammenfassung der Ergebnisse der Mortalität***

Insgesamt zeigten sich für die Endpunktkategorie Mortalität, gemessen als Gesamtüberleben, keine signifikanten Unterschiede zwischen Luspatercept und Placebo, jeweils in Kombination mit BSC.

## **Morbidität**

In der Endpunktkategorie Morbidität wurde der Zusatznutzen hinsichtlich der Vermeidung bzw. Einsparung von EK-Transfusionen anhand der Endpunkte Transfusionsfreiheit und Reduktion der Transfusionslast (gemessen als HI-E gemäß IWG 2006), sowie der Symptomatik (gemessen anhand des EORTC QLQ-C30-Fragebogens) und Hospitalisierungen bewertet. Änderungen der klinisch relevanten Parameter Hb-Spiegel zur Bemessung der Verbesserung der Erythropoese und Vermeidung der Anämie sowie Eisenparameter zur Abschätzung der Hämosiderose wurden ergänzend berichtet.

## ***Transfusionsfreiheit***

Das Erreichen einer Transfusionsfreiheit ist das bedeutendste Therapieziel im vorliegenden Anwendungsgebiet und stellt einen patientenrelevanten Endpunkt dar. Eine Transfusionsfreiheit entspricht einem Fehlen von Anämie-Symptomen, die EK-Transfusionen indizieren würden, und geht mit einer Vermeidung von transfusionsassoziierten Nebenwirkungen und Folgekomplikationen einher. Das Erreichen einer Transfusionsfreiheit war primärer Endpunkt der Studie MEDALIST. Das IQWiG und der G-BA haben in vorangegangenen Verfahren das Erreichen einer Transfusionsfreiheit bereits als patientenrelevanten Endpunkt anerkannt [21-23].

Unter Luspatercept war die Wahrscheinlichkeit eine Transfusionsfreiheit zu erreichen etwa drei- (bezogen auf eine Transfusionsfreiheit über mindestens acht Wochen) bis zehnmal (bezogen auf eine Transfusionsfreiheit über  $\geq 24$  Wochen) höher als unter Placebo. Dies entspricht einer absoluten Risikodifferenz (Luspatercept – Placebo) von 24,8 %-Punkten (bezogen auf eine Transfusionsfreiheit über mindestens acht Wochen) bzw. 11,8 %-Punkten (bezogen auf eine Transfusionsfreiheit über  $\geq 24$  Wochen) zugunsten der mit Luspatercept behandelten Patienten. Unabhängig davon, wie lange die Mindestdauer der Transfusionsfreiheit definiert war (acht, 12, 16 oder 24 Wochen), ergab sich in allen Fällen ein statistisch signifikanter und im Ausmaß großer Unterschied zwischen den beiden Behandlungsarmen (siehe Tabelle 4-24). Die betrachteten Mindestzeiträume einer Transfusionsfreiheit entsprechen den derzeitig anerkannten und etablierten Kriterien für klinisch relevante Therapieeffekte im vorliegenden Anwendungsgebiet und stellen patientenrelevante Effekte dar. Darüber hinaus entspricht eine Transfusionsfreiheit im Kontext einer zuvor transfusionsbedürftigen Erkrankung einer bislang nicht erreichten, großen Verbesserung des therapielevanten Nutzens: Dies trifft insbesondere zu, da eine anhaltende Freiheit von Symptomen erzielt wird und gleichzeitig das Risiko für teils schwerwiegende Nebenwirkungen und Folgekomplikationen vermieden werden kann.

Ergänzend dargestellte Analysen untermauern diesen deutlichen medizinischen Zusatznutzen. So erreichten mit Luspatercept behandelte Patienten statistisch signifikant schneller eine Transfusionsfreiheit als Patienten unter Placebo (siehe Tabelle 4-25). Darüber hinaus zeigte sich hinsichtlich der Dauer transfusionsfreier Intervalle ein großer Vorteil, der einem unter Luspatercept im Vergleich zu Placebo um fast 55 % reduzierten Risiko entspricht, eine erneute EK-Transfusion zu benötigen (siehe Tabelle 4-26). Aus den Swimmer-Plots der Transfusionszeitpunkte und Intervalle einer Transfusionsfreiheit wird zudem deutlich, dass Patienten im Luspatercept-Arm, deren Transfusionsfreiheit durch erneut benötigte EK-Transfusionen unterbrochen wurde, anschließend wieder transfusionsfrei wurden, d. h. mehrfach ansprachen. Unter Placebo erzielten die meisten Patienten, falls sie überhaupt ansprachen, hingegen nur ein einmaliges Ansprechen; mehr als zwei Intervalle einer Transfusionsfreiheit wurden nicht beobachtet (siehe Abbildung 4-10). Aus der bisherigen Erfahrung mit transfusionspflichtigen Patienten mit MDS ist die „erstmalige erneute Transfusion nach Erreichen einer Transfusionsfreiheit“ üblicherweise gleichbedeutend mit „Wirkungsverlust“. Dieses Paradigma ist mit Luspatercept nicht mehr gültig.

Insgesamt zeigten sich hinsichtlich des Erreichens einer Transfusionsfreiheit folgende Effekte:

Unter Luspatercept

- erreichten deutlich mehr Patienten eine langfristige Transfusionsfreiheit, die zudem deutlich früher eintrat,
- erreichten mehr Patienten multiple Phasen der Transfusionsfreiheit und
- die Transfusionsfreiheit hielt deutlich länger an

als unter Placebo.

In Abschnitt 4.3.1.3.1 ergänzend dargestellte Ergebnisse zum Hb-Spiegel belegen zudem, dass Luspatercept effektiv in die Regulation der defekten Erythropoese eingreift, die Erythrozyten-Reifung aktiviert und somit eine Verbesserung des Hb-Spiegels und der Anämie bewirkt. Dieser Effekt setzte bereits während des ersten Behandlungszyklus mit Luspatercept ein und war über den gesamten Studienzeitraum hinweg stabil (siehe Tabelle 4-27 und Abbildung 4-11). Im vorliegenden Kontext ist diese Regulation der Erythropoese insbesondere relevant, als dass nur durch eine stabile Erhöhung des Hb-Werts eine Reduktion der Transfusionslast bis hin zur Transfusionsfreiheit erreicht werden kann. Anders ausgedrückt bedeutet dies, dass eine Transfusionsfreiheit nicht auf Kosten des Hb-Werts und der Anämie erreicht wird, sondern die effektive Aktivierung der Erythrozyten-Reifung widerspiegelt.

### ***Reduktion der Transfusionslast***

Neben dieser deutlichen Überlegenheit von Luspatercept hinsichtlich einer kompletten Transfusionsfreiheit zeigten sich zudem konsistent statistisch signifikante und klinisch relevante Effekte hinsichtlich einer Reduktion der Transfusionslast. Aufgrund der großen Belastung der Patienten durch regelmäßige EK-Transfusionen und der damit verbundenen Risiken für Transfusionsreaktionen, darunter auch die sekundäre Hämosiderose, stellt bereits die Reduktion der Transfusionslast eine deutliche Verbesserung des therapielevanten Nutzens dar.

Im vorliegenden Anwendungsgebiet ist eine klinisch relevante Reduktion der Transfusionslast im Sinne eines erythroiden Ansprechens (HI-E) als etablierter patientenrelevanter Endpunkt definiert [3, 38]. In Abhängigkeit der Transfusionslast ist ein HI-E als klinisch relevante Reduktion der transfundierten EK-Einheiten um  $\geq 4$  EK-Einheiten pro acht Wochen bzw. eine stabile Erhöhung des Hb-Werts über acht Wochen um  $\geq 1,5$  g/dl definiert (in MEDALIST wurde eine Erhöhung des Hb-Werts nur bei gleichzeitiger Transfusionsfreiheit als Ansprechen gewertet). Sowohl hinsichtlich eines HI-E insgesamt als auch bei Betrachtung des HI-E innerhalb der Populationen mit unterschiedlicher Transfusionslast vor Behandlungsbeginn, zeigten sich konstant statistisch signifikante Vorteile für Luspatercept (siehe Tabelle 4-30). Darüber hinaus trat ein HI-E bei mit Luspatercept behandelten Patienten nicht nur häufiger, sondern auch deutlich früher auf als unter Placebo (siehe Tabelle 4-31).

Auswertungen zur mittleren Änderung der Transfusionslast bestätigen zudem diesen deutlichen Effekt hinsichtlich einer Einsparung von EK-Transfusionen. So benötigten Patienten im Luspatercept-Arm im Mittel über 16 Wochen 3,5 EK-Einheiten weniger als unter Placebo (siehe Tabelle 4-32). Für einen transfusionsabhängigen Patienten mit einer durchschnittlichen Transfusionslast von 11 EK-Einheiten/16 Wochen entspricht diese Reduktion der Transfusionslast einer Einsparung der Menge benötigter EK-Einheiten um etwa ein Drittel und somit auch einer deutlichen Reduktion des zusätzlich aufgenommenen Eisens.

Auch die Transfusionshäufigkeit, d. h. die Anzahl an Transfusionszeitpunkten in einem bestimmten Zeitraum, konnte durch Luspatercept deutlich reduziert werden. Für einen transfusionsabhängigen Patienten, der beispielsweise alle drei Wochen eine EK-Transfusion erhalten hatte, entspricht die unter Luspatercept erreichte mittlere Änderung der Transfusionshäufigkeit fast einer Halbierung der Transfusionshäufigkeit von etwa fünf Zeitpunkten auf drei Zeitpunkte pro 16 Wochen (siehe Tabelle 4-33). Dies entspricht einer deutlichen zeitlichen Entlastung für den Patienten und einer Vereinfachung seines mit regelmäßigen Transfusionen zusammenhängenden relevanten organisatorischen Aufwands.

### ***Reduktion der Eisenüberladung***

Das Auftreten einer sekundären Eisenüberladung ist eine typische Komplikation polytransfundierter Patienten, die mit einer erhöhten Morbidität und Mortalität einhergeht. Trotz der Verfügbarkeit und des Einsatzes von Eisenchelatoren, ist eine Eisenüberladung oftmals nicht adäquat behandelbar, sodass es zu Siderose-bedingten Organschäden kommen kann. Effektiver als eine medikamentöse Chelierung überschüssigen Eisens ist die Vermeidung der übermäßigen Eisenzufuhr durch Vermeidung oder Minimierung von Transfusionen.

Der sFerritin-Spiegel dient als klinisch relevanter Parameter zur Abschätzung von Trends im Eisenhaushalt. Bereits innerhalb der ersten 24 Wochen der Behandlung konnte der sFerritin-Spiegel unter Luspatercept im Vergleich zu Placebo deutlich gesenkt werden (siehe Tabelle 4-36). Der Effekt war statistisch signifikant. Diese Ergebnisse weisen auf positive Effekte der Vermeidung von Transfusionen durch Luspatercept hinsichtlich der Minimierung des Risikos für Langzeitkomplikationen, insbesondere der Siderose-bedingten Morbidität, hin. Auch für den Bedarf an Eisenchelatoren deutete sich bereits ein positiver Trend an. Da die Dosierung von Eisenchelatoren in Abhängigkeit des sFerritin-Spiegels erfolgt, sind Dosismodifikationen vor allem im langfristigen Verlauf zu erwarten.

### ***Symptomatik gemessen anhand des EORTC QLQ-C30-Fragebogens***

Die Erhebung krankheitsbedingter Symptome anhand des EORTC QLQ-C30-Fragebogens stellt einen patientenrelevanten Endpunkt dar, da er die für die Patienten spürbaren Auswirkungen der Erkrankung sowie der Therapie direkt abbildet.

Untersuchungen zur Verbesserung der Symptomatik, gemessen als Verbesserung der Symptomskalen des EORTC QLQ-C30-Fragebogens um die MCID von zehn Punkten, zeigten über alle betrachteten Symptome keine klinisch relevanten Unterschiede zwischen den beiden Behandlungsarmen (siehe Tabelle 4-41). Die Symptomatik war somit zwischen Luspatercept bei gleichzeitiger Vermeidung von EK-Transfusionen und Placebo unter Beibehaltung der konventionellen Transfusionstherapie vergleichbar. Durch die adäquate Behandlung der Patienten im Rahmen der BSC war das Auftreten von Symptomen auch nicht zu erwarten gewesen.

### ***Hospitalisierungen***

Aufgrund der direkten Auswirkungen von Hospitalisierungen auf die Wahrnehmung von Funktionen und Aktivitäten der Patienten sind stationäre Aufnahmen patientenrelevant.

Zwischen den beiden Behandlungsarmen zeigten sich keine statistisch signifikanten Unterschiede hinsichtlich stationärer Aufnahmen jeglicher Ursache während der Behandlungsphase bis Woche 24 (siehe Tabelle 4-48). Hauptgrund für Hospitalisierungen waren in beiden Behandlungsarmen UE. Auch bei Betrachtung der stationären Aufnahmen aufgrund von UE und/oder Transfusionen zeigten sich keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen (siehe Tabelle 4-50). Dies trifft auch bei längerer Beobachtung über Woche 25 hinaus und trotz des Ungleichgewichts zwischen den beiden Behandlungsarmen aufgrund der Selektion der Patienten zu Woche 25 zu (siehe Tabelle 4-52).

### ***Zusammenfassung der Ergebnisse der Morbidität***

Im vorliegenden Anwendungsgebiet der Niedrigrisiko-MDS tritt bei ca. 70 – 80 % der Patienten als Erstmanifestation eine Anämie infolge eines Erythrozyten-Reifungs-Defekts auf, welche auch als Kernsymptom angesehen werden kann und im Zentrum aller Komplikationen steht [2-4]. Die Anämie beeinflusst den klinischen Zustand des Patienten maßgeblich und macht die Transfusionstherapie notwendig. Entsprechend ist eine Verbesserung der Anämie durch eine Korrektur des Erythrozyten-Reifungs-Defekts und somit eine Verbesserung des klinischen Zustands, der sich unmittelbar in einer Vermeidung oder zumindest Reduktion der Transfusionslast sowie stabilen Erhöhung der Hb-Werte widerspiegelt, das höchste Therapieziel. Langfristig kann dadurch auch der Eisenhaushalt normalisiert werden und das Gesamtüberleben verlängert werden.

Zusammengefasst zeigten sich für Luspatercept in der Kategorie Morbidität ausschließlich positive Ergebnisse und eine deutliche Überlegenheit gegenüber Placebo hinsichtlich des Erreichens einer Transfusionsfreiheit und einer Reduktion der Transfusionslast. Diese patientenrelevanten Therapieziele konnten unter Luspatercept nicht nur häufiger, sondern auch deutlich schneller erreicht werden als unter Placebo. Der Verbesserung des klinischen Zustands der Patienten liegt eine effektive Korrektur des Erythrozyten-Reifungs-Defekts und Verbesserung der Erythropoese zugrunde, welche sich in stabil erhöhten Hb-Werten – auch bei Reduktion bzw. Vermeidung von EK-Transfusionen – widerspiegelt. Eine Vermeidung von Transfusionen entspricht somit einer Stabilisierung des Hb-Werts und Vermeidung von Symptomen, die Anlass für eine EK-Transfusion darstellen würden, und bedeutet zugleich eine Vermeidung von transfusionsbedingten Nebenwirkungen und Folgekomplikationen. Das Potential von Luspatercept hinsichtlich einer Reduktion der sekundären Eisenüberladung durch Einsparung an EK-Transfusionen, d. h. langfristiger Vermeidung einer sekundären Hämosiderose und damit verbundenen Komplikationen, zeichnet sich zudem bereits in den Ergebnissen zum sFerritin-Spiegel ab. Statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen hinsichtlich Symptomatik und Hospitalisierungen lagen nicht vor.

Unter Luspatercept zeigte sich eine bedeutsame Verbesserung der Anämie und deren Kontrolle. Insgesamt liegt somit ein erheblicher Vorteil von Luspatercept gegenüber Placebo, jeweils in Kombination mit BSC, hinsichtlich der Endpunktkategorie Morbidität vor.

### **Lebensqualität**

Endpunkte, die die Lebensqualität abbilden, sind patientenrelevant, da sie die Auswirkungen der Erkrankung sowie der Behandlung auf das Befinden des Patienten direkt abbilden. Die Lebensqualität wurde in MEDALIST anhand des generischen Fragebogens für onkologische Erkrankungen EORTC QLQ-C30 sowie des krankheitsspezifischen Fragebogens QoL-E erhoben. Die Ergebnisse der Funktionsskalen des EORTC QLQ-C30-Fragebogens sowie der Domänen und Summenscores des QoL-E-Fragebogens wurden zur Ableitung des Zusatznutzens herangezogen. In MEDALIST war in beiden Behandlungsarmen eine adäquate supportive Therapie, insbesondere die bedarfsgerechte Transfusionstherapie mit EK, ggf. in Kombination mit Eisenchelatoren, erlaubt, die geeignet ist, die Beeinträchtigungen der Lebensqualität durch die Anämie zu minimieren. Im vorliegenden Anwendungsgebiet stellt insbesondere der Erhalt der Lebensqualität und Autonomie ein patientenrelevantes Therapieziel dar.

Dieses Therapieziel, die Lebensqualität aufrecht zu erhalten wurde unter Luspatercept auch bei einer Minimierung der supportiven Maßnahmen, insbesondere der Transfusionstherapie, erreicht. Die Lebensqualität war über den gesamten Studienverlauf hinweg konstant, klinisch relevante Unterschiede zwischen den beiden Behandlungsarmen lagen nicht vor.

Durch die Vermeidung von Transfusionen unter Luspatercept zeigte sich zudem ein positiver Einfluss auf die Beeinträchtigungen durch die Transfusionsabhängigkeit im Alltag. Fast die Hälfte aller mit Luspatercept behandelten Patienten gaben zu Woche 24 an, im Alltag durch die Transfusionsabhängigkeit überhaupt nicht beeinträchtigt zu sein. Im Placebo-Arm waren es nur rund ein Fünftel aller Patienten, die diese Einschätzung teilten. Der Unterschied war statistisch signifikant und spiegelt die Linderung der transfusionsassoziierten Belastungen im Alltag durch Luspatercept wider.

### ***Zusammenfassung der Ergebnisse der Lebensqualität***

Insgesamt zeigten sich für die Endpunktkategorie Lebensqualität keine klinisch relevanten Unterschiede zwischen den beiden Behandlungsarmen. Die Lebensqualität wurde in beiden Behandlungsarmen über den gesamten Studienverlauf stabil aufrechterhalten und war vergleichbar. Die Vermeidung von Transfusionen durch Luspatercept schlägt sich in einer für die Patienten spürbaren Vermeidung alltäglicher transfusionsbedingter Beeinträchtigungen nieder.

## Sicherheit

Die Betrachtung von UE ist bei der Bewertung des medizinischen Nutzens und Zusatznutzens unerlässlich, um die Verträglichkeit und Sicherheit eines Arzneimittels im Behandlungskontext zu bewerten, und stellt somit einen patientenrelevanten Endpunkt dar.

Prinzipiell wären unter der aktiven, zielgerichteten Behandlung mit Luspatercept im Vergleich zu Placebo, jeweils in Kombination mit BSC, auch höhere Raten an Patienten mit UE zu erwarten gewesen. Unter der Behandlung mit Luspatercept zeigten sich aber gegenüber Placebo keine statistisch signifikanten Unterschiede in den Gesamtraten an jeglichen UE, SUE, schweren UE oder Therapieabbrüchen aufgrund von UE, sodass von einer sehr guten Verträglichkeit von Luspatercept im Vergleich zu Placebo und mit keiner Einschränkung für die mit Luspatercept behandelten Patienten auszugehen ist.

Statistisch signifikante aber als geringfügig und damit nicht relevant zu klassifizierende Unterschiede lagen bezogen auf die UE-Kategorie jegliche UE für die SOC „Gastrointestinale Erkrankungen“, „Erkrankungen des Nervensystems“, „Augenerkrankungen“ und die beiden PT „Asthenie“ und „Schwindelgefühl“ sowie für die UE-Kategorie schwere UE für die SOC „Erkrankungen des Nervensystems“ vor (siehe Tabelle 4-72). Diese Unterschiede spiegeln sich zudem nicht in einer Verschlechterung der Lebensqualität oder Symptomatik (gastrointestinale Nebenwirkungen) gemessen anhand des EORTC QLQ-C30-Fragebogens wider. Bezogen auf schwere UE und SUE gab es auch auf Ebene der SOC und PT keine statistisch signifikanten Unterschiede zwischen den beiden Behandlungsarmen. Therapieabbrüche aufgrund von UE traten nur vereinzelt auf, statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen lagen auf Ebene der SOC und PT ebenfalls nicht vor.

Auch die in Abschnitt 4.3.1.3.1 ergänzend dargestellten UE von besonderem Interesse (SMQ „Malignitäten“, SMQ „Prämaligne Erkrankungen“, PT „Neutropenie“, PT „Thrombozytopenie“) traten nur vereinzelt auf, statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen lagen nicht vor (siehe Tabelle 4-82). Ein negativer Einfluss von Luspatercept auf die Entstehung von prämaligen oder malignen Erkrankungen sowie die Pathophysiologie der Erkrankung, insbesondere hinsichtlich weiterer Manifestationen von Zytopenien wie Neutropenie-bedingter Infektionen oder Thrombozytopenie-bedingter Blutungen, konnte nicht festgestellt werden.

Die Endpunkte Zeit bis zur Progression in eine AML sowie das Voranschreiten der Erkrankung in ein Hochrisiko-MDS stellen weitere relevante Sicherheitsaspekte dar und wurden daher ebenfalls gesondert berichtet. Auch hier traten Ereignisse insgesamt nur sehr selten auf, statistisch signifikante Unterschiede zwischen den beiden Behandlungsarmen lagen nicht vor. Die mediane Zeit bis zur Progression in eine AML wurde in beiden Behandlungsarmen aufgrund der geringen Ereigniszahlen nicht erreicht (siehe Tabelle 4-83 f.). Ein negativer Einfluss von Luspatercept auf die Pathophysiologie und Progression der Erkrankung konnte nicht festgestellt werden.

### ***Zusammenfassung der Ergebnisse der Sicherheit***

Insgesamt zeigten sich für die Endpunktkategorie Sicherheit weder Hinweise auf einen größeren noch auf einen geringeren Nutzen von Luspatercept im Vergleich zum Placebo-Arm, jeweils in Kombination mit BSC. Das Sicherheits- und Verträglichkeitsprofil von Luspatercept ist somit als günstig einzustufen und führt zu keiner Einschränkung in der Behandlung.

### **Fazit**

Die Ergebnisse zum medizinischen Nutzen und Zusatznutzen von Luspatercept beruhen auf der doppelblinden, randomisiert kontrollierten Phase-III-Studie MEDALIST. Die Aussagesicherheit der Ergebnisse ist aufgrund der hohen Evidenzstufe der Studie (1b) sowie des geringen Verzerrungspotentials auf Studien- und Endpunktebene als hoch einzustufen. Entsprechend lässt sich basierend auf den Ergebnissen aus MEDALIST ein Hinweis auf einen Zusatznutzen ableiten. Folgende Aspekte sind bei der Bewertung des Zusatznutzens relevant:

- Bei MDS handelt es sich um seltene Erkrankungen des blutbildenden Systems im Knochenmark mit einer erheblichen Krankheitslast und Beeinträchtigung der Lebensqualität. Dabei spielen sowohl erkrankungsbedingte als auch therapiebedingte, d. h. transfusionsbedingte klinische und psychosoziale Komplikationen eine große Rolle. Im Zentrum aller Komplikationen steht die chronische Anämie infolge eines Erythrozyten-Reifungs-Defekts, die auch als Kernproblem der Erkrankung angesehen werden kann. Diese bestimmt den klinischen Zustand der Patienten maßgeblich und macht eine Therapie notwendig.
- Für Patienten im Anwendungsgebiet von Luspatercept gibt es einen großen therapeutischen Bedarf nach einem Arzneimittel, das den Erythrozyten-Reifungs-Defekt korrigiert und dadurch die Anämie verbessert, die Transfusionslast senkt oder vollständig vermeidet und zu einer besseren Kontrolle der Erkrankung und Aufrechterhaltung der Lebensqualität und Autonomie der Patienten führt. Derzeit steht als einzige zugelassene Therapieoption die supportive Transfusionstherapie mit EK, ggf. in Kombination mit Eisenchelatoren, zur Verfügung. Diese kann den Mangel an körpereigenen reifen Erythrozyten zwar temporär kompensieren, den Hb-Wert transient erhöhen und die Symptomatik der Anämie vorübergehend mildern, korrigiert jedoch den Erythrozyten-Reifungs-Defekt und die ineffektive Erythropoese nicht. Eine anhaltende, stabile Korrektur des Hb-Spiegels und somit andauernde Vermeidung der Anämie-Symptomatik ist somit nicht zu erreichen. Darüber hinaus geht die Transfusionstherapie mit EK mit akuten Nebenwirkungen und langfristigen Folgekomplikationen einher.
- Aus Sicht der Patienten stellen regelmäßige Transfusionen eine Beeinträchtigung des alltäglichen Lebens dar. Eine Transfusionsunabhängigkeit oder zumindest Reduktion der Transfusionslast wird von MDS-Patienten ausdrücklich gewünscht. Langfristig können durch die Reduktion der Transfusionslast auch transfusionsassoziierte Folgekomplikationen vermieden und die Morbidität und Mortalität des Patienten gesenkt werden.

- Luspatercept ist der erste Vertreter der neuen Wirkstoffklasse der ERA und damit der erste innovative Therapieansatz, der gezielt gegen den der Anämie zugrundeliegenden Erythrozyten-Reifungs-Defekt wirkt und damit die Anämie, die Kernsymptomatik bei Niedrigrisiko-MDS verbessert. Damit deckt Luspatercept den großen therapeutischen Bedarf im vorliegenden Anwendungsgebiet.
- Mit Luspatercept können MDS-Patienten im Anwendungsgebiet:
  - schneller, häufiger und anhaltender eine Transfusionsfreiheit erreichen als mit Placebo. Dies bedeutet zum einen das Fehlen von Symptomen, die einen Anlass zu einer Transfusion darstellen, und zum anderen die Vermeidung von Folgekomplikationen, die im Rahmen einer Transfusionstherapie auftreten können. Für die Patienten geht eine Vermeidung von Transfusionen zudem mit einem Zeitgewinn und einer Vermeidung der transfusionsbedingten Belastung im Alltag einher.
  - die Anämie durch eine Korrektur des Erythrozyten-Reifungs-Defekts und Förderung der Erythropoese verbessern.
  - die generische Lebensqualität, die aufgrund von MDS oft beeinträchtigt ist, stabil aufrechterhalten. Die Anämie wird durch Luspatercept so gut kontrolliert und verbessert, dass die Patienten auch ohne Transfusionen keine relevante Beeinträchtigung ihrer Lebensqualität durch die Anämie verspüren.
  - eine spürbare Erleichterung hinsichtlich der transfusionsbedingten Störungen im Alltag erreichen.
  - all diese Wirksamkeitsparameter bei einem insgesamt günstigen Sicherheits- und Verträglichkeitsprofil erzielen, das zu keiner Einschränkung von mit Luspatercept behandelten Patienten führt.

Grundsätzlich gilt entsprechend § 5 Absatz 7 des 5. Kapitels der VerfO für Luspatercept zur Behandlung von erwachsenen Patienten mit transfusionsabhängiger Anämie aufgrund von MDS mit Ringsideroblasten, mit sehr niedrigem, niedrigem oder intermediärem Risiko, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind, als belegt. Aufgrund der hohen Evidenzstufe und niedrigen Verzerrung auf Studien- und Endpunktebene ist die Aussagesicherheit der Ergebnisse hoch. Die deutlichen, bisher nicht erreichten Vorteile in der Wirksamkeit bei vergleichbarer Lebensqualität und keiner Einschränkung durch UE belegen den Zusatznutzen von Luspatercept im Vergleich zu Placebo.

**Insgesamt lässt sich so das Ausmaß des Zusatznutzens unter Angabe der Aussagekraft der Nachweise quantifizieren und führt zu einem Hinweis auf einen erheblichen Zusatznutzen für Luspatercept.**

#### 4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-115: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit transfusionsabhängiger Anämie aufgrund von MDS mit Ringsideroblasten, mit sehr niedrigem, niedrigem oder intermediärem Risiko, die auf eine EPO-basierte Therapie nicht zufriedenstellend angesprochen haben oder dafür nicht geeignet sind	Erheblicher Zusatznutzen
EPO = Erythropoetin; MDS = Myelodysplastische Syndrome	

#### 4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

##### 4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht zutreffend.

##### 4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Nicht zutreffend.

#### 4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

*Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.*

Nicht zutreffend.

#### 4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005<sup>18</sup>, Molenberghs 2010<sup>19</sup>). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrundeliegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006<sup>20</sup>) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%- Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006<sup>21</sup>) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

---

<sup>18</sup> Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.

<sup>19</sup> Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.

<sup>20</sup> Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.

<sup>21</sup> Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

*Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.*

Nicht zutreffend.

#### 4.6 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. Celgene (2019): Clinical Study Report ACE-536-MDS-001: A phase 3, double-blind, randomized study to compare the efficacy and safety of luspatercept (ACE-536) versus placebo for the treatment of anemia due to very low, low, or intermediate risk myelodysplastic syndromes in subjects with ring sideroblasts who require red blood cell transfusions.
2. Deutsche Gesellschaft für Hämatologie und Medizinische Onkologie e.V. (DGHO) (2020): Myelodysplastische Syndrome (MDS). [Zugriff: 10.01.2020]. URL: [https://www.onkopedia.com/de/onkopedia/guidelines/myelodysplastische-syndrome-mds/@\\_@guideline/html/index.html](https://www.onkopedia.com/de/onkopedia/guidelines/myelodysplastische-syndrome-mds/@_@guideline/html/index.html).
3. Fenaux P, Haase D, Sanz GF, Santini V, Buske C, Group EGW (2014): Myelodysplastic syndromes: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*; 25 Suppl 3:iii57-69.
4. National Comprehensive Cancer Network (NCCN) (2020): NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) Myelodysplastic Syndromes. Version 2.2020 - February 28, 2020. [Zugriff: 27.05.2020]. URL: [https://www.nccn.org/professionals/physician\\_gls/default.aspx#mds](https://www.nccn.org/professionals/physician_gls/default.aspx#mds).
5. Germing U, Oliva EN, Hiwase D, Almeida A (2019): Treatment of Anemia in Transfusion-Dependent and Non-Transfusion-Dependent Lower-Risk MDS: Current and Emerging Strategies. *HemaSphere*; 3(6):e314.
6. Germing U, Schroeder T, Kaivers J, Kundgen A, Kobbe G, Gattermann N (2019): Novel therapies in low- and high-risk myelodysplastic syndrome. *Expert Rev Hematol*:1-16.
7. Malcovati L, Della Porta MG, Cazzola M (2006): Predicting survival and leukemic evolution in patients with myelodysplastic syndrome. *Haematologica*; 91(12):1588-90.
8. Platzbecker U, Hofbauer LC, Ehninger G, Hölig K (2012): The clinical, quality of life, and economic consequences of chronic anemia and transfusion support in patients with myelodysplastic syndromes. *Leuk Res*; 36(5):525-36.
9. Gattermann N (2018): Iron overload in myelodysplastic syndromes (MDS). *Int J Hematol*; 107(1):55-63.
10. Caocci G, Baccoli R, Ledda A, Littera R, La Nasa G (2007): A mathematical model for the evaluation of amplitude of hemoglobin fluctuations in elderly anemic patients affected by myelodysplastic syndromes: correlation with quality of life and fatigue. *Leuk Res*; 31(2):249-52.
11. Koutsavlis I (2016): Transfusion Thresholds, Quality of Life, and Current Approaches in Myelodysplastic Syndromes. *Anemia*; 2016:8494738-.
12. Agberemi R, MDS UK Patient Support Group (2018): Impact of myelodysplastic syndrome and its treatment on quality of life in patients. [Zugriff: 23.01.2020]. URL: [https://mdspatientsupport.org.uk/wp-content/uploads/2018/11/1111\\_MDS\\_Newsletter\\_Nov2018\\_FINAL.pdf](https://mdspatientsupport.org.uk/wp-content/uploads/2018/11/1111_MDS_Newsletter_Nov2018_FINAL.pdf).

13. Lawrence L (2017): Why Blood Counts Matter in Myelodysplastic Syndrome. [Zugriff: 23.01.2020]. URL: <https://www.curetoday.com/publications/cure/2017/hematology-2-2017/why-blood-counts-matter-in-myelodysplastic-syndrome>.
14. Biehl J (2019): What Does Cancer-Related Fatigue Feel Like? [Zugriff: 23.01.2020]. URL: <https://www.curetoday.com/community/jane-biehl/2019/10/what-does-cancerrelated-fatigue-feel-like>.
15. Thomas ML (2012): The impact of myelodysplastic syndromes on quality of life: lessons learned from 70 voices. The journal of supportive oncology; 10(1):37-44.
16. Blood and Beyond (2020): Rethinking blood use in Europe to improve outcomes for patients. [Zugriff: 13.05.2020]. URL: [https://media.bloodandbeyond.com/wp-content/uploads/BloodandBeyond\\_Rethinking\\_blood\\_use\\_in\\_Europe.pdf](https://media.bloodandbeyond.com/wp-content/uploads/BloodandBeyond_Rethinking_blood_use_in_Europe.pdf).
17. Szende A, Schaefer C, Goss TF, Heptinstall K, Knight R, Lubbert M, et al. (2009): Valuation of transfusion-free living in MDS: results of health utility interviews with patients. Health Qual Life Outcomes; 7:81.
18. Celgene GmbH (2018): Luspatercept MEDALIST Post-hoc Analysen - Datenschnitt 08.05.2018.
19. Celgene GmbH (2018): Luspatercept MEDALIST Post-hoc-Analysen HRQoL - Datenschnitt 08.05.2018.
20. Celgene GmbH (2019): Luspatercept MEDALIST Post-hoc Analysen - Datenschnitt 01.07.2019.
21. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) (2019): IQWiG-Berichte – Nr. 833 Ravulizumab (paroxysmale nächtliche Hämoglobinurie) – Nutzenbewertung gemäß § 35a SGB V. [Zugriff: 09.04.2020]. URL: [https://www.g-ba.de/downloads/92-975-3182/2019-08-01\\_Nutzenbewertung-IQWiG\\_Ravulizumab\\_D-463.pdf](https://www.g-ba.de/downloads/92-975-3182/2019-08-01_Nutzenbewertung-IQWiG_Ravulizumab_D-463.pdf).
22. Gemeinsamer Bundesausschuss (G-BA) (2020): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Ravulizumab. Vom 6. Februar 2020. [Zugriff: 09.04.2020]. URL: [https://www.g-ba.de/downloads/40-268-6320/2020-02-06\\_AM-RL-XII\\_Ravulizumab\\_D-463\\_TrG.pdf](https://www.g-ba.de/downloads/40-268-6320/2020-02-06_AM-RL-XII_Ravulizumab_D-463_TrG.pdf).
23. Gemeinsamer Bundesausschuss (G-BA) (2020): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO. Wirkstoff: Autologe CD34+ hämatopoetische Stammzellen. [Zugriff: 09.04.2020]. URL: [https://www.g-ba.de/downloads/92-975-3407/2019-11-15\\_Nutzenbewertung-G-BA\\_CD34+Zellen\\_D-497.pdf](https://www.g-ba.de/downloads/92-975-3407/2019-11-15_Nutzenbewertung-G-BA_CD34+Zellen_D-497.pdf).
24. Celgene Corporation (2005): CC-5013-MDS-004 - Lenalidomide Versus Placebo in Myelodysplastic Syndromes With a Deletion 5q[31] Abnormality (NCT00179621). Stand des Eintrags: 14.04.2011. [Zugriff: 13.05.2020]. URL: <https://clinicaltrials.gov/ct2/show/record/NCT00179621>
25. Celgene (2003): CC-5013-MDS-003 - Lenalidomide Safety/Efficacy in Myelodysplastic Syndromes (MDS) Associated With a Deletion (Del)(5q) Cytogenetic Abnormality (NCT00065156). Stand des Eintrags: 19.11.2019. [Zugriff: 12.03.2020]. URL: <https://clinicaltrials.gov/ct2/show/record/NCT00065156>

26. Fenaux P, Santini V, Spiriti MAA, Giagounidis A, Schlag R, Radinoff A, et al. (2018): A phase 3 randomized, placebo-controlled study assessing the efficacy and safety of epoetin- $\alpha$  in anemic patients with low-risk MDS. *Leukemia*; 32(12):2648-58.
27. Malcovati L, Della Porta MG, Strupp C, Ambaglio I, Kuendgen A, Nachtkamp K, et al. (2011): Impact of the degree of anemia on the outcome of patients with myelodysplastic syndrome and its integration into the WHO classification-based Prognostic Scoring System (WPSS). *Haematologica*; 96(10):1433-40.
28. Bundesärztekammer (BÄK) (2014): Querschnitts-Leitlinien (BÄK) zur Therapie mit Blutkomponenten und Plasmaderivaten 4. überarbeitete und aktualisierte Auflage 2014. [Zugriff: 20.01.2020]. URL: [https://www.bundesaeztekammer.de/fileadmin/user\\_upload/downloads/QLL\\_Haemotherapie\\_2014.pdf](https://www.bundesaeztekammer.de/fileadmin/user_upload/downloads/QLL_Haemotherapie_2014.pdf).
29. Bundesärztekammer (BÄK) (2017): Richtlinie zur Gewinnung von Blut und Blutbestandteilen und zur Anwendung von Blutprodukten (Richtlinie Hämotherapie). [Zugriff: 20.01.2020]. URL: [https://www.bundesaeztekammer.de/fileadmin/user\\_upload/downloads/pdf-Ordner/MuE/Richtlinie\\_Haemotherapie\\_E\\_A\\_2019.pdf](https://www.bundesaeztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/MuE/Richtlinie_Haemotherapie_E_A_2019.pdf).
30. Gattermann N (2012): Pathophysiological and clinical aspects of iron chelation therapy in MDS. *Current pharmaceutical design*; 18(22):3222-34.
31. Jansen AJ, Essink-Bot ML, Beckers EA, Hop WC, Schipperus MR, Van Rhenen DJ (2003): Quality of life measurement in patients with transfusion-dependent myelodysplastic syndromes. *Br J Haematol*; 121(2):270-4.
32. Oliva EN, Finelli C, Santini V, Poloni A, Liso V, Cilloni D, et al. (2012): Quality of life and physicians' perception in myelodysplastic syndromes. *Am J Blood Res*; 2(2):136-47.
33. Stauder R, Yu G, Koinig KA, Bagguley T, Fenaux P, Symeonidis A, et al. (2018): Health-related quality of life in lower-risk MDS patients compared with age- and sex-matched reference populations: a European LeukemiaNet study. *Leukemia*; 32(6):1380-92.
34. Giagounidis A (2017): Current treatment algorithm for the management of lower-risk MDS. *Hematology Am Soc Hematol Educ Program*; 2017(1):453-9.
35. Institut für Qualität und Wirtschaftlichkeit (IQWiG) (2017): Allgemeine Methoden Version 5.0. [Zugriff: 26.04.2019]. URL: [https://www.iqwig.de/download/Allgemeine-Methoden\\_Version-5-0.pdf](https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf).
36. Celgene (2020): Reblozyl 25 mg/75 mg Pulver zur Herstellung einer Injektionslösung; Fachinformation. Stand: Juni 2020 [Zugriff: 09.07.2020]. URL: <http://www.fachinfo.de>.
37. Greenberg PL, Tuechler H, Schanz J, Sanz G, Garcia-Manero G, Solé F, et al. (2012): Revised International Prognostic Scoring System for Myelodysplastic Syndromes. *Blood*; 120(12):2454-65.
38. Cheson BD, Greenberg PL, Bennett JM, Lowenberg B, Wijermans PW, Nimer SD, et al. (2006): Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. *Blood*; 108(2):419-25.
39. Balducci L (2006): Transfusion independence in patients with myelodysplastic syndromes: impact on outcomes and quality of life. *Cancer*; 106(10):2087-94.
40. Malcovati L, Germing U, Kuendgen A, Della Porta MG, Pascutto C, Invernizzi R, et al. (2007): Time-dependent prognostic scoring system for predicting survival and leukemic evolution in myelodysplastic syndromes. *J Clin Oncol*; 25(23):3503-10.

41. Malcovati L, Porta MGD, Pascutto C, Invernizzi R, Boni M, Travaglino E, et al. (2005): Prognostic Factors and Life Expectancy in Myelodysplastic Syndromes Classified According to WHO Criteria: A Basis for Clinical Decision Making. *Journal of Clinical Oncology*; 23(30):7594-603.
42. Harnan S, Ren S, Gomersall T, Everson-Hock ES, Sutton A, Dhanasiri S, et al. (2016): Association between Transfusion Status and Overall Survival in Patients with Myelodysplastic Syndromes: A Systematic Literature Review and Meta-Analysis. *Acta haematologica*; 136(1):23-42.
43. de Swart L, Crouch S, Hoeks M, Smith A, Langemeijer S, Fenaux P, et al. (2019): Impact of red blood cell transfusion dose density on progression-free survival in lower-risk myelodysplastic syndromes patients. *Haematologica*:haematol.2018.212217.
44. Vamvakas EC, Blajchman MA (2009): Transfusion-related mortality: the ongoing risks of allogeneic blood transfusion and the available strategies for their prevention. *Blood*; 113(15):3406-17.
45. Abetz L, Baladi JF, Jones P, Rofail D (2006): The impact of iron overload and its treatment on quality of life: results from a literature review. *Health Qual Life Outcomes*; 4:73.
46. Killick SB (2017): Iron chelation therapy in low risk myelodysplastic syndrome. *Br J Haematol*; 177(3):375-87.
47. Shammo JM, Komrokji RS (2018): Clinical consequences of iron overload in patients with myelodysplastic syndromes: the case for iron chelation therapy. *Expert Rev Hematol*; 11(7):577-86.
48. Novartis Pharma (2006): EXJADE® Filmtabletten; Fachinformation. Stand: Juli 2019 [Zugriff: 06.05.2020]. URL: <http://www.fachinfo.de>.
49. Novartis Pharma (2003): Desferal®; Fachinformation. Stand: März 2020 [Zugriff: 06.05.2020]. URL: <http://www.fachinfo.de>.
50. Angelucci E, Li J, Greenberg PL, Depei W, Hou M, Montaña Figueroa E, et al. (2018): Safety and Efficacy, Including Event-Free Survival, of Deferasirox Versus Placebo in Iron-Overloaded Patients with Low- and Int-1-Risk Myelodysplastic Syndromes (MDS): Outcomes from the Randomized, Double-Blind Telesto Study. *Blood*; 132(Supplement 1):234-.
51. Angelucci E, Li J, Greenberg P, Wu D, Hou M, Montaña Figueroa E, et al. (2020): Iron Chelation in Transfusion-Dependent Patients With Low- to Intermediate-1-Risk Myelodysplastic Syndromes. *Annals of Internal Medicine*; 172(8):513-22.
52. Gemeinsamer Bundesausschuss (G-BA) (2019): Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV Beratungsanforderung 2019-B-054. Luspatercept zur Behandlung von Anämie aufgrund Myelodysplastischen Syndromen.
53. Platzbecker U, Fenaux P, Ades L, Giagounidis A, Santini V, van de Loosdrecht AA, et al. (2019): Proposals for revised IWG 2018 hematological response criteria in patients with MDS included in clinical trials. *Blood*; 133(10):1020-30.
54. Cheson BD, Bennett JM, Kantarjian H, Pinto A, Schiffer CA, Nimer SD, et al. (2000): Report of an international working group to standardize response criteria for myelodysplastic syndromes. *Blood*; 96(12):3671-4.
55. European Medicines Agency (EMA) (2013): Assessment report. Invented name: Revlimid. Procedure No. EMEA/H/C/000717/II/0056. [Zugriff: 24.04.2019]. URL: [https://www.ema.europa.eu/en/documents/variation-report/revlimid-h-c-717-ii-0056-epar-assessment-report-variation\\_en.pdf](https://www.ema.europa.eu/en/documents/variation-report/revlimid-h-c-717-ii-0056-epar-assessment-report-variation_en.pdf).

56. Pileggi C, Di Sanzo M, Mascaro V, Marafioti MG, Costanzo FS, Pavia M (2017): Role of serum ferritin level on overall survival in patients with myelodysplastic syndromes: Results of a meta-analysis of observational studies. *PLoS One*; 12(6):e0179016.
57. Pinchon DJ, Stanworth SJ, Doree C, Brunskill S, Norfolk DR (2009): Quality of life and use of red cell transfusion in patients with myelodysplastic syndromes. A systematic review. *Am J Hematol*; 84(10):671-7.
58. Gemeinsamer Bundesausschuss (G-BA) (2019): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 Verfo. Wirkstoff: Brentuximab Vedotin (nAWG). Datum der Veröffentlichung: 17. Juni 2019. [Zugriff: 02.07.2019]. URL: [https://www.g-ba.de/downloads/92-975-3004/2019-03-15\\_Nutzenbewertung-G-BA\\_Brentuximab-Vediotin\\_D-449.pdf](https://www.g-ba.de/downloads/92-975-3004/2019-03-15_Nutzenbewertung-G-BA_Brentuximab-Vediotin_D-449.pdf).
59. Gemeinsamer Bundesausschuss (G-BA) (2019): Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 Verfo. Wirkstoff: Blinatumomab (AWG C). Datum der Veröffentlichung: 15. Mai 2019. [Zugriff: 02.07.2019]. URL: [https://www.g-ba.de/downloads/92-975-2967/2019-02-15\\_Nutzenbewertung-G-BA\\_Blinatumomab-D-429.pdf](https://www.g-ba.de/downloads/92-975-2967/2019-02-15_Nutzenbewertung-G-BA_Blinatumomab-D-429.pdf).
60. Gemeinsamer Bundesausschuss (G-BA) (2018): Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Telotristatethyl (Karzinoid-Syndrom-bedingte Diarrhö). Vom 5. April 2018. [Zugriff: 02.07.2019]. URL: [https://www.g-ba.de/downloads/39-261-3276/2018-04-05\\_AM-RL-XII\\_Telotristatethyl\\_D-318\\_BAnz.pdf](https://www.g-ba.de/downloads/39-261-3276/2018-04-05_AM-RL-XII_Telotristatethyl_D-318_BAnz.pdf).
61. Gemeinsamer Bundesausschuss (G-BA) (2018): Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Telotristatethyl. Vom 5. April 2018. [Zugriff: 02.07.2019]. URL: [https://www.g-ba.de/downloads/40-268-4915/2018-04-05\\_AM-RL-XII\\_Telotristatethyl\\_D-318\\_TrG.pdf](https://www.g-ba.de/downloads/40-268-4915/2018-04-05_AM-RL-XII_Telotristatethyl_D-318_TrG.pdf).
62. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, et al. (2001): The EORTC QLQ-C30 Scoring Manual (3<sup>rd</sup> Edition). [Zugriff: 24.04.2019]. URL: <https://www.eortc.org/app/uploads/sites/2/2018/02/SCmanual.pdf>.
63. Osoba D, Rodrigues G, Myles J, Zee B, Pater J (1998): Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*; 16(1):139-44.
64. Oliva EN, Nobile F, Dimitrov BD (2013): Development and validation of QOL-E© instrument for the assessment of health-related quality of life in myelodysplastic syndromes. *Central European Journal of Medicine*; 8(6):835-44.
65. U.S. Department of Health and Human Services (2009): Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0. [Zugriff: 15.08.2019]. URL: [https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE\\_4.03/CTCAE\\_4.03\\_2010-06-14\\_QuickReference\\_8.5x11.pdf](https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf).
66. Kantarjian H, O'Brien S, Ravandi F, Cortes J, Shan J, Bennett JM, et al. (2008): Proposal for a new risk model in myelodysplastic syndrome that accounts for events not considered in the original International Prognostic Scoring System. *Cancer*; 113(6):1351-61.

67. European Medicines Agency (EMA) (2019): Guideline on the investigation of subgroups in confirmatory clinical trials. [Zugriff: 25.04.2019]. URL: [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-subgroups-confirmatory-clinical-trials\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-subgroups-confirmatory-clinical-trials_en.pdf).
68. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) (2017): ICH Harmonised Guideline. General principles for planning and design of multi-regional clinical trials. E17. Final version. [Zugriff: 12.05.2020]. URL: [https://database.ich.org/sites/default/files/E17EWG\\_Step4\\_2017\\_1116.pdf](https://database.ich.org/sites/default/files/E17EWG_Step4_2017_1116.pdf).
69. Park S, Grabar S, Kelaidi C, Beyne-Rauzy O, Picard F, Bardet V, et al. (2008): Predictive factors of response and survival in myelodysplastic syndrome treated with erythropoietin and G-CSF: the GFM experience. *Blood*; 111(2):574-82.
70. Jädersten M, Malcovati L, Dybedal I, Porta MGD, Invernizzi R, Montgomery SM, et al. (2008): Erythropoietin and Granulocyte-Colony Stimulating Factor Treatment Associated With Improved Survival in Myelodysplastic Syndrome. *Journal of Clinical Oncology*; 26(21):3607-13.
71. Park S, Kelaidi C, Meunier M, Casadevall N, Gerds AT, Platzbecker U (2019): The prognostic value of serum erythropoietin in patients with lower-risk myelodysplastic syndromes: a review of the literature and expert opinion. *Ann Hematol*;
72. Pfizer (2007): Retacrit™ Injektionslösung in Fertigspritze; Fachinformation. Stand: August 2019 [Zugriff: 06.05.2020]. URL: <http://www.fachinfo.de>.
73. Janssen (2008): Erypo® FS 1000 I.E./0,5 ml/ 2000 I.E./0,5 ml/ 3000 I.E./0,3 ml/ 4000 I.E./0,4 ml/ 10 000 I.E./ml/ 40 000 I.E./ml; Fachinformation. Stand: März 2019 [Zugriff: 06.05.2020]. URL: <http://www.fachinfo.de>.
74. Celgene (2015): ACE-536-MDS-001 - A Phase 3, Double-blind, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) Versus Placebo for the Treatment of Anemia Due to the IPSS-R Very Low, Low, or Intermediate Risk Myelodysplastic Syndromes in Subjects With Ring Sideroblasts Who Require Red Blood Cell Transfusions. (NCT02631070). Stand des Eintrags: 23.03.2020. [Zugriff: 13.05.2020]. URL: <https://clinicaltrials.gov/ct2/show/record/NCT02631070?term=luspatercept+OR+ACE-536+OR+%22ACE-536%22+OR+ACE536&draw=1&rank=5>
75. Celgene Corporation (2015): ACE-536-MDS-001 - A Phase 3, Double-Blind, Randomized Study To Compare The Efficacy And Safety Of Luspatercept (ACE-536) Versus Placebo For The Treatment Of Anemia Due To IPSS-R Very Low, Low, Or Intermediate Risk Myelodysplastic Syndromes In Subjects With Ring Syderoblasts Who Require Red Blood Cell Transfusions (2015-003454-41). [Zugriff: 13.05.2020]. URL: <https://www.clinicaltrialsregister.eu/ctr-search/trial/2015-003454-41/DE>
76. Celgene Corporation (2015): ACE-536-MDS-001 - A Phase 3, Double-Blind, Randomized Study To Compare The Efficacy And Safety Of Luspatercept (ACE-536) Versus Placebo For The Treatment Of Anemia Due To IPSS-R Very Low, Low, Or Intermediate Risk Myelodysplastic Syndromes In Subjects With Ring Syderoblasts Who Require Red Blood Cell Transfusions - MEDALIST (EUCTR2015-003454-41). Stand des Eintrags: 16.12.2019. [Zugriff: 29.04.2019]. URL: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003454-41-DE>

77. Celgene (2015): ACE-536-MDS-001 - A Phase 3, Double-blind, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) Versus Placebo for the Treatment of Anemia Due to the IPSS-R Very Low, Low, or Intermediate Risk Myelodysplastic Syndromes in Subjects With Ring Sideroblasts Who Require Red Blood Cell Transfusions. - WHO ICTRP (NCT02631070). Stand des Eintrags: 30.03.2020. [Zugriff: 29.04.2019]. URL: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02631070>
78. Celgene Corporation United States (XXXX): ACE-536-MDS-001 - A Phase 3, Double-Blind, Randomized Study To Compare The Efficacy And Safety Of Luspatercept (ACE-536) Versus Placebo For The Treatment Of Anemia Due To IPSS-R Very Low, Low, Or Intermediate Risk Myelodysplastic Syndromes In Subjects With Ring Syderoblasts Who Require Red Blood Cell Transfusions (2015-003454-41). [Zugriff: 18.05.2020]. URL: <https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html>
79. Fenaux P, Platzbecker U, Mufti GJ, Garcia-Manero G, Buckstein R, Santini V, et al. (2020): Luspatercept in Patients with Lower-Risk Myelodysplastic Syndromes. *New England Journal of Medicine*; 382(2):140-51.
80. Celgene (2018): ACE-536-MDS-002 - Efficacy and Safety Study of Luspatercept (ACE-536) Versus Epoetin Alfa for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in ESA Naïve Subjects Who Require Red Blood Cell Transfusions (COMMANDS) (NCT03682536). Stand des Eintrags: 13.05.2020. [Zugriff: 18.05.2020]. URL: <https://clinicaltrials.gov/ct2/show/record/NCT03682536?view=record>
81. Celgene Corporation (2018): ACE-536-MDS-002 - A Phase 3, Open -label, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) versus Epoetin Alfa for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in ESA Naïve Subjects Who Require Red Blood Cell Transfusions. (2017-003190-34). [Zugriff: 18.05.2020]. URL: [https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract\\_number:2017-003190-34](https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-003190-34)
82. Celgene Corporation (2018): ACE-536-MDS-002 - A Phase 3, Open -label, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) versus Epoetin alfa for the Treatment of Anemia due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in ESA Naive Subjects who require Red Blood Cell Transfusions. - The Commands Trial - WHO ICTRP (2017-003190-34). Stand des Eintrags: 16.12.2019. [Zugriff: 18.05.2020]. URL: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2017-003190-34-NL>
83. Celgene Corporation (2018): ACE-536-MDS-002 - A Phase 3, Open -label, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) versus Epoetin Alfa for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in ESA Naive Subjects Who Require Red Blood Cell Transfusions. - The Commands Trial - ICTRP (2017-003190-34). Stand des Eintrags: 01.02.2020. [Zugriff: 18.05.2020]. URL: <http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2017-003190-34-DE>

84. Celgene Corporation United States (XXXX): ACE-536-MDS-002 - A Phase 3, Open - label, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) versus Epoetin Alfa for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in ESA Naïve Subjects Who Require Red Blood Cell Transfusions. (2017-003190-34). [Zugriff: 18.05.2020]. URL: <https://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.html>
85. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. (2016): The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*; 127(20):2391-405.
86. Greenberg P, Cox C, LeBeau MM, Fenau P, Morel P, Sanz G, et al. (1997): International Scoring System for Evaluating Prognosis in Myelodysplastic Syndromes. *Blood*; 89(6):2079-88.
87. de Swart L, Smith A, Johnston TW, Haase D, Droste J, Fenau P, et al. (2015): Validation of the revised international prognostic scoring system (IPSS-R) in patients with lower-risk myelodysplastic syndromes: a report from the prospective European LeukaemiaNet MDS (EUMDS) registry. *Br J Haematol*; 170(3):372-83.
88. Castelli R, Schiavon R, Rossi V, Deliliers GL (2018): Management of anemia in low-risk myelodysplastic syndromes treated with erythropoiesis-stimulating agents newer and older agents. *Med Oncol*; 35(5):76.
89. Kubasch AS, Platzbecker U (2019): Setting Fire to ESA and EMA Resistance: New Targeted Treatment Options in Lower Risk Myelodysplastic Syndromes. *Int J Mol Sci*; 20(16)
90. Hellström-Lindberg E, Negrin R, Stein R, Krantz S, Lindberg G, Vardiman J, et al. (1997): Erythroid response to treatment with G-CSF plus erythropoietin for the anaemia of patients with myelodysplastic syndromes: proposal for a predictive model. *British Journal of Haematology*; 99(2):344-51.
91. Germing U, Strupp C, Kundgen A, Bowen D, Aul C, Haas R, et al. (2004): No increase in age-specific incidence of myelodysplastic syndromes. *Haematologica*; 89(8):905-10.
92. Gattermann N, Kundgen A, Kellermann L, Zeffel M, Paessens B, Germing U (2013): The impact of age on the diagnosis and therapy of myelodysplastic syndromes: results from a retrospective multicenter analysis in Germany. *Eur J Haematol*; 91(6):473-82.
93. Neukirchen J, Schoonen WM, Strupp C, Gattermann N, Aul C, Haas R, et al. (2011): Incidence and prevalence of myelodysplastic syndromes: data from the Dusseldorf MDS-registry. *Leuk Res*; 35(12):1591-6.
94. Celgene GmbH (2019): Luspatercept MEDALIST Post-hoc Analysen - Datenschnitt 07.01.2019.
95. Celgene (2019): Health-related quality of life report Luspatercept (ACE-536). ACE-536-MDS-001. A phase 3, double-blind, randomized study to compare the efficacy and safety of Luspatercept (ACE-536) versus placebo for the treatment of anemia due to IPSS-R very low, low or intermediate risk myelodysplastic syndromes in subjects with ring sideroblasts who require red blood cell transfusions. The "Medalist" trial.
96. Cocks K, King MT, Velikova G, de Castro G, Jr., Martyn St-James M, Fayers PM, et al. (2012): Evidence-based guidelines for interpreting change scores for the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *Eur J Cancer*; 48(11):1713-21.

97. Wong SS, Wilczynski NL, Haynes RB (2006): Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. *J Med Libr Assoc*; 94(4):451-5.

**Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche**

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

<b>Datenbankname</b>	EMBASE	
<b>Suchoberfläche</b>	Ovid	
<b>Datum der Suche</b>	07.11.2016	
<b>Zeitsegment</b>	1974 to 2016 November 04	
<b>Suchfilter</b>	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle <sup>22</sup> ] – Strategy minimizing difference between sensitivity and specificity	
<b>#</b>	<b>Suchbegriffe</b>	<b>Ergebnis</b>
1	Diabetes Mellitus/	552986
2	Non Insulin Dependent Diabetes Mellitus/	195234
3	(diabet* or niddm or t2dm).ab,ti.	714228
4	or/1-3	847068
5	linagliptin*.mp.	1562
6	(random* or double-blind*).tw.	1193849
7	placebo*.mp.	388057
8	or/6-7	1382838
9	and/4,5,8	633

<sup>22</sup> Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

**Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel****Suche nach RCT mit dem zu bewertenden Arzneimittel in MEDLINE**

<b>Datenbankname</b>	MEDLINE	
<b>Suchoberfläche</b>	Ovid MEDLINE(R) 1946 to May Week 2 2020, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations 1946 to May 15, 2020, Ovid MEDLINE(R) Daily Update May 15, 2020	
<b>Datum der Suche</b>	18.05.2020	
<b>Zeitsegment</b>	Keine Einschränkung	
<b>Suchfilter</b>	Filter für RCT nach Wong et al. 2006 [97]	
<b>#</b>	<b>Suchbegriffe</b>	<b>Ergebnis</b>
1	Luspatercept*.mp. <sup>a</sup>	40
2	reblozyl*.mp.	1
3	„ACE-536“.mp.	9
4	„ACE536“.mp.	0
5	„ACE 536“.mp.	9
6	1 or 2 or 3 or 4 or 5	45
7	randomized controlled trial.pt.	505639
8	randomized.mp.	827537
9	placebo.mp.	211536
10	7 or 8 or 9	889481
11	6 and 10	5
<sup>a</sup> für Luspatercept lag kein MeSH-Term vor RCT = randomisiert kontrollierte Studie ( <i>randomized controlled trial</i> )		

**Suche nach RCT mit dem zu bewertenden Arzneimittel in EMBASE**

<b>Datenbankname</b>	EMBASE	
<b>Suchoberfläche</b>	Ovid EMBASE 1974 to 2020 May 15	
<b>Datum der Suche</b>	18.05.2020	
<b>Zeitsegment</b>	Keine Einschränkung	
<b>Suchfilter</b>	Filter für RCT nach Wong et al. 2006 [97]	
<b>#</b>	<b>Suchbegriffe</b>	<b>Ergebnis</b>
1	Exp Luspatercept/	187
2	Luspatercept*.mp. <sup>a</sup>	193
3	reblozyl*.mp.	1
4	„ACE-536“.mp.	94
5	„ACE536“.mp.	0
6	„ACE 536“.mp.	94
7	1 or 2 or 3 or 4 or 5 or 6	217
8	Random*.tw.	1532608
9	Placebo*.mp.	452808
10	Double-blind*.tw.	208828
11	8 or 9 or 10	1786653
12	7 and 11	35
RCT = randomisiert kontrollierte Studie ( <i>randomized controlled trial</i> )		

**Suche nach RCT mit dem zu bewertenden Arzneimittel in der Cochrane Library**

<b>Datenbankname</b>	Cochrane Central Register of Controlled Trials	
<b>Suchoberfläche</b>	The Cochrane Library	
<b>Datum der Suche</b>	18.05.2020	
<b>Zeitsegment</b>	Keine Einschränkung	
<b>Suchfilter</b>	Kein Suchfilter verwendet	
<b>#</b>	<b>Suchbegriffe</b>	<b>Ergebnis</b>
1	(luspatercept*):ti,ab,kw <sup>a</sup>	28
2	(reblozyl*):ti,ab,kw	0
3	(„ACE-536“):ti,ab,kw OR („ACE536“):ti,ab,kw OR („ACE 536“):ti,ab,kw	14
4	#1 OR #2 OR #3	29
<sup>a</sup> für Luspatercept lag kein MeSH-Term vor		

**Anhang 4-A2: Suche nach RCT für indirekte Vergleiche**

Nicht zutreffend.

**Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien**

Nicht zutreffend.

**Anhang 4-A4: Suche nach weiteren Untersuchungen**

Nicht zutreffend.

**Anhang 4-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken**

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. clinicaltrials.gov), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

<b>Studienregister/ Studienergebnisdatenbank</b>	International Clinical Trials Registry Platform Search Portal
<b>Internetadresse</b>	<a href="http://apps.who.int/trialsearch/">http://apps.who.int/trialsearch/</a>
<b>Datum der Suche</b>	07.11.2016
<b>Eingabeoberfläche</b>	Standard Search
<b>Suchstrategie</b>	linagliptin OR BI 1356
<b>Treffer</b>	169

**Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel****Suche nach RCT mit dem zu bewertenden Arzneimittel in Clinicaltrials.gov**

<b>Studienregister/ Studienergebnisdatenbank</b>	clinicaltrials.gov
<b>Internetadresse</b>	<a href="http://www.clinicaltrials.gov">http://www.clinicaltrials.gov</a>
<b>Datum der Suche</b>	18.05.2020
<b>Suchoberfläche</b>	Standard Search
<b>Suchstrategie</b>	luspatercept OR reblozyl OR ACE-536
<b>Treffer</b>	14

**Suche nach RCT mit dem zu bewertenden Arzneimittel im EU Clinical Trials Register (EU-CTR)**

<b>Studienregister/ Studienergebnisdatenbank</b>	EU Clinical Trials Register (EU-CTR)
<b>Internetadresse</b>	<a href="https://www.clinicaltrialsregister.eu/">https://www.clinicaltrialsregister.eu/</a>
<b>Datum der Suche</b>	18.05.2020
<b>Suchoberfläche</b>	Basic Search <sup>(1)</sup>
<b>Suchstrategie</b>	luspatercept* OR reblozyl* OR ACE-536
<b>Treffer</b>	11
(1) Keine Verwendung der Advanced Search bzw. Search tools, um die Suche nicht einzuschränken	

**Suche nach RCT mit dem zu bewertenden Arzneimittel im ICTRP Search Portal der WHO**

<b>Studienregister/ Studienergebnisdatenbank</b>	ICTRP Search Portal der WHO
<b>Internetadresse</b>	<a href="http://apps.who.int/trialsearch/">http://apps.who.int/trialsearch/</a>
<b>Datum der Suche</b>	18.05.2020
<b>Suchoberfläche</b>	Standard Search
<b>Suchstrategie</b>	luspatercept OR reblozyl OR ACE-536
<b>Treffer</b>	63 Einträge zu 30 Studien

**Anhang 4-B2: Suche nach RCT für indirekte Vergleiche**

Nicht zutreffend.

**Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien**

Nicht zutreffend.

**Anhang 4-B4: Suche nach weiteren Untersuchungen**

Nicht zutreffend.

**Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)**

*Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der /den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.*

**Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel**

Nicht zutreffend.

**Anhang 4-C2: Suche nach RCT für indirekte Vergleiche**

Nicht zutreffend.

**Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien**

Nicht zutreffend.

**Anhang 4-C4: Suche nach weiteren Untersuchungen**

Nicht zutreffend.

#### Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken)

Listen Sie nachfolgend die durch die Studienregistersuche(n)/ Studienergebnisdatenbanksuche(n) identifizierten, aber ausgeschlossenen Registereinträge auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

#### Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nr.	Referenz	Ausschlussgrund
1	Celgene (2019): ACE-536-B-THAL-004 U1111-1241-4168 (Other Identifier: WHO) 2019-000208-13 (EudraCT Number) - A Phase 2A Open-label Multicenter Study to Evaluate the Safety and Pharmacokinetics of Luspatercept in Pediatric Subjects Who Require Regular Red Blood Cell Transfusions Due to B-Thalassemia - ClinicalTrials.gov (NCT04143724). Stand des Eintrags: 20.12.2019. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT04143724">https://ClinicalTrials.gov/ct2/show/record/NCT04143724</a>	A1
2	Celgene (2019): ACE-536-LTFU-001 U1111-1235-8123 (Registry Identifier: WHO) 2018-002915-93 (EudraCT Number) - A Phase 3b, Open Label, Single-arm Rollover Study to Evaluate Long Term Safety in Subjects Who Have Participated in Other Luspatercept (Ace-536) Clinical Trials. - ClinicalTrials.gov (NCT04064060). Stand des Eintrags: 30.04.2020. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT04064060">https://ClinicalTrials.gov/ct2/show/record/NCT04064060</a>	A4
3	Celgene (2019): ACE-536-MDS-003 U1111-1224-6268 (Registry Identifier: WHO) - A Phase 2, Multicenter, Single-arm Study to Evaluate the Efficacy, Pharmacokinetics, and Safety of Luspatercept (ACE-536) for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in Japanese Subjects Who Are Not Requiring Red Blood Cell Transfusion - ClinicalTrials.gov (NCT03900715). Stand des Eintrags: 14.11.2019. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT03900715">https://ClinicalTrials.gov/ct2/show/record/NCT03900715</a>	A1
4	Celgene (2018): ACE-536-MDS-002 U1111-1218-1810 (Registry Identifier: WHO) 2017-003190-34 (EudraCT Number) - A Phase 3, Open-label, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) Versus Epoetin Alpha for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Due to Myelodysplastic Syndromes (MDS) ESA in Native Subjects Who Require Red Blood Cell Transfusions - ClinicalTrials.gov (NCT03682536). Stand des Eintrags: 13.05.2020. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT03682536">https://ClinicalTrials.gov/ct2/show/record/NCT03682536</a>	kein Ausschlussgrund/ eingeschlossen
5	Celgene (2017): ACE-536-B-THAL-002 U1111-1202-7068 (Registry Identifier: WHO) 2015-003225-33 (EudraCT Number) - A Phase 2, Double-Blind, Placebo Controlled Multicenter Study to Determine the Efficacy and Safety of Luspatercept (ACE-536) in Adults With Non-Transfusion Dependent Beta (B)-Thalassemia - ClinicalTrials.gov (NCT03342404). Stand des Eintrags: 20.04.2020. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT03342404">https://ClinicalTrials.gov/ct2/show/record/NCT03342404</a>	A1

Nr.	Referenz	Ausschlussgrund
6	Celgene (2017): ACE-536-MF-001 2017-000322-35 (EudraCT Number) U1111-1197-1202 (Registry Identifier: WHO) - A Phase-2 Study To Determine Efficacy and Safety of Luspatercept in Subjects With Myeloproliferative Neoplasm-Associated Myelofibrosis and Anemia With or Without Red Blood Cell-Transfusion Dependence - ClinicalTrials.gov (NCT03194542). Stand des Eintrags: 27.02.2020. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT03194542">https://ClinicalTrials.gov/ct2/show/record/NCT03194542</a>	A1
7	Celgene (2015): ACE-536-B-THAL-001 - A Phase 3, Double-Blind, Placebo Controlled Multicenter Study to Determine the Efficacy and Safety of Luspatercept (ACE-536) in Adults With Transfusion Dependent Beta (B)-Thalassemia - ClinicalTrials.gov (NCT02604433). Stand des Eintrags: 27.01.2020. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT02604433">https://ClinicalTrials.gov/ct2/show/record/NCT02604433</a>	BELIEVE
8	Celgene (2015): ACE-536-MDS-001 - A Phase 3, Double-blind, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) Versus Placebo for the Treatment of Anemia Due to the IPSS-R Very Low, Low, or Intermediate Risk Myelodysplastic Syndromes in Subjects With Ring Sideroblasts Who Require Red Blood Cell Transfusions. - ClinicalTrials.gov (NCT02631070). Stand des Eintrags: 23.03.2020. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT02631070">https://ClinicalTrials.gov/ct2/show/record/NCT02631070</a>	kein Ausschlussgrund/ eingeschlossen
9	Acceleron Pharma, Inc. (2014): A536-05 2014-001280-13 (EudraCT Number) - An Open-Label Extension Study to Evaluate the Long-Term Effects of ACE-536 for the Treatment of Anemia in Patients With Low or Intermediate-1 Risk Myelodysplastic Syndromes (MDS) Previously Enrolled in Study A536-03 - ClinicalTrials.gov (NCT02268383). Stand des Eintrags: 10.06.2019. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT02268383">https://ClinicalTrials.gov/ct2/show/record/NCT02268383</a>	A4
10	Acceleron Pharma, Inc. (2014): A536-06 - An Open-Label Extension Study to Evaluate the Long-Term Effects of ACE-536 in Patients With $\beta$ -Thalassemia Previously Enrolled in Study A536-04 - ClinicalTrials.gov (NCT02268409). Stand des Eintrags: 07.01.2020. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT02268409">https://ClinicalTrials.gov/ct2/show/record/NCT02268409</a>	A4
11	Acceleron Pharma, Inc. (2012): A536-04 2012-002499-15 (EudraCT Number) - A Phase 2, Open-Label, Ascending Dose Study to Evaluate the Effects of ACE-536 in Patients With Beta-Thalassemia Intermedia - ClinicalTrials.gov (NCT01749540). Stand des Eintrags: 14.12.2016. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT01749540">https://ClinicalTrials.gov/ct2/show/record/NCT01749540</a>	A4
12	Acceleron Pharma, Inc. (2012): A536-03 - A Phase 2, Open Label, Ascending Dose Study of ACE-536 for the Treatment of Anemia in Patients With Low or Intermediate-1 Risk Myelodysplastic Syndromes (MDS) - ClinicalTrials.gov (NCT01749514). Stand des Eintrags: 11.07.2019. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT01749514">https://ClinicalTrials.gov/ct2/show/record/NCT01749514</a>	A4
13	Acceleron Pharma, Inc. (2011): A536-02 - A Phase 1 Randomized, Double-Blind, Placebo-Controlled, Multiple Ascending Dose Study to Evaluate the Safety, Tolerability, Pharmacokinetics, and Pharmacodynamic Effects of ACE-536 in Healthy Postmenopausal Women - ClinicalTrials.gov (NCT01432717). Stand des Eintrags: 07.11.2012. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT01432717">https://ClinicalTrials.gov/ct2/show/record/NCT01432717</a>	A1
14	Celgene (2015): ACE-536-B-THAL-003 - Burden of Beta-Thalassemia - Quality of Life and Health Care Resource Utilization- A Prospective Observational Study - ClinicalTrials.gov (NCT02626689). Stand des Eintrags: 07.04.2017. [Zugriff: 18.05.2020]. URL: <a href="https://ClinicalTrials.gov/ct2/show/record/NCT02626689">https://ClinicalTrials.gov/ct2/show/record/NCT02626689</a>	A4

Nr.	Referenz	Ausschlussgrund
15	Acceleron Pharma Inc. (2014): A536-05 - An Open-Label Extension Study to Evaluate the Long-Term Effects of ACE-536 for the Treatment of Anemia in Patients with Low or Intermediate-1 Risk Myelodysplastic Syndromes (MDS) Previously Enrolled in Study A536-03 - EU-CTR (2014-001280-13). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-001280-13">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-001280-13</a>	A4
16	Acceleron Pharma Inc. (2012): A536-03 - Myelodysplastic Syndromes (MDS) - EU-CTR (2012-002523-14). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-002523-14">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-002523-14</a>	A4
17	Celgene Corporation (2019): ACE-536-B-THAL-004 - A phase 2a study to evaluate the safety and pharmacokinetics of Luspatercept (ACE-536) in paediatric subjects who require regular red blood cell transfusions due to beta thalassemia - EU-CTR (2019-000208-13). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2019-000208-13">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2019-000208-13</a>	A1
18	Celgene Corporation (2019): ACE-536-LTFU-001 - A phase 3B, open-label, single-arm, rollover study to evaluate long-term safety in subjects who have participated in other Luspatercept (ACE-536) clinical trials. - EU-CTR (2018-002915-93). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-002915-93">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2018-002915-93</a>	A4
19	Celgene Corporation (2017): ACE-536-B-THAL-002 - A Phase 2, Double-Blind, Randomized, Placebo-Controlled, Multicenter Study to Determine the Efficacy and Safety of Luspatercept (ACE-536) versus Placebo in Adults with Non Transfusion Dependent Beta ( $\beta$ )-Thalassemia - EU-CTR (2015-003225-33). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003225-33">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003225-33</a>	A1
20	Celgene Corporation (2015): ACE-536-B-THAL-001 - A Phase 3, Double-Blind, Randomized, Placebo-Controlled, Multicenter Study to Determine the Efficacy and Safety of Luspatercept (ACE-536) Versus Placebo in Adults Who Require Regular Red Blood Cell Transfusions Due to Beta-Thalassemia - EU-CTR (2015-003224-31). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003224-31">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003224-31</a>	BELIEVE
21	Celgene Corporation (2017): ACE-536-MF-001 - A Phase 2, Multicenter, Open-Label Study to Evaluate the Efficacy and Safety of Luspatercept (ACE-536) in Subjects With Myeloproliferative Neoplasm-Associated Myelofibrosis and Anemia With and Without Red Blood Cell-Transfusion Dependence - EU-CTR (2017-000322-35). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-000322-35">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-000322-35</a>	A1
22	Celgene Corporation (2015): ACE-536-MDS-001 - A Phase 3, Double-Blind, Randomized Study To Compare The Efficacy And Safety Of Luspatercept (ACE-536) Versus Placebo For The Treatment Of Anemia Due To IPSS-R Very Low, Low, Or Intermediate Risk Myelodysplastic Syndromes In Subjects With Ring Syderoblasts Who Require Red Blood Cell Transfusions - EU-CTR (2015-003454-41). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003454-41">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-003454-41</a>	kein Ausschlussgrund/geschlossen

Nr.	Referenz	Ausschlussgrund
23	Celgene Corporation (2018): ACE-536-MDS-002 - A Phase 3, Open -label, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) versus Epoetin Alfa for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in ESA Naïve Subjects Who Require Red Blood Cell Transfusions. - EU-CTR (2017-003190-34). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-003190-34">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-003190-34</a>	kein Ausschlussgrund/ eingeschlossen
24	ACCELERON PHARMA INC (2012): A536-04 - A Phase 2, Open-Label, Ascending Dose Study to Evaluate the Effects of ACE-536 in Patients with $\beta$ -Thalassemia Intermedia - EU-CTR (2012-002499-15). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-002499-15">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-002499-15</a>	A4
25	Acceleron Pharma, Inc. (2014): A536-06 - An Open-Label Extension Study to Evaluate the Long-Term Effects of ACE-536 in Patients with $\beta$ -Thalassemia Previously Enrolled in Study A536-04 - EU-CTR (2014-001281-94). [Zugriff: 18.05.2020]. URL: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-001281-94">https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-001281-94</a>	A4
26	Acceleron Pharma, Inc. (2014): A536-06 - An Open-Label Extension Study to Evaluate the Long-Term Effects of ACE-536 in Patients with $\beta$ -Thalassemia Previously Enrolled in Study A536-04 - WHO ICTRP (2014-001281-94). Stand des Eintrags: 01.11.2014. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-001281-94-GR">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-001281-94-GR</a>	A4
27	ACCELERON PHARMA INC (2012): A536-04 - A Phase 2, Open-Label, Ascending Dose Study to Evaluate the Effects of ACE-536 in Patients with $\beta$ -Thalassemia Intermedia - WHO ICTRP (2012-002499-15). Stand des Eintrags: 01.04.2016. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-002499-15-IT">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-002499-15-IT</a>	A4
28	Acceleron Pharma, Inc. (2014): A536-06 - An Open-Label Extension Study to Evaluate the Long-Term Effects of ACE-536 in Patients with $\beta$ -Thalassemia Previously Enrolled in Study A536-04 - WHO ICTRP (2014-001281-94). Stand des Eintrags: 01.07.2015. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-001281-94-IT">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-001281-94-IT</a>	A4
29	Celgene Corporation (2017): ACE-536-B-THAL-002 - A Phase 2, Double-Blind, Randomized, Placebo-Controlled, Multicenter Study to Determine the Efficacy and Safety of Luspatercept (ACE-536) versus Placebo in Adults with Non Transfusion Dependent Beta ( $\beta$ )-Thalassemia - The BEYOND Study - WHO ICTRP (2015-003225-33). Stand des Eintrags: 01.12.2017. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003225-33-GR">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003225-33-GR</a>	A1
30	Celgene (2015): ACE-536-B-THAL-003 - To Document the Burden of Illness on the Quality of Life and the Impact on Healthcare Utilization in (Beta) $\beta$ -thalassemia Subjects Who Are Transfusion Dependent (TD) and Non-transfusion Dependent (NTD) Receiving Standard of Care (NCT02626689). Stand des Eintrags: 16.12.2017. [Zugriff: 18.05.2020]. URL: <a href="https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02626689">https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02626689</a>	A4
31	Celgene K.K (2019): ACE-536-MDS-003 - An Efficacy and Safety Study of Luspatercept (ACE-536) for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in Japanese Subjects Who Are Not Requiring Red Blood Cell Transfusion (NCT03900715). Stand des Eintrags: 25.11.2019. [Zugriff: 18.05.2020]. URL: <a href="https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03900715">https://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03900715</a>	A1



Nr.	Referenz	Ausschlussgrund
41	Celgene (2015): ACE-536-MDS-001 - A Phase 3, Double-blind, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) Versus Placebo for the Treatment of Anemia Due to the IPSS-R Very Low, Low, or Intermediate Risk Myelodysplastic Syndromes in Subjects With Ring Sideroblasts Who Require Red Blood Cell Transfusions. - WHO ICTRP (NCT02631070). Stand des Eintrags: 03.03.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02631070">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02631070</a>	kein Ausschlussgrund/ eingeschlossen
42	Celgene (2019): 2019-000208-13 ACE-536-B-THAL-004 U1111-1241-4168 - A Phase 2A Open-label Multicenter Study to Evaluate the Safety and Pharmacokinetics of Luspatercept in Pediatric Subjects Who Require Regular Red Blood Cell Transfusions Due to B-Thalassemia - WHO ICTRP (NCT04143724). Stand des Eintrags: 06.01.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04143724">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04143724</a>	A1
43	Celgene (2015): ACE-536-B-THAL-001 - A Phase 3, Double-Blind, Placebo Controlled Multicenter Study to Determine the Efficacy and Safety of Luspatercept (ACE-536) in Adults With Transfusion Dependent Beta (B)-Thalassemia - WHO ICTRP (NCT02604433). Stand des Eintrags: 01.02.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02604433">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02604433</a>	BELIEVE
44	Celgene Corporation (2019): IND 112562 - A Phase 3B, open-label, single-arm, rollover study to evaluate long-term safety in subjects who have participated in other luspatercept (ACE-536) clinical trials - WHO ICTRP (LBCTR2019100218). Stand des Eintrags: 02.04.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=LBCTR2019100218">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=LBCTR2019100218</a>	A4
45	Celgene Corporation (2018): 2017-003190-34-GB ACE-536-MDS-002 - A Phase 3, Open -label, Randomized Study to Compare the Efficacy and Safety of Luspatercept (ACE-536) versus Epoetin Alfa for the Treatment of Anemia Due to IPSS-R Very Low, Low or Intermediate Risk Myelodysplastic Syndromes (MDS) in ESA Na <sup>+</sup> -ve Subjects Who Require Red Blood Cell Transfusions. - The Commands Trial - WHO ICTRP (2017-003190-34). Stand des Eintrags: 01.02.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2017-003190-34-DE">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2017-003190-34-DE</a>	kein Ausschlussgrund/ eingeschlossen
46	Celgene (2017): 2017-000322-35 ACE-536-MF-001 U1111-1197-1202 - A Phase-2 Study To Determine Efficacy and Safety of Luspatercept in Subjects With Myeloproliferative Neoplasm-Associated Myelofibrosis and Anemia With or Without Red Blood Cell-Transfusion Dependence - WHO ICTRP (NCT03194542). Stand des Eintrags: 09.03.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03194542">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03194542</a>	A1
47	Celgene (2017): 2015-003225-33 ACE-536-B-THAL-002 U1111-1202-7068 - A Phase 2, Double-Blind, Placebo Controlled Multicenter Study to Determine the Efficacy and Safety of Luspatercept (ACE-536) in Adults With Non-Transfusion Dependent Beta (B)-Thalassemia - WHO ICTRP (NCT03342404). Stand des Eintrags: 02.04.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03342404">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03342404</a>	A1
48	Acceleron Pharma Inc. (2014): A536-05 - An Open-Label Extension Study to Evaluate the Long-Term Effects of ACE-536 for the Treatment of Anemia in Patients with Low or Intermediate-1 Risk Myelodysplastic Syndromes (MDS) Previously Enrolled in Study A536-03 - WHO ICTRP (2014-001280-13). Stand des Eintrags: 04.05.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-001280-13-DE">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2014-001280-13-DE</a>	A4

Nr.	Referenz	Ausschlussgrund
49	Acceleron Pharma, Inc. (2013): 1704/01 2012-002523-14 EK309092012 NCT01749514 - A Phase 2, Open Label, Ascending Dose Study of ACE-536 for the Treatment of Anemia in Patients with Low or Intermediate-1 Risk Myelodysplastic Syndromes (MDS) - A536-03 - WHO ICTRP (DRKS00004531). Stand des Eintrags: 01.05.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00004531">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00004531</a>	A4
50	Acceleron Pharma, Inc. (2011): A536-02 - A Phase 1 Randomized, Double-Blind, Placebo-Controlled, Multiple Ascending Dose Study to Evaluate the Safety, Tolerability, Pharmacokinetics, and Pharmacodynamic Effects of ACE-536 in Healthy Postmenopausal Women - WHO ICTRP (NCT01432717). Stand des Eintrags: 01.02.2015. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01432717">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01432717</a>	A1
51	Celgene Corporation (2015): ACE-536-B-THAL-001 - A Phase 3, Double-Blind, Randomized, Placebo-Controlled, Multicenter Study to Determine the Efficacy and Safety of Luspatercept (ACE-536) Versus Placebo in Adults Who Require Regular Red Blood Cell Transfusions Due to Beta-Thalassemia - The Believe Trial - WHO ICTRP (2015-003224-31). Stand des Eintrags: 04.12.2018. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003224-31-GB">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003224-31-GB</a>	BELIEVE
52	Celgene Corporation (2019): ACE-536-LTFU-001 - A phase 3B, open-label, single-arm, rollover study to evaluate long-term safety in subjects who have participated in other Luspatercept (ACE-536) clinical trials. - WHO ICTRP (2018-002915-93). Stand des Eintrags: 01.12.2019. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-002915-93-SE">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2018-002915-93-SE</a>	A4
53	Celgene Corporation (2015): ACE-536-MDS-001 - A Phase 3, Double-Blind, Randomized Study To Compare The Efficacy And Safety Of Luspatercept (ACE-536) Versus Placebo For The Treatment Of Anemia Due To IPSS-R Very Low, Low, Or Intermediate Risk Myelodysplastic Syndromes In Subjects With Ring Syderoblasts Who Require Red Blood Cell Transfusions - MEDALIST - WHO ICTRP (2015-003454-41). Stand des Eintrags: 01.12.2019. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003454-41-DE">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-003454-41-DE</a>	kein Ausschlussgrund/ eingeschlossen
54	Celgene (2019): 2018-002915-93 ACE-536-LTFU-001 U1111-1235-8123 - A Phase 3b, Open Label, Single-arm Rollover Study to Evaluate Long Term Safety in Subjects Who Have Participated in Other Luspatercept (Ace-536) Clinical Trials. - WHO ICTRP (NCT04064060). Stand des Eintrags: 01.05.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04064060">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT04064060</a>	A4
55	Acceleron Pharma, Inc. (2015): 2014-001280-13 EK250062014 NCT02268383 - A Phase 2, Open-Label Extension Study to Evaluate the Long-Term Effects of ACE 536 for the Treatment of Anemia in Patients with Low or Intermediate-1 Risk Myelodysplastic Syndromes (MDS) Previously Enrolled in Study A536-03 - A536-05 - WHO ICTRP (DRKS00007625). Stand des Eintrags: 01.05.2020. [Zugriff: 18.05.2020]. URL: <a href="http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00007625">http://apps.who.int/trialsearch/Trial2.aspx?TrialID=DRKS00007625</a>	A4

**Anhang 4-D2: Suche nach RCT für indirekte Vergleiche**

Nicht zutreffend.

**Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien**

Nicht zutreffend.

**Anhang 4-D4: Suche nach weiteren Untersuchungen**

Nicht zutreffend.

**Anhang 4-E: Methodik der eingeschlossenen Studien – RCT**

*Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-116 inklusive eines Flow-Charts für den Patientenfluss.*

*Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-116 inklusive eines Flow-Charts für den Patientenfluss.*

Tabelle 4-116 (Anhang): Studiendesign und -methodik für Studie MEDALIST

Item <sup>a</sup>	Charakteristikum	Studieninformation
<b>Studienziel</b>		
<b>2 b</b>	Genaue Ziele, Fragestellung und Hypothesen	<p>Zielsetzung und Fragestellung:</p> <ul style="list-style-type: none"> <li>Untersuchung der Wirksamkeit von Luspatercept im Vergleich Placebo hinsichtlich des Erreichens einer Transfusionsunabhängigkeit bei der Behandlung von Anämie aufgrund von MDS mit sehr niedrigem, niedrigem oder intermediärem Risiko mit Ringsideroblasten und Transfusionsabhängigkeit</li> </ul> <p>Hypothesen:</p> <ul style="list-style-type: none"> <li>Nullhypothese: Die Wahrscheinlichkeit, eine Transfusionsunabhängigkeit zu erreichen, ist bei Behandlung mit Luspatercept und Placebo gleich.</li> <li>Alternativhypothese: Die Behandlung mit Luspatercept führt im Vergleich zu Placebo zu einer erhöhten Wahrscheinlichkeit, eine Transfusionsunabhängigkeit zu erreichen.</li> </ul>
<b>Methoden</b>		
<b>3</b>	Studiendesign	
<b>3a</b>	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p><b>Design:</b> zweiarmig, parallel, aktiv-kontrolliert</p> <p><b>Zuteilung der Studienteilnehmer:</b> randomisiert im Verhältnis 2:1 (Luspatercept + BSC: Placebo + BSC)</p> <p><b>Verblindung:</b> doppelblind</p> <p><b>Studiendauer:</b> mindestens 24 Wochen</p> <p><b>Studienorganisation:</b> multizentrisch</p> <p><b>Phase:</b> III</p>
<b>3b</b>	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p><b>Amendment 0.1 DE</b> (länderspezifisch; Deutschland; 31.03.2016):</p> <ul style="list-style-type: none"> <li>Konkretisierung der Protokolldefinition „hochwirksame Methoden der Kontrazeption“ zur Anpassung an die Empfehlungen zur Kontrazeption und Schwangerschaftstestung in klinischen Studien</li> <li>Aufnahme von Zusatzanforderungen hinsichtlich der Bestätigung des HIV-, Hepatitis-B-, Hepatitis-C-Status der Probanden durch das Studienzentrum vor Studienteilnahme</li> <li>Sprachliche Hinzufügung des expliziten Ausschlusses von verletzbaren Personen (inhaftierte, institutionalisierte Personen)</li> </ul> <p><b>Amendment 1</b> (21.09.2016):</p> <ul style="list-style-type: none"> <li>Hinzufügung der Überwachung von anderen malignen/prämalignen Erkrankungen als „wichtige medizinische Ereignisse“</li> <li>Hinzufügung eines Leitfadens für die Zentren zur Erhebung der Transfusionsdaten</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• Hinzufügung weitere exploratorischer Untersuchungen (z.B. CRP)</li> <li>• Einschluss der Kriterien zur Dosismodifikation und zum Therapieabbruch hinsichtlich einer Leukozytenerhöhung und Krankheitsprogression gemäß IWG-MDS-Kriterien</li> <li>• Überarbeitetes WHO-Klassifizierungssystem und Einschluss des FAB-Klassifikationssystems für die Baseline MDS-Diagnosestellung</li> <li>• Verlängerung der Erhebung von Transfusionsdaten nach Abbruch der Studienmedikation auf 16 Wochen nach letzter Dosis oder End-of-treatment-Visit (je nachdem, was später auftritt)</li> <li>• Hinzufügung von Informationen bezüglich der Bewertung des Nutzen-Risikos und Gesamt-Nutzen-Risikos der Probanden in der Studie</li> </ul> <p>Im Amendment 1 wurden einige Ein- und Ausschlusskriterien konkretisiert und überarbeitet, darunter</p> <ul style="list-style-type: none"> <li>• eine Modifizierung hinsichtlich der kontrazeptiven Maßnahmen zur Anpassung an die Definition von „hochwirksamen Methoden zur Kontrazeption“ gemäß Clinical Trial Facilitation Group Leitlinien</li> <li>• die Aufnahme des oberen Hämoglobin-Grenzwerts von 10,0 g/dl vor Transfusion als Einschlusskriterium</li> <li>• die Konkretisierung der Ausschlusskriterien, um die Studienteilnahme von Patienten mit &gt;2xULN Serumbilirubin zu ermöglichen, sofern ein bekanntes oder diagnostiziertes Gilbert Syndrom vorliegt</li> <li>• die Überarbeitung des Ausschlusskriteriums, um Patienten mit signifikanter kardialer Dysfunktion (bekannte Ejektionsfraktion &lt;35%), gemessen mittels lokalen EKGs oder MUGA innerhalb von sechs Monaten vor Randomisierung, auszuschließen</li> <li>• die Überarbeitung des Ausschlusskriteriums bezüglich erlaubter Medikationen, um den Einschluss von Patienten zu ermöglichen, die zuvor eine subtherapeutische Behandlung mit HMA oder Lenalidomid erhalten haben</li> <li>• die Überarbeitung des Einschlusskriteriums bezüglich vorheriger ESA-Therapie durch Kürzung des ESA/G-CSF/GM-CSF-Washout-Fensters vor Randomisierung von sechs auf vier Wochen</li> <li>• die Überarbeitung der Einschlusskriterien, um die Verwendung von Studienmedikationen vor Randomisierung zu erlauben</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Darüber hinaus wurden weitere kleinere Anpassungen/ Konkretisierungen im Protokoll vorgenommen und relevante Abschnitte des Protokolls anhand der Investigator's Brochure Edition 8 angepasst.</p> <p><b>Amendment 1.1 FR</b> (länderspezifisch; Frankreich; 21.09.2016):</p> <ul style="list-style-type: none"> <li>• Überarbeitung des Einschlusskriteriums bezüglich ESA, um den Grenzwert für Serum-EPO auf &gt; 500 U/l in ESA-naiven Patienten anzupassen</li> </ul> <p><b>Amendment 2</b> (09.05.2017) und <b>Amendment 2.1 FR</b> (länderspezifisch; Frankreich; 19.04.2017):</p> <ul style="list-style-type: none"> <li>• Überarbeitung der Kriterien zur Dosismodifikation und Abbruch der Studienmedikation gemäß Forderung des Steuerungsgremiums, um „Progression in eine AML oder MDS mit hohem/sehr hohem Risiko nach IPSS-R“ auszuschließen und das Kriterium bezüglich potentieller Fälle von Leukozytose zu konkretisieren</li> <li>• Konkretisierung der Erhebung von ADA- und PK-Daten während des Follow-ups nach Behandlungsende für alle Patienten</li> <li>• Überarbeitung zur Konkretisierung der Zeitvorgabe der Woche 25-Prozedur und der Bewertung der MDS</li> <li>• Verlängerung der Follow-up-Periode nach Behandlungsende von „mindestens zwei Jahren“ auf „mindestens drei Jahre“ nach letzter Dosis Studienmedikation</li> </ul> <p>Darüber hinaus wurden weitere kleinere Anpassungen/ Konkretisierungen im Protokoll vorgenommen.</p>
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p><b>Einschlusskriterien:</b></p> <ul style="list-style-type: none"> <li>• Alter <math>\geq 18</math> Jahre bei Gabe der Einverständniserklärung</li> <li>• Verständnis und freiwillige Gabe der Einverständniserklärung vor jeglichen studienbezogenen Untersuchungen/Prozeduren</li> <li>• dokumentierte MDS-Diagnose gemäß WHO/FAB-Klassifizierung mit sehr niedrigem, niedrigem oder intermediärem Risiko nach IPSS-R-Klassifizierung und <ul style="list-style-type: none"> <li>○ einem Anteil von Ringsideroblasten an erythroiden Vorläuferzellen im Knochenmark von <math>\geq 15\%</math> oder <math>\geq 5\%</math> (aber <math>&lt; 15\%</math>) bei Vorliegen einer SF3B1-Mutation,</li> <li>○ <math>&lt; 5\%</math> Blasten im Knochenmark,</li> <li>○ Anzahl peripherer weißer Blutkörperchen <math>&lt; 13.000/\mu\text{l}</math></li> </ul> </li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• resistent/refraktär oder intolerant oder nicht geeignet für eine vorherige ESA-Behandlung, definiert anhand eines der folgenden Kriterien <ul style="list-style-type: none"> <li>○ <i>resistent/refraktär</i>: dokumentiertes Nicht-Ansprechen oder nicht anhaltendes Ansprechen auf vorherige ESA-enthaltende Behandlung, entweder als Monotherapie oder in Kombination mit bspw. G-CSF (Definition der ESA-enthaltenden Behandlung: entweder rekombinantes humanes Erythropoetin <math>\geq 40.000</math> IU/Woche für mindestens 8 Dosen oder Dosisäquivalenten, oder Darbepoetin-<math>\alpha</math> <math>\geq 500</math> <math>\mu</math>g q3w für mindestens 4 Dosen oder Dosisäquivalenten)</li> <li>○ <i>intolerant</i>: dokumentierter Behandlungsabbruch einer vorherigen ESA-enthaltenden Behandlung, entweder als Monotherapie oder in Kombination, bspw. mit G-CSF, zu jeder Zeit nach Behandlungsbeginn aufgrund einer Intoleranz oder UE</li> <li>○ <i>ungeeignet</i>: geringe Chance auf ein Ansprechen auf eine ESA-Therapie basierend auf Serum-EPO-Spiegeln <math>&gt;200</math> U/l für Patienten, die zuvor noch nicht mit ESA behandelt wurden</li> </ul> </li> <li>• mit ESA oder G-CSF/GM-CSF vorbehandelte Patienten: keine Behandlung mit diesen Substanzen für mindestens vier Wochen vor Randomisierung</li> <li>• EK-Transfusionsabhängigkeit, definiert als <ul style="list-style-type: none"> <li>○ durchschnittlicher Transfusionsbedarf von <math>\geq</math> zwei Einheiten EK/acht Wochen, bestätigt über einen Zeitraum von mindestens 16 Wochen unmittelbar vor Randomisierung</li> <li>○ Hämoglobin-Werte zum Zeitpunkt oder innerhalb von sieben Tagen vor Gabe einer EK-Transfusion von <math>\leq 10</math> g/dl, um die Transfusion hinsichtlich dieses Einschlusskriteriums zu zählen. EK-Transfusionen bei einem Hb-Wert <math>&gt;10</math> g/dl vor Transfusion oder für elektive Operationen wurden nicht als Transfusion im Sinne des Einschlusskriteriums berücksichtigt</li> <li>○ kein zusammenhängendes 56-Tage-Intervall innerhalb der 16 Wochen unmittelbar vor Randomisierung, in welchem keine EK-Transfusion verabreicht wurde</li> </ul> </li> <li>• ECOG-Status von 0, 1 oder 2</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• Frauen im gebärfähigen Alter waren definiert als geschlechtsreife Frauen, die (1) keine Hysterektomie oder bilaterale Oophorektomie hatten oder (2) noch nicht natürlicherweise postmenopausal (Amenorrhö nach Krebs war kein Ausschluss der Gebärfähigkeit) für mindestens 24 zusammenhängende Monate waren (d.h. eine Menstruation innerhalb der vorhergehenden 24 Monate). Frauen im gebärfähigen Alter müssen <ul style="list-style-type: none"> <li>○ zwei negative, durch den Prüfarzt bestätigte Schwangerschaftstests vor Beginn der Studienmedikation vorliegen haben (außer der Screening-Schwangerschaftstest wurde innerhalb von 72 Stunden des Tages eins des Zyklus eins durchgeführt) und Einverständnis geben, während der Studie und nach Behandlungsende weitere Schwangerschaftstest durchzuführen</li> <li>○ falls sexuell aktiv, zustimmen und in der Lage sein, hoch wirksame Methoden der Kontrazeption (definiert als hormonelle Kontrazeption, Intrauterinpeessar, Tubenligatur oder einen vasktomierter Partner) ohne Unterbrechung fünf Wochen vor Behandlungsbeginn, während der Behandlung mit Studienmedikation und für 12 Wochen nach Beendigung der Studienmedikation einzuhalten</li> </ul> </li> <li>• Männliche Probanden müssen einwilligen, während der Studienteilnahme, Dosisunterbrechungen und für mindestens 12 Wochen nach Abbruch der Studienmedikation bei Geschlechtsverkehr mit schwangeren Frauen oder Frauen im gebärfähigen Alter Kondome für Männer zu verwenden (definiert als Latexkondom oder Nicht-Latexkondom, jedoch nicht aus natürlichen (tierischen) Membranen) – selbst bei erfolgreicher Vasektomie</li> <li>• Einverständnis, die Studienvisiten und andere Anforderungen des Protokolls zu befolgen</li> </ul>

		<p><b>Ausschlusskriterien:</b></p> <ul style="list-style-type: none"> <li>• Vorherige Therapie mit krankheitsmodifizierenden Substanzen für die zugrundeliegende MDS-Erkrankung (z.B. Immunmodulatoren (wie Lenalidomid), hypomethylierende Substanzen oder Immunsuppressiva)       <ul style="list-style-type: none"> <li>○ Einschluss von Patienten nach Ermessen des Prüfarztes erlaubt, sofern nicht mehr als zwei Dosen hypomethylierender Substanz oder nicht mehr als eine Kalenderwoche Lenalidomid verabreicht wurde und die letzte Dosis <math>\geq</math> fünf Wochen vor Randomisierung verabreicht wurde</li> </ul> </li> <li>• Vorherige Behandlung mit Luspatercept oder Sotatercept (ACE-011)</li> <li>• MDS mit Del(5q)-Anomalie</li> <li>• Sekundäre MDS, d.h. MDS, die durch chemische Verletzung oder Chemotherapie und/oder Bestrahlung einer anderen Erkrankung entstanden ist</li> <li>• Bekannte klinisch signifikante Anämie aufgrund von Eisen-, Vitamin B12- oder Folsäuremangel, oder autoimmuner oder vererbter hämolytischen Anämie, oder gastrointestinaler Blutung       <ul style="list-style-type: none"> <li>○ Eisenmangel bestimmt als Serum Ferritin <math>\leq 15 \mu\text{g/l}</math> und weitere Untersuchungen, sofern klinisch indiziert (z.B. berechnete Transferrin-Sättigung [Eisen/Gesamteisenbindungskapazität <math>\leq 20\%</math>] oder Knochenmarksfärbung für Eisen)</li> </ul> </li> <li>• Vorherige allogene oder autologe Stammzelltransplantation</li> <li>• Bekannte Vorgeschichte einer AML-Diagnose</li> <li>• Verwendung der folgenden Substanzen innerhalb der fünf Wochen vor Randomisierung:       <ul style="list-style-type: none"> <li>○ zytotoxische Chemotherapie</li> <li>○ Kortikosteroide, außer Patienten mit stabiler oder sinkender Dosis über mindestens eine Woche vor Randomisierung zur Behandlung von anderen medizinischen Konditionen als MDS</li> <li>○ Eisenchelattherapie, außer Patienten mit stabiler oder sinkender Dosis über mindestens acht Wochen vor Randomisierung</li> <li>○ andere hämatopoetische Wachstumsfaktoren für Erythrozyten (z.B. Interleukin-3)</li> <li>○ Prüfsubstanzen oder -geräte oder zugelassene Behandlungen für experimentelle Zwecke. Sofern die Halbwertszeit der vorherigen Prüfsubstanz bekannt war, wurde deren Anwendung innerhalb von fünf</li> </ul> </li> </ul>
--	--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

		<p>Halbwertszeiten vor Randomisierung oder fünf Wochen, je nachdem was länger war, ausgeschlossen.</p> <ul style="list-style-type: none"> <li>• Unkontrollierte Hypertonie, definiert als wiederholte Erhöhungen des diastolischen Blutdrucks <math>\geq 100</math> mm Hg trotz adäquater Behandlung</li> <li>• Absolute Neutrophilenzahl (ANC) <math>&lt; 500/\mu\text{l}</math> (<math>0,5 \times 10^9/\text{l}</math>)</li> <li>• Thrombozytenzahl <math>&lt; 50.000/\mu\text{l}</math> (<math>50 \times 10^9/\text{l}</math>)</li> <li>• Geschätzte glomeruläre Filtrationsrate oder Kreatinin-Clearance <math>&lt; 40</math> ml/min</li> <li>• Aspartataminotransferase (AST)/Serum-Glutamat-Oxalacetat-Transaminase (SGOT) oder Alaninaminotransferase (ALT)/Serum-Glutamat-Pyruvat-Transaminase (SGPT) <math>\geq 3,0</math> x oberer Normalwert (<i>upper limit of normal</i>, ULN)</li> <li>• Gesamtbilirubin <math>\geq 2,0</math> x ULN <ul style="list-style-type: none"> <li>○ höhere Werte waren erlaubt, wenn diese einer aktiven Zerstörung der Erythrozytenvorläuferzellen im Knochenmark zuzuschreiben sein konnten (d.h. ineffektive Erythropoese) oder bei Vorliegen oder bekannter Vorgeschichte eines Gilbert-Syndroms</li> <li>○ Patienten wurden ausgeschlossen, wenn eine autoimmune hämolytische Anämie vorlag, manifestiert als korrigierte Retikulozytenzahl <math>&gt; 2\%</math> mit entweder einem positiven Coombs Test oder über 50% indirektem Bilirubin</li> </ul> </li> <li>• Vorgeschichte von malignen Erkrankungen, mit Ausnahme von MDS, außer der Patient war für <math>\geq</math> fünf Jahre krankheitsfrei (einschließlich Beendigung jeglicher aktiven oder adjuvanten Behandlung der vorherigen malignen Erkrankung). Patienten mit den folgenden Erkrankungen (in der Vorgeschichte/andauernd) waren erlaubt: <ul style="list-style-type: none"> <li>○ Basalzell- oder Plattenepithelzellkarzinom der Haut</li> <li>○ Carcinoma in situ der Zervix</li> <li>○ Carcinoma in situ der Brust</li> <li>○ zufälliger histologischer Befund eines Prostatakrebses (T1a oder T1b des TNM-Staging-Systems)</li> </ul> </li> <li>• Große Operation innerhalb der acht Wochen vor Randomisierung. Patienten müssen sich vollständig von jeglichen vorangegangenen Operationen vor Randomisierung erholt haben.</li> <li>• Vorgeschichte eines Schlaganfalls, einer tiefen Venenthrombose, einer pulmonalen oder</li> </ul>
--	--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>arteriellen Embolie innerhalb der 6 Monate vor Randomisierung</p> <ul style="list-style-type: none"> <li>• Schwangere oder stillende Frauen</li> <li>• Herzinfarkt, unkontrollierte Angina, unkontrollierte Herzinsuffizienz oder unkontrollierte Arrhythmie, festgestellt durch den Prüfarzt innerhalb von sechs Monaten vor Randomisierung. Patienten mit einer bekannten Ejektionsfraktion von &lt;35%, bestätigt durch ein lokales EKG oder MUGA innerhalb der sechs Monate vor Randomisierung, wurden ausgeschlossen.</li> <li>• Unkontrollierte systemische Pilz-, bakterielle oder virale Infektion (definiert durch anhaltende Zeichen/Symptome der Infektion ohne Besserung trotz adäquater antibiotischer, antiviraler und/oder anderer Behandlung), bekannte HIV-Infektion, bekannter Nachweis einer aktiven, infektiösen Hepatitis B und/oder bekannter Nachweis einer aktiven Hepatitis C. Lokale Untersuchung zur Bestätigung des HIV-, HBV- und HCV-Status sollte nicht länger als vier Wochen vom Datum der ICF-Unterschrift entfernt liegen.</li> <li>• Vorgeschichte einer schwerwiegenden allergischen oder anaphylaktischen Reaktion oder Hypersensitivität gegenüber rekombinanten Proteinen oder Hilfsstoffen der Prüfsubstanz</li> <li>• Patienten mit jeglicher bedeutenden medizinischen Kondition, Laborwertanomalien, psychiatrischer Erkrankung oder mit Einstufung gemäß lokalen Vorschriften als gefährdet (z.B. inhaftiert oder institutionalisiert), die die Teilnahme der Studie verhindern würde</li> <li>• Patienten mit jeglicher Kondition, einschließlich des Vorliegens von Laborwertanomalien, welche den Patienten einem nicht akzeptablen Risiko aussetzen würde, sollte er/sie in der Studie teilnehmen</li> <li>• Patienten mit jeglicher Kondition oder begleitender Medikation, die in der Lage war, die Interpretierbarkeit der Studiendaten zu zunichte zu machen</li> </ul>
4b	Studienorganisation und Ort der Studiendurchführung	<p><b>Studiendurchführung:</b> PPD, LLC (Wilmington, NC, US)</p> <p><b>Koordinierender Prüfarzt:</b> Prof. Uwe Platzbecker (Universitätsklinikum Carl Gustav Carus Dresden, Medizinische Klinik und Poliklinik I, Dresden, Deutschland)</p> <p><b>Orte der Studiendurchführung:</b> 65 Zentren in 11 Ländern (Belgien, Deutschland, Frankreich, Großbritannien, Italien, Kanada, Niederlande, Schweden, Spanien, Türkei, Vereinigte Staaten von Amerika)</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
5	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p><b>Luspatercept + BSC:</b></p> <p>Initialdosis: 1,0 mg/kg Körpergewicht s. c. q3w (Verabreichung an Tag 1 jedes 21-Tage-Behandlungszyklus)</p> <p>Maximaldosis: 1,75 mg/kg Körpergewicht</p> <p>Dosisreduktion: um ~25%/Verabreichung</p> <p><b>Placebo + BSC:</b></p> <p>Volumenäquivalent Placebo (0,9 % Natriumchlorid-Lösung) zu Luspatercept s. c. q3w (Verabreichung an Tag eins jedes 21-Tage-Behandlungszyklus)</p> <p><b>BSC</b> umfasste eine bedarfsgerechte, supportive Transfusionstherapie mit Erythrozytenkonzentraten in Kombination mit Eisenchelatoren</p>
6	Zielkriterien	
6a	<p>Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten</p>	<p><b>Primärer Endpunkt:</b></p> <ul style="list-style-type: none"> <li>• Erythrozyten-Transfusionsfreiheit für mindestens acht Wochen (im Studienzeitraum Woche 1-24)</li> </ul> <p><b>Sekundäre Endpunkte:</b></p> <ul style="list-style-type: none"> <li>• Erythrozyten-Transfusionsfreiheit für mindestens 12 Wochen (im Studienzeitraum Woche 1-24 und 1-48)</li> <li>• Erythrozyten-Transfusionsfreiheit für mindestens acht Wochen (im Studienzeitraum Woche 1-48)</li> <li>• Reduktion der transfundierten EK-Einheiten über 16 Wochen (im Studienzeitraum Woche 9-24 und 33-48)</li> <li>• Erreichen des mHI-E per IWG-Kriterien 2006 (im Studienzeitraum Woche 1-24 und 1-48)</li> <li>• Durchschnittliche Hämoglobinerhöhung um mindestens 1,0 g/dl (im Studienzeitraum Woche 1-24 und 1-48)</li> <li>• Dauer der Erythrozyten-Transfusionsfreiheit (im Studienzeitraum Woche 1-48 und 1-EOT (<i>end of treatment</i>))</li> <li>• Gesundheitsbezogene Lebensqualität gemessen anhand des EORTC QLQ-C30-Fragebogens (im Studienzeitraum Woche 1-48 und Baseline-EOT)</li> <li>• Erreichen der HI-N per IWG-Kriterien 2006 (im Studienzeitraum Woche 1-24 und 1-48)</li> <li>• Erreichen der HI-P per IWG-Kriterien 2006 (im Studienzeitraum Woche 1-24 und 1-48)</li> <li>• Durchschnittliche Verringerung des Serum-Ferritin-Spiegels (im Studienzeitraum Woche 9-24 und 33-48)</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• Durchschnittliche Verringerung des Verbrauchs an Eisenchelatoren (im Studienzeitraum Woche 9-24 und 33-48)</li> <li>• Zeit bis zum Erreichen einer Erythrozyten-Transfusionsfreiheit von mindestens acht Wochen (im Studienzeitraum Woche 1-24 und 1-48)</li> <li>• Progression in eine AML (vom Zeitpunkt der Randomisierung bis mindestens drei Jahre nach der letzten Dosis und im Studienzeitraum Woche 1-48)</li> <li>• Gesamtüberleben (vom Zeitpunkt der Randomisierung bis mindestens drei Jahre nach der letzten Dosis und im Studienzeitraum Woche 1-48)</li> <li>• Sicherheit (vom Zeitpunkt des Screenings bis 42 Tage nach der letzten Dosis und im Studienzeitraum Woche 1-48)</li> <li>• Populations-PK-Modell (vom Zeitpunkt der Randomisierung bis ein Jahr nach der ersten Dosis)</li> <li>• Verhältnis der Exposition zum Ansprechen (vom Zeitpunkt der Randomisierung bis ein Jahr nach der ersten Dosis)</li> <li>• Anti-Drug-Antibodies (vom Zeitpunkt der Randomisierung bis ein Jahr nach der ersten Dosis)</li> </ul> <p><b>Explorative Endpunkte:</b></p> <ul style="list-style-type: none"> <li>• Biomarker der TGF-<math>\beta</math>-Superfamilie (z.B. GDF11) und andere explorative Marker (z.B. CRP) (im Studienzeitraum Baseline-EOT)</li> <li>• MDS-assoziierte molekulare Mutationen (z.B. SF3B1) (im Studienzeitraum Baseline-EOT)</li> <li>• Health Ressource Utilization (im Studienzeitraum Baseline-EOT)</li> <li>• Gesundheitsbezogene Lebensqualität gemessen anhand des QoL-E (im Studienzeitraum Baseline-EOT)</li> </ul>
<b>6b</b>	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Nicht zutreffend.
<b>7</b>	Fallzahl	
<b>7a</b>	Wie wurden die Fallzahlen bestimmt?	Bei einer Fallzahl von 210 Patienten (140 Luspatercept (ACE-536), 70 Placebo) beträgt die Power, um einen Unterschied hinsichtlich der Ansprechrate (Erythrozyten-Transfusionsfreiheit über mindestens acht Wochen im Studienzeitraum von Woche 1-24) von 0,30 im Interventionsarm (Luspatercept + BSC) und 0,10 im Kontrollarm (Placebo + BSC) aufzudecken 90% (Einseitiges Signifikanzniveau: Alpha = 0,025; Teststatistiken zur Differenz der Anteile unter Verwendung gepoolter Varianzschätzung und 10% Dropout-Rate).

Item <sup>a</sup>	Charakteristikum	Studieninformation
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Eine Interimsanalyse (<i>Futility Analysis</i>) war für den primären Endpunkt und sekundären Hauptendpunkt geplant, sobald 105 Patienten (50% der geplanten Ereignisse des primären Endpunktes) die primäre Studienphase (erste 24 Wochen der doppelblinden Behandlung) abgeschlossen haben oder die Behandlung vor Erreichen der 24 Wochen der doppelblinden Behandlung abgebrochen haben.</p> <p>Ein Anspruch auf Überlegenheit basierend auf den Wirksamkeitsergebnissen der Interimsanalyse war nicht geplant, sodass der Typ-I-Fehler für die finale Analyse auf einem einseitigen Niveau von 0,025 gehalten wurde.</p>
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Zentrale Randomisierung mittels Interactive Response Technology (IRT) System im Verhältnis 2:1 (Luspatercept:Placebo) durch Algorithmus unter Berücksichtigung präspezifizierter Kriterien, darunter doppelblinde Behandlung und Stratifizierung der Randomisierung.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	<p>Stratifizierung anhand der folgenden Faktoren:</p> <ul style="list-style-type: none"> <li>• Erythrozyten-Transfusionslast zur Baseline <ul style="list-style-type: none"> <li>○ <math>\geq 6</math> EK-Einheiten pro acht Wochen (Durchschnitt der zwei Acht-Wochen-Intervalle unmittelbar vor Randomisierung)</li> <li>○ <math>&lt; 6</math> EK-Einheiten pro acht Wochen (Durchschnitt der zwei Acht-Wochen-Intervalle unmittelbar vor Randomisierung)</li> </ul> </li> <li>• IPSS-R zur Baseline <ul style="list-style-type: none"> <li>○ sehr niedrig, niedrig</li> <li>○ intermediär</li> </ul> </li> </ul>
9	<p>Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment)</p> <p>Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war</p>	<p>Ausgewähltem Personal an jedem Studienzentrum wurden Passwort-geschützte, kodierte Identifizierungsnummern zugeordnet, die zur Einwahl in das IRT-System autorisierten, um Patienten einzuschließen. Die Randomisierung erfolgte anhand eines Algorithmus durch das IRT-System. Die Randomisierungsnummer entsprach einem bestimmten Behandlungsarm innerhalb eines Stratum. Die Randomisierungsnummer alleine war nicht ausreichend, um eine Entblindung zu ermöglichen, da dafür eine vollständige Entblindung der kodierten Informationen gekoppelt mit der Randomisierungsnummer nötig war.</p> <p>Für die Zuordnung der Behandlung zu jedem Dosisstart und bei jeder Dosisreduktion oder -titration musste das Personal am Zentrum das IRT-System kontaktieren, um die neue Dosierung aufzunehmen und die neue Behandlung zugewiesen zu bekommen.</p>

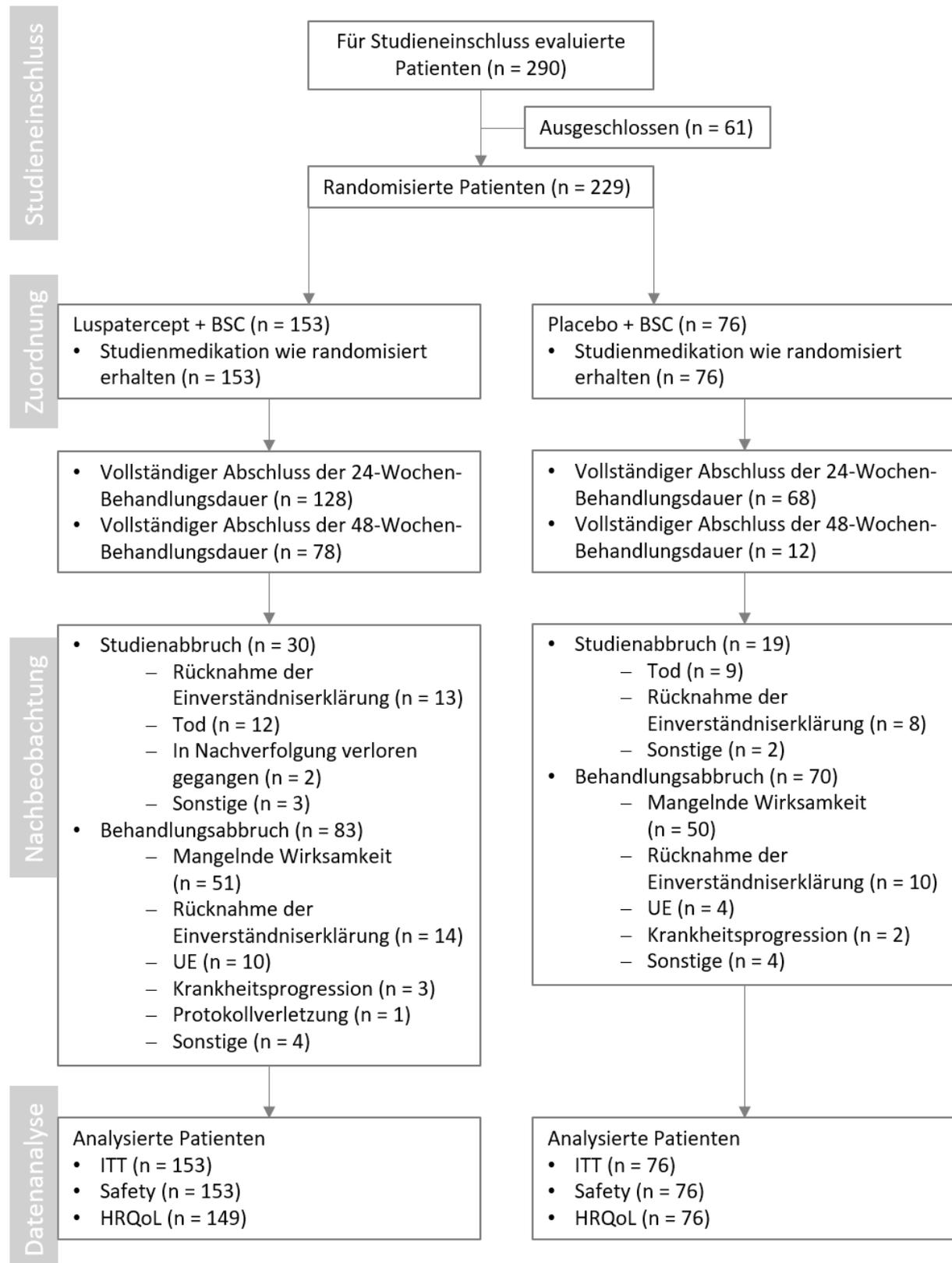
Item <sup>a</sup>	Charakteristikum	Studieninformation
<b>10</b>	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Randomisierung erfolgte durch das IRT-System, welches Patienten fortlaufend entsprechend der präspezifizierten Kriterien in ein geeignetes Stratum einsortierte.
<b>11</b>	Verblindung	
<b>11a</b>	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	a) ja b) ja c) ja  Die Verblindung wurde für alle Beteiligten (mit Ausnahme bestimmter Personen, wie z.B. dem Apotheker und dem bioanalytischen Labor) während der gesamten Studiendauer aufrecht erhalten bis alle Patienten die Studie beendet hatten oder dem Zeitpunkt, an dem die Studie entblindet wurde (gemäß DMC Empfehlung) und die Datenbank gesperrt wurde.
<b>11b</b>	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Für den Kontrollarm wurde 0,9%ige Natriumchlorid-Lösung zur Injektion in einer exakt gleich aussehenden Spritze wie die, die für rekonstituiertes Luspatercept verwendet wurde, vom ausgewählten Studienpersonal (z.B. dem Apotheker) aufgezo- gen, um ein passendes Placebo herzustellen. Die Abgabe von Luspatercept und Placebo an den Prüfarzt und dessen Angestellte erfolgte verblindet.
<b>12</b>	Statistische Methoden	
<b>12a</b>	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<b>Population für die Wirksamkeitsendpunkte:</b> <u>Intention-to-treat(ITT)-Population:</u> alle randomisierten Patienten, unabhängig davon, ob die Ein-/Ausschlusskriterien erfüllt waren oder die Studienmedikation erhalten wurde <b>Population für die Sicherheitsendpunkte:</b> <u>Safety Population:</u> alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben <b>Population für die Endpunkte der Lebensqualität:</b> <u>Auswertbare Population gesundheitsbezogener Lebensqualität (HRQoL-evaluable population):</u> alle Patienten der ITT-Population, die den EORTC QLQ-C30-Fragebogen zur Baseline (d.h. Tag eins des ersten Zyklus oder Screening Visite) und mindestens ein postbaseline Untersuchung vollständig (definiert als mindestens 50% beantworteter Fragen) ausgefüllt haben

		<p><b>Population für die Pharmakokinetik-Endpunkte:</b>  <u>Pharmakokinetik-Population:</u> alle Patienten, die mindestens eine Dosis Luspatercept erhalten haben und messbare Luspatercept-Serumkonzentrationen aufwiesen</p> <p><b>Multiplizität:</b>  Ein sequentieller Gatekeeping-Ansatz wurde für die Hypothesentestung multipler Endpunkte (prim. Endpunkt (Erythrozyten-Transfusionsfreiheit über mindestens acht Wochen) und sek. Hauptendpunkt (Erythrozyten-Transfusionsfreiheit über mindestens 12 Wochen) Wirksamkeitsendpunkt) verwendet, um den gesamten Fehler 1. Art zu kontrollieren. Der prim. Endpunkt und sek. Hauptendpunkt wird sequentiell, in einer präspezifizierten Reihenfolge (zuerst Testung des prim. Endpunkts zum einseitigen Signifikanzniveau von 0,025) durchgeführt. Um das einseitige Signifikanzniveau von 0,025 für alle Endpunkte der Erythrozyten-Transfusionsfreiheit zu erhalten, kann ein Rückschluss bezogen auf die sek. Hauptendpunkte nur erfolgen, sofern die Überlegenheit von Luspatercept für den prim. Endpunkt gezeigt werden konnte.  Von dem präspezifizierten sequentiellen Testen der Erythrozyten-Transfusionsfreiheit abgesehen, wurden keine weiteren Alpha-Anpassungen für Multiplizität gemacht.</p> <p><b>Analyse des primären Endpunkts:</b>  Ansprechrate einer Erythrozyten-Transfusionsfreiheit (keine EK-Transfusionen) in einem Zeitraum von mindestens 56 Tagen im Studienzeitraum von Woche 1-24  Cochran-Mantel-Haenszel (CMH)-Test mit einseitigem Signifikanzniveau von 0,025, stratifiziert nach durchschnittlicher Transfusionslast und IPSS-R zur Baseline</p> <p><b>Analyse des sekundären Hauptendpunkts:</b>  Ansprechrate einer Erythrozyten-Transfusionsfreiheit (keine EK-Transfusionen) in einem Zeitraum von mindestens 12 Wochen im Studienzeitraum von Woche 1-24 und Woche 1-48  CMH-Test mit einseitigem Signifikanzniveau von 0,025, stratifiziert nach durchschnittlicher Transfusionslast und IPSS-R zur Baseline (sequentielle Testung nach prim. Endpunkt)</p> <p><b>Analyse weiterer sekundärer Endpunkte:</b>  Deskriptive Analyse weiterer sekundärer Endpunkte. Für Ereigniszeitanalysen wurde die Kaplan-Meier-Methode verwendet. Kategorische Endpunkte wurden als Anzahl und Prozent dargestellt.</p>
--	--	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Item <sup>a</sup>	Charakteristikum	Studieninformation
		Für die Analyse der gesundheitsbezogenen Lebensqualität wurde der Behandlungseffekt mittels eines ANCOVA-Modells, adjustiert nach dem Baseline-Score der Domänen Stratifizierungsfaktoren der Randomisierung, von Woche eins zu Woche 25, 48 und EOT untersucht.
<b>12b</b>	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p>Durchführung von Subgruppenanalysen für die klinische Aktivität anhand der Stratifizierungsfaktoren und anderen Baseline-Charakteristika:</p> <ul style="list-style-type: none"> <li>• Altersklasse</li> <li>• Geschlecht</li> <li>• ethnische Zugehörigkeit</li> <li>• MDS WHO Klassifikation zur Baseline</li> <li>• Zeit seit initialer MDS-Diagnose zur Baseline</li> <li>• ECOG-Performance-Status zur Baseline</li> <li>• Durchschnittlicher Erythrozyten-Transfusionsbedarf zur Baseline</li> <li>• IPSS-R zur Baseline</li> <li>• Thrombozytenzahl zur Baseline</li> <li>• Nierenfunktionsstatus (Kreatinin-Clearance)</li> <li>• Serum-Erythropoetin-Spiegel (U/l) zur Baseline</li> <li>• Längste Dauer einer vorherigen ESA-Therapie</li> <li>• SF3B1-Mutationsstatus</li> </ul>
<b>Resultate</b>		
<b>13</b>	Patientenfluss (inklusive Flow-Charts zur Veranschaulichung im Anschluss an die Tabelle)	
<b>13a</b>	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p><b>Luspatercept + BSC vs. Placebo + BSC:</b></p> <p>a) 153 vs. 76 (total: 229) b) 153 vs. 76 (total: 229) c) 153 vs. 76 (total: 229)</p>
<b>13b</b>	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p><b>Luspatercept + BSC:</b></p> <ul style="list-style-type: none"> <li>• Widerruf des Patienten (8,5 %)</li> <li>• Tod (7,8 %)</li> <li>• <i>Lost to follow-up</i> (1,3 %)</li> <li>• Sonstige (2,0 %)</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<b>Placebo + BSC:</b> <ul style="list-style-type: none"> <li>• Widerruf des Patienten (10,5 %)</li> <li>• Tod (11,8 %)</li> <li>• Sonstige (2,6 %)</li> </ul>
<b>14</b>	Aufnahme / Rekrutierung	
<b>14a</b>	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	First subject first visit: 09.02.2016 Data cut-off date: 08.05.2018
<b>14b</b>	Informationen, warum die Studie endete oder beendet wurde	Nicht zutreffend (MEDALIST wurde noch nicht beendet).
<b>a: nach CONSORT 2010.</b>		

*Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.*



BSC = Best-Supportive-Care; HRQoL = gesundheitsbezogene Lebensqualität (*health-related quality of life*); ITT = Intention-to-Treat; UE = unerwünschtes Ereignis

Abbildung 4-64: Patientenfluss in MEDALIST zum Datenschnitt vom 08.05.2018

**Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten**

Der nachfolgend dargestellte Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- Verzerrungsaspekte auf Endpunktebene. In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich sollen die Bögen studienbezogen ausgefüllt werden. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

*Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.*

Tabelle 4-117 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MEDALIST

**Studie: MEDALIST**

**Tabelle: Liste der für die Bewertung herangezogenen Quellen**

Genauere Benennung der Quelle	Kürzel
Studienbericht	(A)
Studienprotokoll	(B)
Studienregistereintrag (Clinicaltrials.gov)	(C)
Studienregistereintrag (EU-CTR)	(D)
Studienregistereintrag (ICTRP)	(E)

**A Verzerrungsaspekte auf Studienebene:**

**Einstufung als randomisierte Studie**

**ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

**nein** → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

---

Es handelt sich um eine randomisierte Studie.  
A, B, C, D, E

---

1.

**für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz**

**ja**     **unklar**     **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Die Gruppenzuteilung erfolgte zentral und rein zufällig mittels eines IRT(Interactive Response Technology)-Systems in die beiden Behandlungsarmen Luspatercept + BSC und Placebo + BSC im Verhältnis 2:1 stratifiziert nach Transfusionslast und IPSS-R zur Baseline. Die Erzeugung der Randomisierungssequenz erfolgte durch einen geeigneten Algorithmus, auf den das IRT-System zugriff.  
A, B

---

**für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen**

**ja**     **unklar**     **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---



---

2.

**für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Die Zuteilung erfolgte durch das IRT-System, sodass Studienteilnehmer und -personal verblindet blieben. Ausgewähltes Studienpersonal konnte sich mithilfe einer Passwort-geschützten Identifizierungsnummer in das IRT-System einwählen, um Patienten in die Studie aufzunehmen. Die Zuteilung erfolgte anhand eines Randomisierungsalgorithmus durch das IRT-System. Die Randomisierungsnummer entsprach einer Behandlung innerhalb eines Stratum, war jedoch nicht ausreichend, um eine Entblindung vorzunehmen.

A

**für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

## 3. Verblindung von Patienten und behandelnden Personen

**Patient:** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

In dieser Studie waren alle Patienten, Studienpersonal und Vertreter der Firma Celgene hinsichtlich der Behandlung verblindet. Davon ausgenommen waren ausgewählte Einzelpersonen (z.B. der Apotheker des Studienzentrums), die eine adäquate Herstellung der Placebo-Gaben sicherstellten. Die Randomisierung, Arzneimittelausgabe, Dosisreduktion/-titrierung und Beendigung der Studienmedikation waren durch das IRT-System gestützt, sodass die Verblindung sichergestellt wurde. Die Behandlungszuweisung durfte nur entblindet werden, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren. Eine Notfall-Entblindung musste anhand eines strikten Protokolls über das IRT-System erfolgen.

A, B

**behandelnde bzw. weiterbehandelnde Personen:** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

In dieser Studie waren alle Patienten, Studienpersonal und Vertreter der Firma Celgene hinsichtlich der Behandlung verblindet. Davon ausgenommen waren ausgewählte Einzelpersonen (z.B. der Apotheker des Studienzentrums), die eine adäquate Herstellung der Placebo-Gaben sicherstellten. Die Randomisierung, Arzneimittelausgabe, Dosisreduktion/-titrierung und Beendigung der Studienmedikation waren durch das IRT-System gestützt, sodass die Verblindung sichergestellt wurde. Die Behandlungszuweisung durfte nur entblindet werden, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren. Eine Notfall-Entblindung musste anhand eines strikten Protokolls über das IRT-System erfolgen.

---

A, B**4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

---

A, B, C, D, E**5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.

---

A, B, C, D, E**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Die Randomisierung erfolgte zentral unter Verwendung eines IRT-Systems. Während der gesamten Studie waren Studienteilnehmer, -personal und -sponsor hinsichtlich der Behandlungszuteilung verblindet. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Verzerrung konnten nicht identifiziert werden.

---

Das Verzerrungspotenzial auf Studienebene wird somit als niedrig eingestuft.

**B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:****Endpunkt: Mortalität – Gesamtüberleben****1. Verblindung der Endpunkterheber**

ja       unklar       nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.

A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips**

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen alle randomisierten Patienten entsprechend der ITT-Population ein.

A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

ja       nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Für den Endpunkt Mortalität – Gesamtüberleben wurde der gesamte Studienverlauf berücksichtigt. Relevante Unterschiede in den Beobachtungszeiten zwischen den beiden Behandlungsarmen liegen nicht vor, da eine Nachverfolgung des Gesamtüberlebens unabhängig von der Behandlungsdauer bis mindestens drei Jahre nach Therapieabbruch vorgesehen war. Das Ereignis Tod jeglicher Ursache stellt zudem ein objektives Ereignis dar, welches nicht weiter durch subjektive Faktoren beeinträchtigt wird.

A, B

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

**niedrig**                       **hoch**

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft. Der Endpunkt Mortalität – Gesamtüberleben wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die die Verzerrung beeinflussen könnten, liegen nicht vor. In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Mortalität – Gesamtüberleben als niedrig eingestuft.

---

**Endpunkt: Morbidität – Transfusionsfreiheit****1. Verblindung der Endpunkterheber**

**ja**                       **unklar**                       **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewährleisten.  
A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips**

**ja**                       **unklar**                       **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen alle randomisierten Patienten entsprechend der ITT-Population ein.  
A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

**ja**                       **unklar**                       **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.  
A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.  
A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.  
Der Endpunkt Morbidität – Transfusionsfreiheit wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.  
In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Morbidität – Transfusionsfreiheit als niedrig eingestuft.

---

**Endpunkt: Morbidität – Reduktion der Transfusionslast****1. Verblindung der Endpunkterheber** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.  
A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen alle randomisierten Patienten entsprechend der ITT-Population ein.  
A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.  
A, B, C, D, E

---

#### 4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja       nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.  
A, B, C, D, E

---

#### Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.  
Der Endpunkt Morbidität – Reduktion der Transfusionslast wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.  
In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Morbidität – Reduktion der Transfusionslast als niedrig eingestuft.

---

#### Endpunkt: Morbidität – Reduktion der Eisenüberladung

##### 1. Verblindung der Endpunkterheber

ja       unklar       nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.  
A, B

---

##### 2. Adäquate Umsetzung des ITT-Prinzips

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen alle randomisierten Patienten entsprechend der ITT-Population ein.  
A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.  
A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.  
A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.  
Der Endpunkt Morbidität – Reduktion der Eisenüberladung wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.  
In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Morbidität – Reduktion der Eisenüberladung als niedrig eingestuft.

---

**Endpunkt: Morbidität – Reduktion der Eisenüberladung –Bedarf an Eisenchelatoren****1. Verblindung der Endpunkterheber** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewährleisten.  
A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen alle randomisierten Patienten entsprechend der ITT-Population ein.

A, B

---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

A, B, C, D, E

---

### 4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja       nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.

A, B, C, D, E

---

### Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Morbidität – Reduktion der Eisenüberladung – Bedarf an Eisenchelatoren wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.

In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Morbidität – Reduktion der Eisenüberladung – Bedarf an Eisenchelatoren als niedrig eingestuft.

---

### Endpunkt: Morbidität – Symptomatik

#### 1. Verblindung der Endpunkterheber

ja       unklar       nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewährleisten.

A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen nur solche Patienten ein, die zu Studienbeginn und mindestens einem Zeitpunkt im Studienverlauf einen EORTC QLQ-C30-Fragebogen ausgefüllt hatten (HRQoL-Population). Zwischen der HRQoL-Population und der ITT-Population liegen keine wesentlichen Unterschiede hinsichtlich der Baseline-Charakteristika vor. Darüber hinaus umfasst die HRQoL-Population > 95 % der Patienten der ITT-Population (Luspatercept + BSC: 149/153 (97,4 %); Placebo + BSC: 76/76 (100,0 %)), sodass von einer guten Vergleichbarkeit der HRQoL- und ITT-Population auszugehen ist. Die Rücklaufquoten bezogen auf die ITT-Population liegen bis Woche 48 über den vom G BA geforderten 70 % und sind damit ausreichend hoch, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen.

A, B

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

A, B, C, D, E

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.

A, B, C, D, E

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Morbidität – Symptomatik wurde verblindet erhoben. Unter Berücksichtigung der herangezogenen Studienpopulation (HRQoL-Population) und der Rücklaufquoten des Fragebogens liegen bis Woche 48 ausreichende Daten vor, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.

In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Morbidität – Symptomatik als niedrig eingestuft.

---

**Endpunkt: Morbidität – Hospitalisierungen****1. Verblindung der Endpunkterheber** ja     unklar     neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.

A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja     unklar     neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen alle randomisierten Patienten entsprechend der ITT-Population ein.

A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja     unklar     neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja     neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.

A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig     hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Morbidität – Hospitalisierungen wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.

In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Morbidität – Hospitalisierungen als niedrig eingestuft.

---

## Endpunkt: Lebensqualität – EORTC QLQ-C30

### 5. Verblindung der Endpunkterheber

ja     unklar     nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.

A, B

---

### 6. Adäquate Umsetzung des ITT-Prinzips

ja     unklar     nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen nur solche Patienten ein, die zu Studienbeginn und mindestens einem Zeitpunkt im Studienverlauf einen EORTC QLQ-C30-Fragebogen ausgefüllt hatten (HRQoL-Population). Zwischen der HRQoL-Population und der ITT-Population liegen keine wesentlichen Unterschiede hinsichtlich der Baseline-Charakteristika vor. Darüber hinaus umfasst die HRQoL-Population > 95 % der Patienten der ITT-Population (Luspatercept + BSC: 149/153 (97,4 %); Placebo + BSC: 76/76 (100,0 %)), sodass von einer guten Vergleichbarkeit der HRQoL- und ITT-Population auszugehen ist. Die Rücklaufquoten bezogen auf die ITT-Population liegen bis Woche 48 über den vom G BA geforderten 70 % und sind damit ausreichend hoch, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen.

A, B

---

### 7. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja     unklar     nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

A, B, C, D, E

---

**8. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.  
A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.  
Der Endpunkt Lebensqualität – EORTC QLQ-C30 wurde verblindet erhoben. Unter Berücksichtigung der herangezogenen Studienpopulation (HRQoL-Population) und der Rücklaufquoten des Fragebogens liegen bis Woche 48 ausreichende Daten vor, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.  
In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Lebensqualität – EORTC QLQ-C30 als niedrig eingestuft.

---

**Endpunkt: Lebensqualität – QoL-E****1. Verblindung der Endpunkterheber** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.  
A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips**

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse des Endpunktes gingen nur solche Patienten ein, die zu Studienbeginn und mindestens einem Zeitpunkt im Studienverlauf einen EORTC QLQ-C30-Fragebogen ausgefüllt hatten (HRQoL-Population). Zwischen der HRQoL-Population und der ITT-Population liegen keine wesentlichen Unterschiede hinsichtlich der Baseline-Charakteristika vor. Darüber hinaus umfasst die HRQoL-Population > 95 % der Patienten der ITT-Population (Luspatercept + BSC: 149/153 (97,4 %); Placebo + BSC: 76/76 (100,0 %)), sodass von einer guten Vergleichbarkeit der HRQoL- und ITT-Population auszugehen ist. Die Rücklaufquoten bezogen auf die ITT-Population liegen bis Woche 48 über den vom G BA geforderten 70 % und sind damit ausreichend hoch, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen.

A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

ja       nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.

A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Lebensqualität – QoL-E wurde verblindet erhoben. Unter Berücksichtigung der herangezogenen Studienpopulation (HRQoL-Population) und der Rücklaufquoten des Fragebogens liegen bis Woche 48 ausreichende Daten vor, um eine valide Bewertung der Ergebnisse des Fragebogens zu ermöglichen. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.

In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Lebensqualität – QoL-E als niedrig eingestuft.

---

**Endpunkt: Sicherheit – jegliche UE****1. Verblindung der Endpunkterheber**

ja       unklar       nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.

A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips**

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse gingen nur solche Patienten ein, die mindestens eine Dosis Studienmedikation erhalten haben (Safety-Population). Patienten werden in dem Behandlungsarm berücksichtigt, dessen Behandlung sie tatsächlich erhalten haben. Die Safety-Population entspricht der ITT-Population (Luspatercept + BSC: 153/153 (100,0 %); Placebo + BSC: 76/76 (100,0 %)), sodass das ITT-Prinzip adäquat umgesetzt wurde.

A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

ja       nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.

A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** **niedrig**                       **hoch**

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.  
Der Endpunkt Sicherheit – jegliche UE wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.  
In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Sicherheit – jegliche UE als niedrig eingestuft.

---

**Endpunkt: Sicherheit – SUE****1. Verblindung der Endpunkterheber** **ja**                       **unklar**                       **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.  
A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips** **ja**                       **unklar**                       **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse gingen nur solche Patienten ein, die mindestens eine Dosis Studienmedikation erhalten haben (Safety-Population). Patienten werden in dem Behandlungsarm berücksichtigt, dessen Behandlung sie tatsächlich erhalten haben. Die Safety-Population entspricht der ITT-Population (Luspatercept + BSC: 153/153 (100,0 %); Placebo + BSC: 76/76 (100,0 %)), sodass das ITT-Prinzip adäquat umgesetzt wurde.  
A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** **ja**                       **unklar**                       **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.  
A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.  
A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.  
Der Endpunkt Sicherheit – SUE wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt.  
Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.  
In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Sicherheit – SUE als niedrig eingestuft.

---

**Endpunkt: Sicherheit – schwere UE****1. Verblindung der Endpunkterheber** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.  
A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse gingen nur solche Patienten ein, die mindestens eine Dosis Studienmedikation erhalten haben (Safety-Population). Patienten werden in dem Behandlungsarm berücksichtigt, dessen Behandlung sie tatsächlich erhalten haben. Die Safety-Population entspricht der ITT-Population (Luspatercept + BSC: 153/153 (100,0 %); Placebo + BSC: 76/76 (100,0 %)), sodass das ITT-Prinzip adäquat umgesetzt wurde.  
A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.  
A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.  
A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.  
Der Endpunkt Sicherheit – schwere UE wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.  
In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Sicherheit – schwere UE als niedrig eingestuft.

---

**Endpunkt: Sicherheit – Therapieabbrüche aufgrund von UE****1. Verblindung der Endpunkterheber** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

In der Studie MEDALIST waren Patienten, Studienpersonal und Mitarbeiter von Celgene verblindet. Davon ausgenommen waren lediglich ausgewählte Einzelpersonen, die die korrekte Umsetzung der Placebogabe sicherstellten (z.B. Apotheker des Studienzentrums). Eine Entblindung war nur gestattet, wenn dies aus Sicht des Prüfarztes absolut notwendig war, um eine sichere Behandlung des Patienten zu gewähren.  
A, B

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

In die Analyse gingen nur solche Patienten ein, die mindestens eine Dosis Studienmedikation erhalten haben (Safety-Population). Patienten werden in dem Behandlungsarm berücksichtigt, dessen Behandlung sie tatsächlich erhalten haben. Die Safety-Population entspricht der ITT-Population (Luspatercept + BSC: 153/153 (100,0 %); Placebo + BSC: 76/76 (100,0 %)), sodass das ITT-Prinzip adäquat umgesetzt wurde.

---

A, B

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf eine ergebnisgesteuerte Berichterstattung.

---

A, B, C, D, E

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es gab keinerlei Hinweise auf sonstige Aspekte, die zu einer Verzerrung führen könnten.

---

A, B, C, D, E

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft.

Der Endpunkt Sicherheit – Therapieabbrüche aufgrund von UE wurde verblindet erhoben, das ITT-Prinzip wurde adäquat umgesetzt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die zu einer Verzerrung führen können, liegen nicht vor.

In der Gesamtschau wird das Verzerrungspotential für den Endpunkt Sicherheit – Therapieabbrüche aufgrund von UE als niedrig eingestuft.

---

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

### Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunkübergreifend; B: endpunktspezifisch).

#### A Verzerrungsaspekte auf Studienebene:

##### Einstufung als randomisierte Studie

**ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

**nein:** Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

---

---

1.

##### für randomisierte Studien:

##### Adäquate Erzeugung der Randomisierungssequenz

**ja:** Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

**unklar:** Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

**nein:** Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

##### für nicht randomisierte Studien:

##### Zeitliche Parallelität der Gruppen

**ja:** Die Gruppen wurden zeitlich parallel verfolgt.

**unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

**nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

2.

**für randomisierte Studien:****Verdeckung der Gruppenzuteilung („allocation concealment“)** **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

 **unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau. **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**für nicht randomisierte Studien:****Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren** **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

 **unklar:** Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau. **nein:** Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**3. Verblindung von Patienten und behandelnden Personen****Patient:** **ja:** Die Patienten waren verblindet. **unklar:** Es finden sich keine diesbezüglichen Angaben. **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---



---

**behandelnde bzw. weiterbehandelnde Personen:**

**ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

**unklar:** Es finden sich keine diesbezüglichen Angaben.

**nein:** Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---



---

**4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte**

*Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).*

*Beispiele zu a und b:*

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

*Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter*

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

*Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:*

- *Ableich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Ableich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

*Zulässige Gründe sind:*

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

##### 5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
  - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
  - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein.
  - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
  - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

---

##### Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

---



---

### B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

**Endpunkt:** \_\_\_\_\_

#### 1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---



---

#### 2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

**unklar:** Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

**nein:** Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---



---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

*Beachte die Hinweise zu Punkt A.4!*

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---



---

### 4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

**ja**

**nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---



---

### Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

*Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.*

**niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

**hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

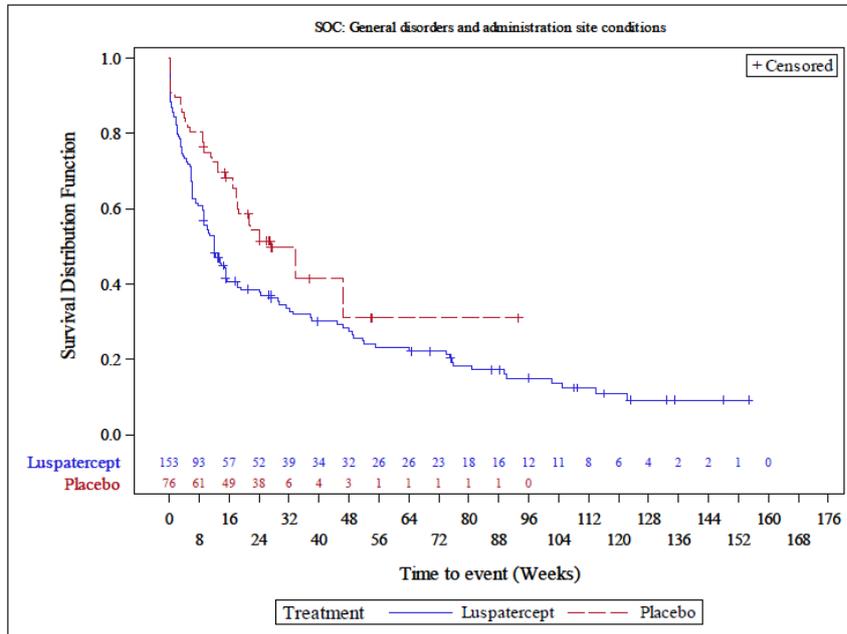
Begründung für die Einstufung:

---

---

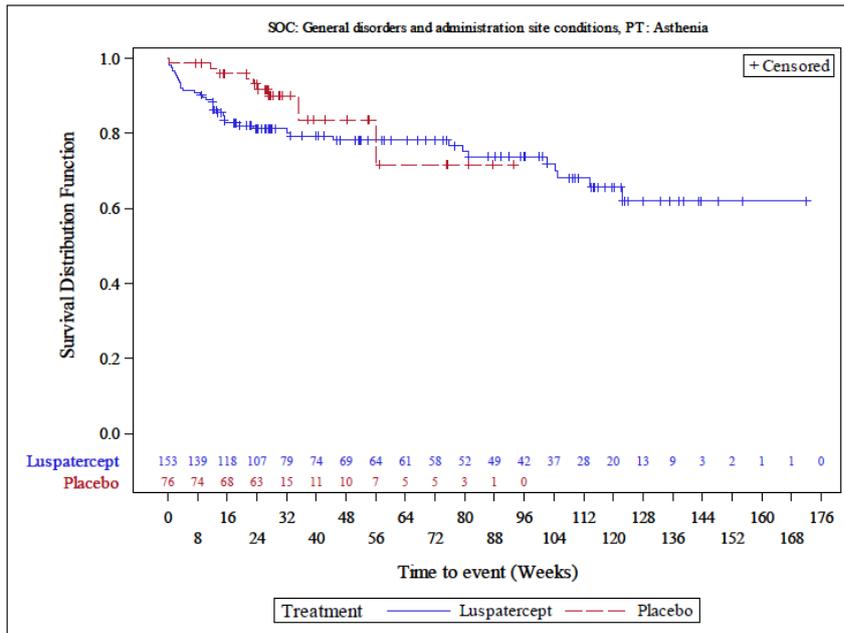
**Anhang 4-G: Kaplan-Meier-Kurven der Zeit bis zum ersten UE auf Ebene der SOC und PT**

**Jegliche UE**



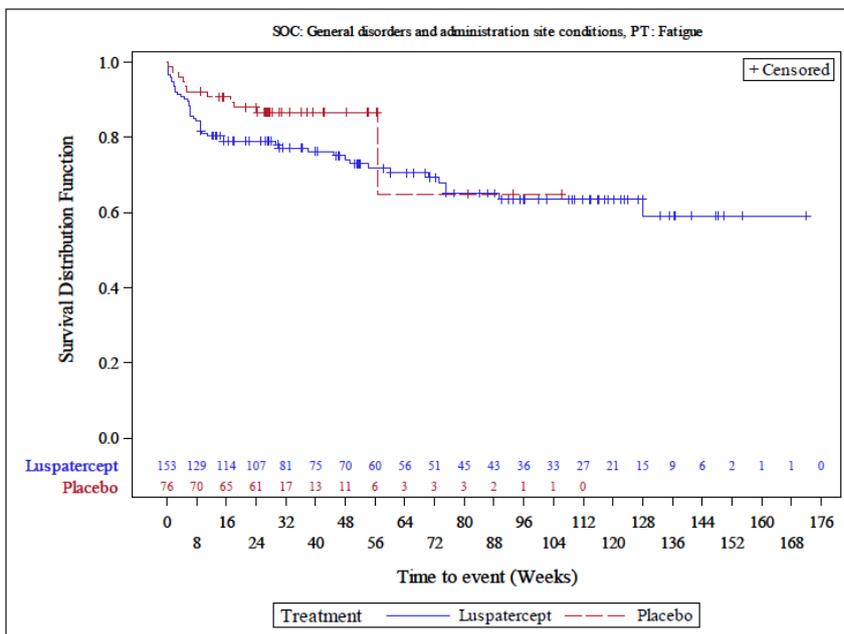
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-65 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



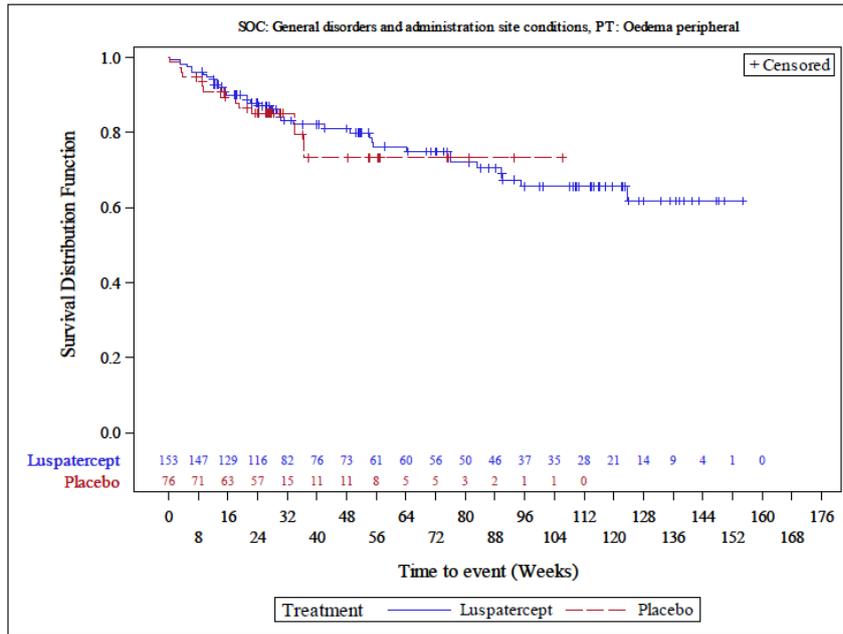
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-66 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Asthenie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



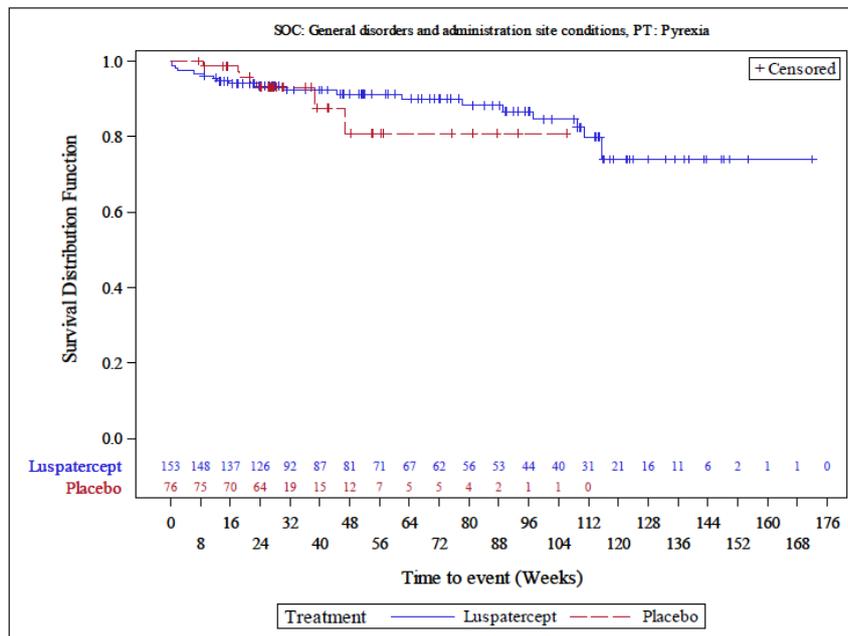
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-67 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Fatigue aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



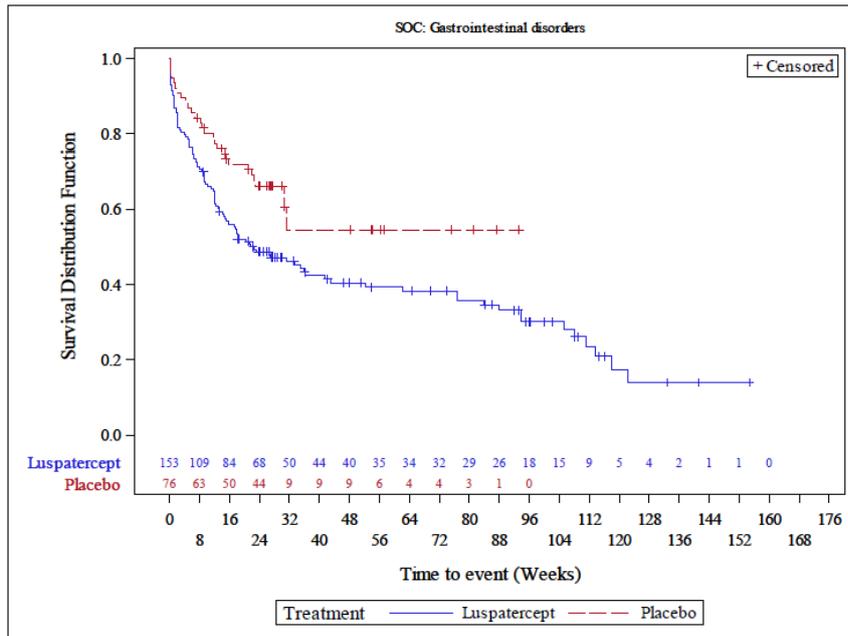
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-68: (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Ödem peripher aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



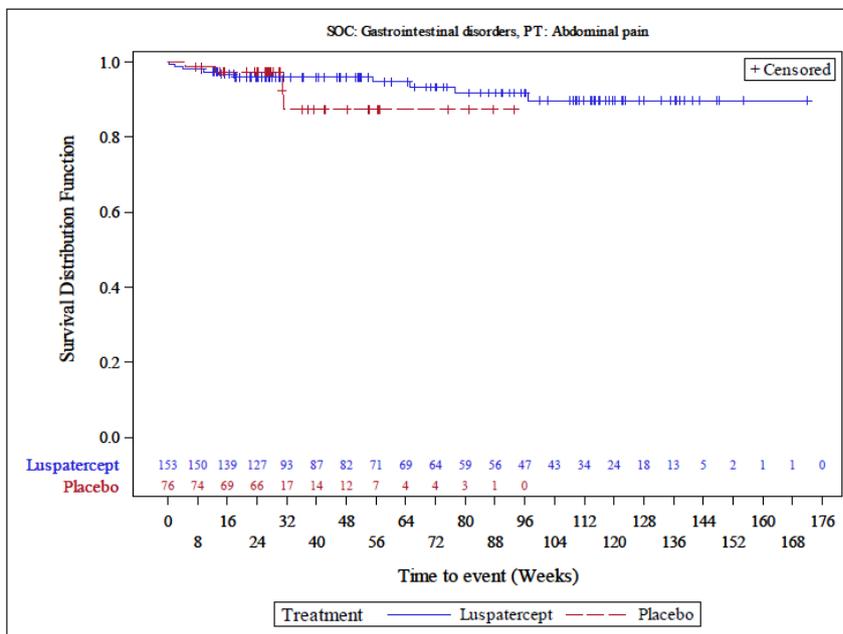
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-69 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Fieber aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



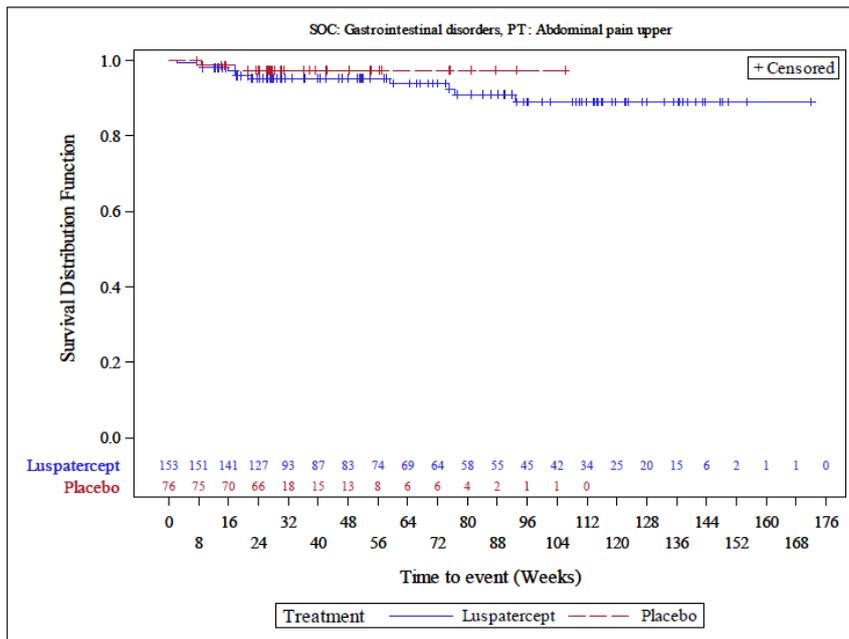
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-70 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Gastrointestinale Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



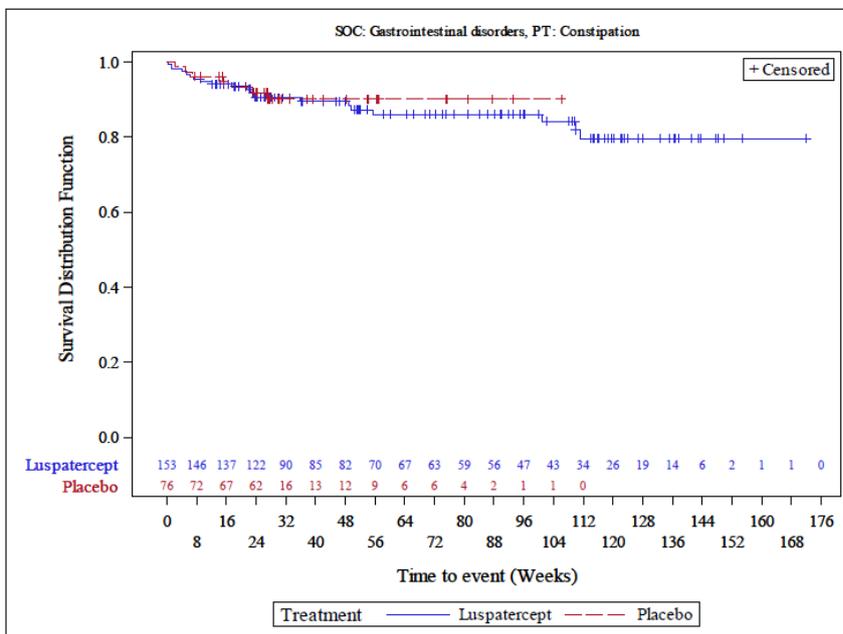
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-71 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Bauchschmerz aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



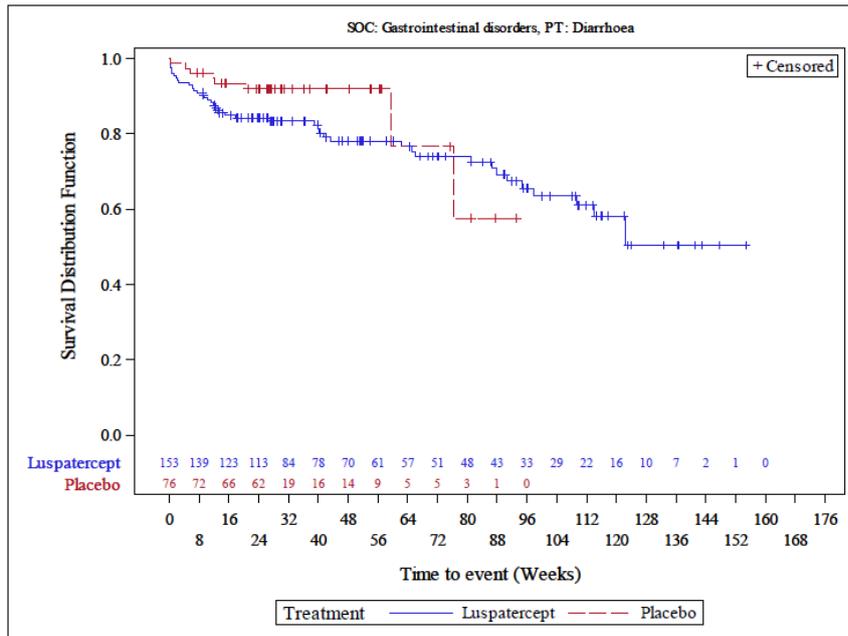
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-72 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Schmerzen im Oberbauch aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



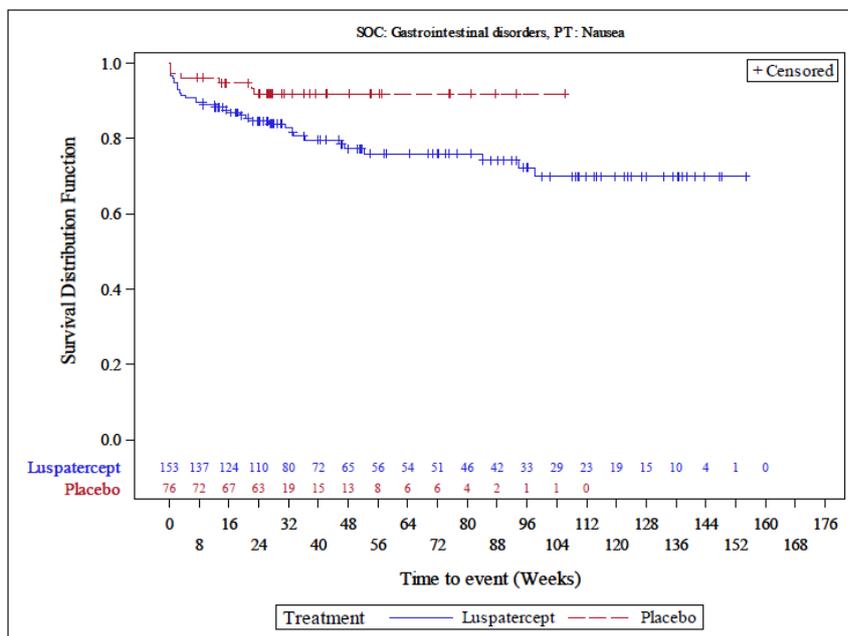
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-73 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Obstipation aus MEDALIST; Woche Woche eins bis zum Datenschnitt vom 01.07.2019



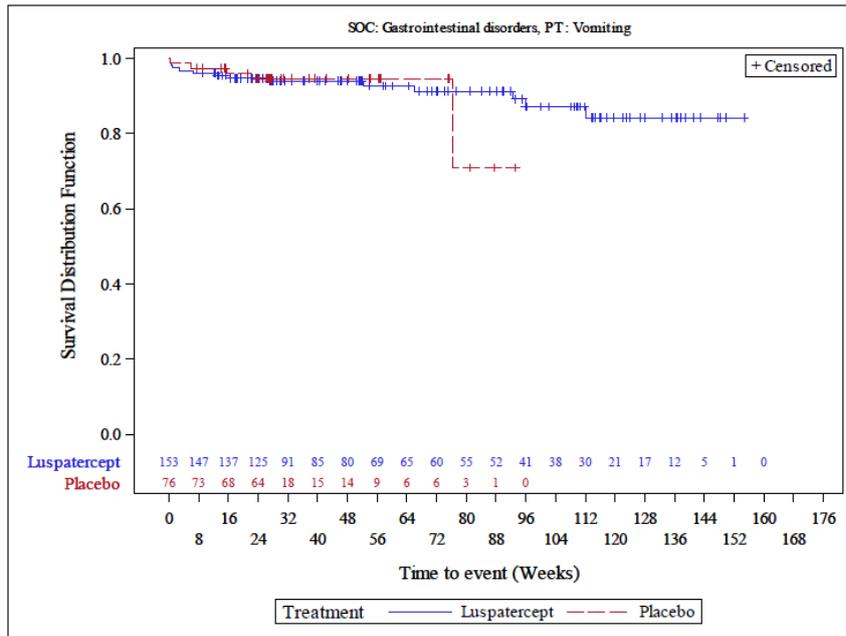
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-74 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Diarrhö aus MEDALIST; Woche Woche eins bis zum Datenschnitt vom 01.07.2019



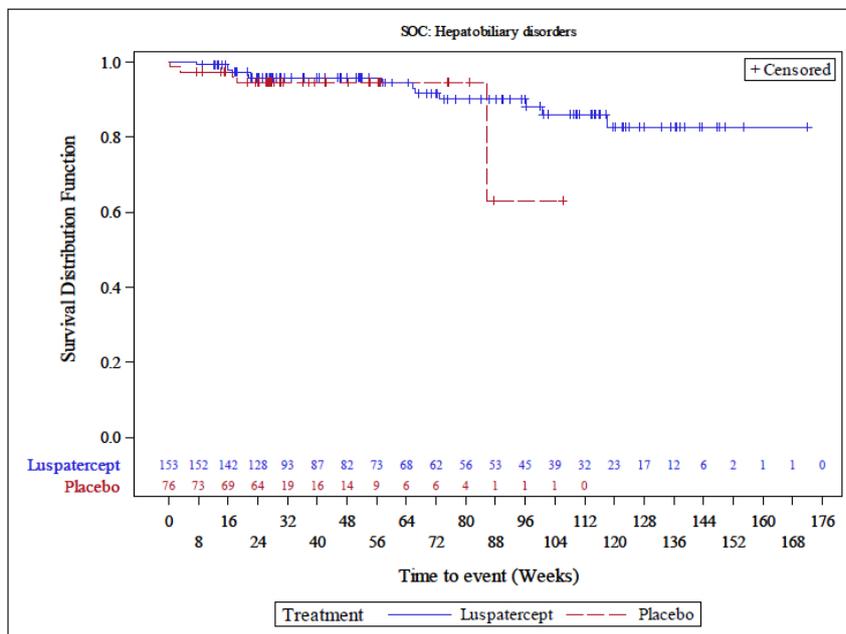
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-75 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Übelkeit aus MEDALIST; Woche Woche eins bis zum Datenschnitt vom 01.07.2019



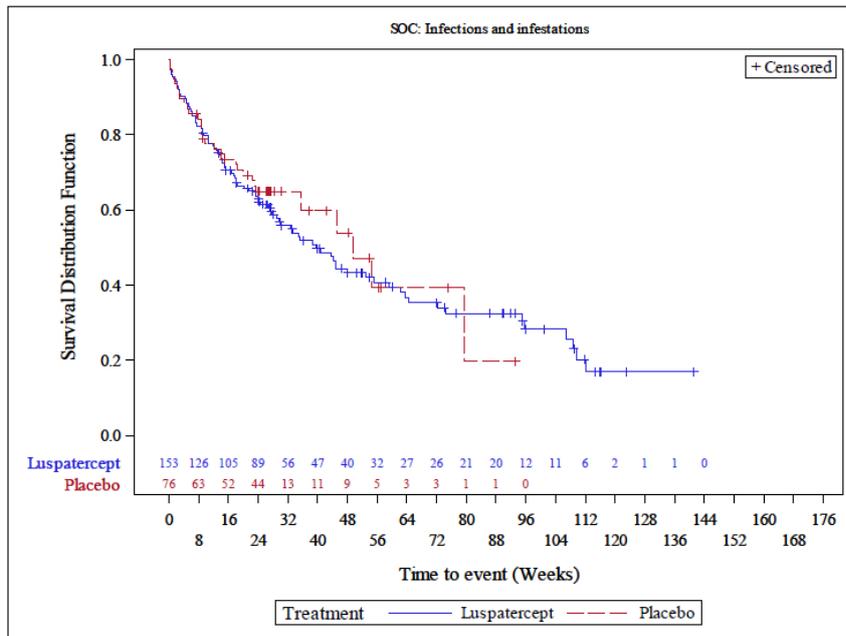
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-76 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Erbrechen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



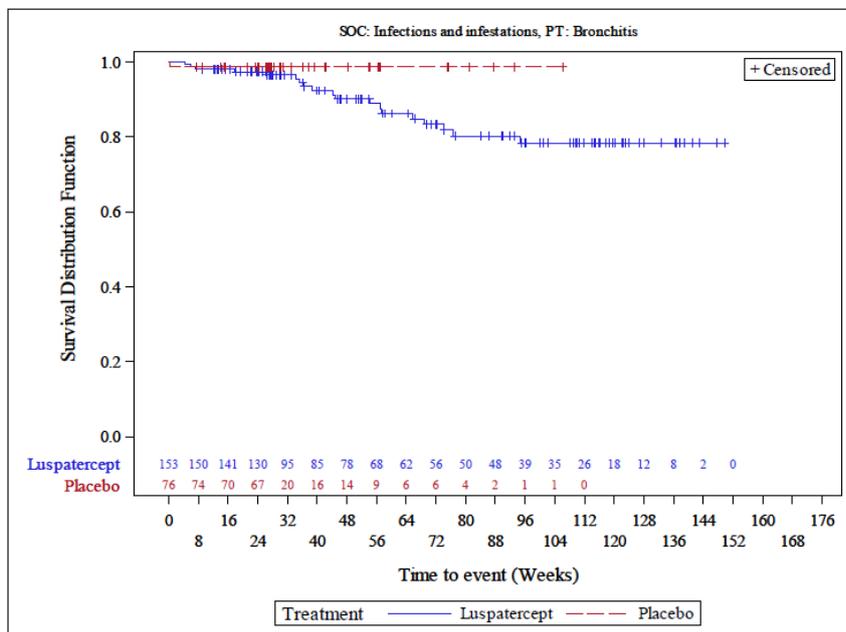
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-77 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Hepatobiliäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



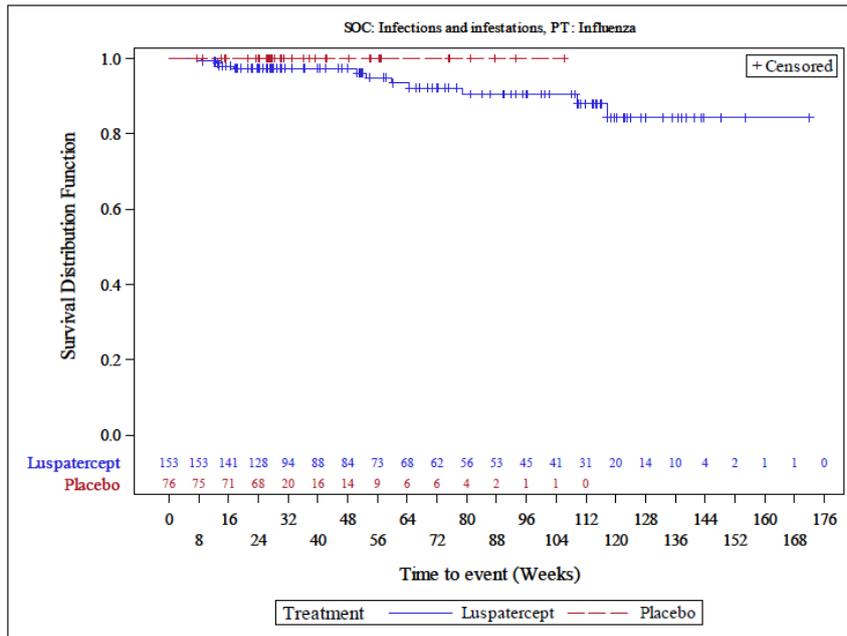
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-78 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Infektionen und parasitäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



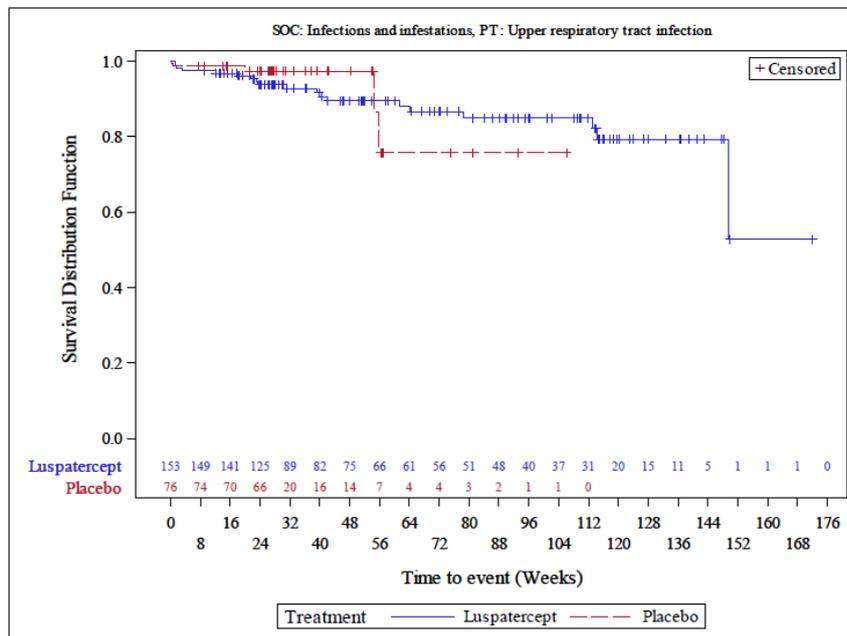
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-79 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Bronchitis aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



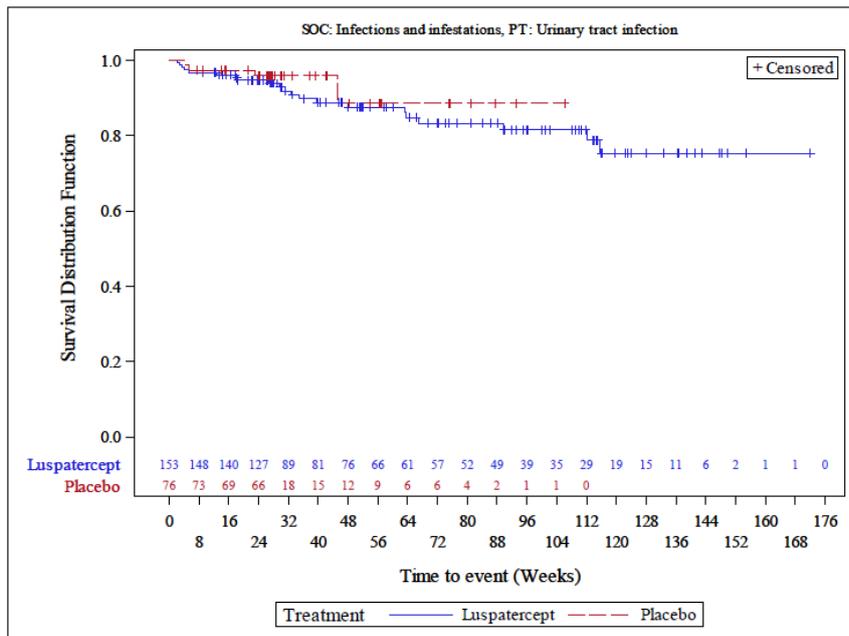
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-80 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Grippe aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



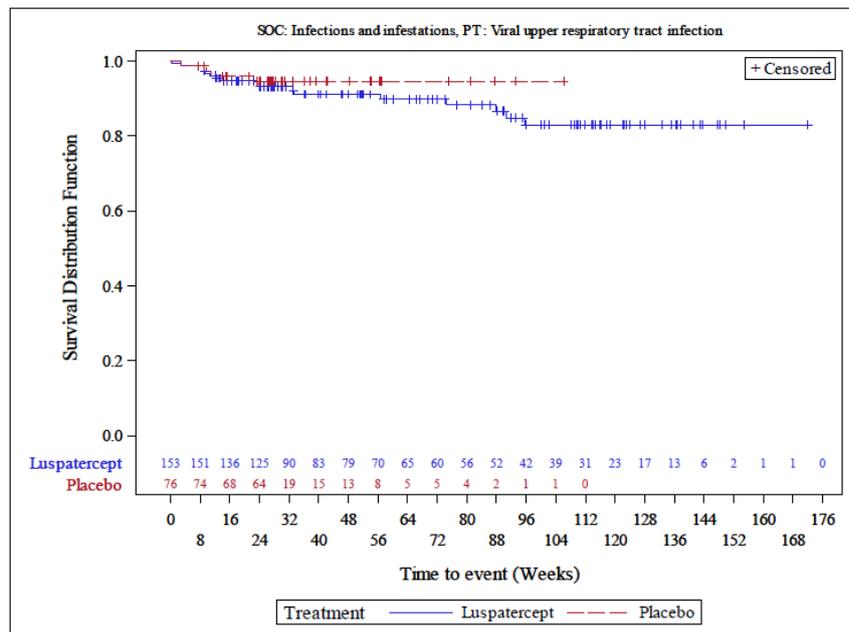
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-81 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Infektion der oberen Atemwege aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



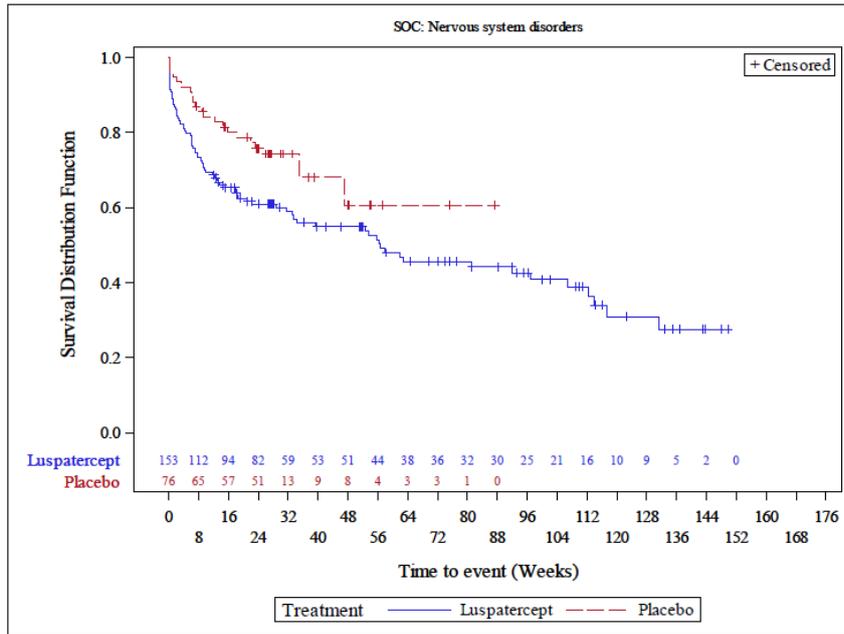
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-82 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Harnwegsinfektion aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



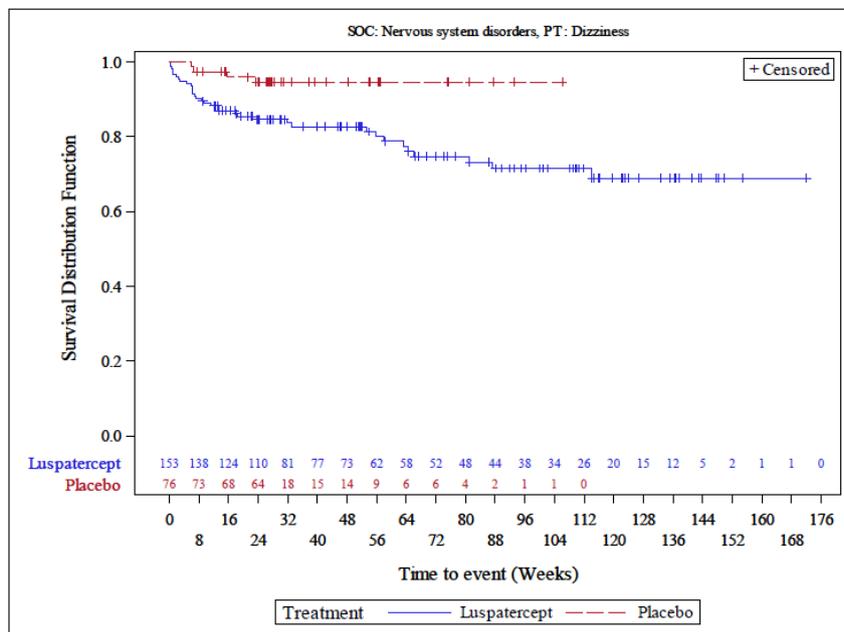
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-83 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT virale Infektion der oberen Atemwege aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



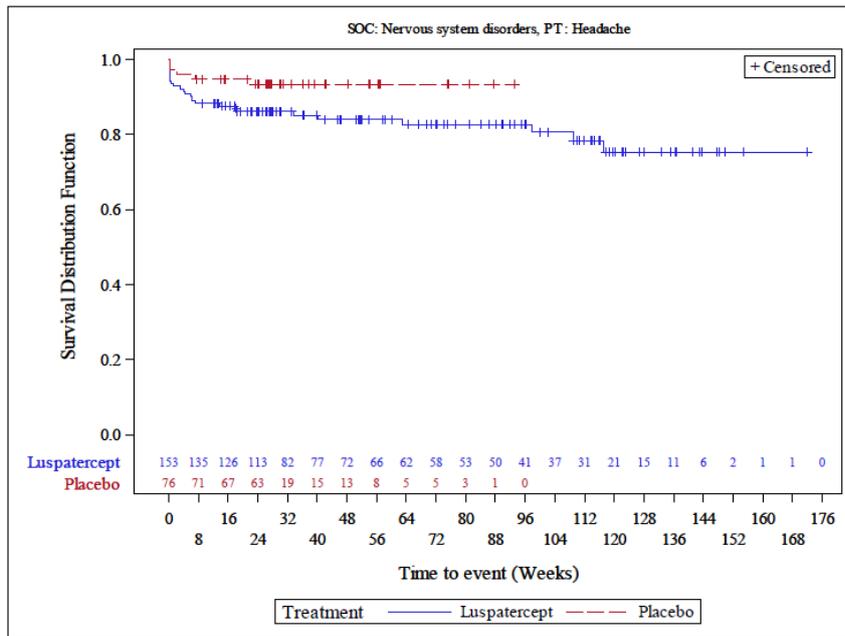
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-84 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen des Nervensystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



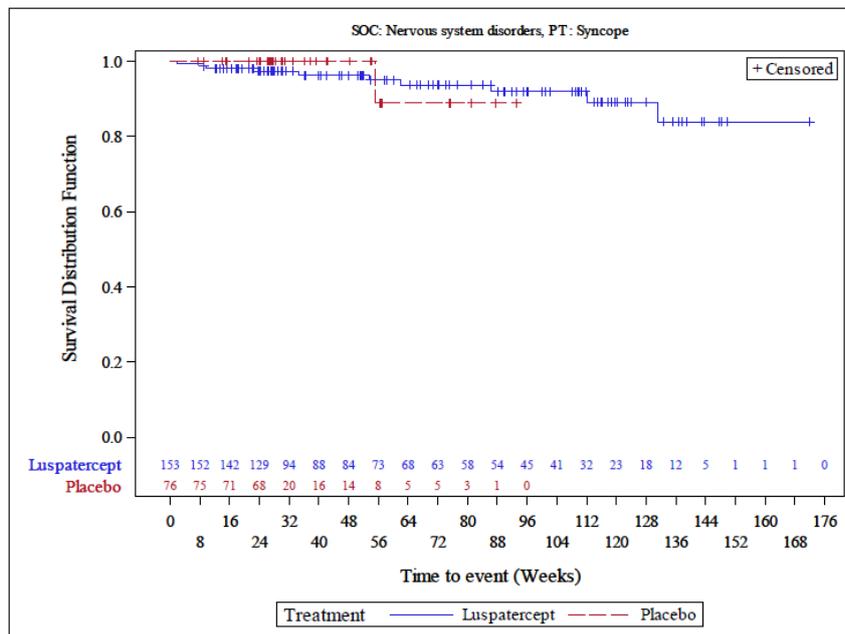
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-85 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Schwindelgefühl aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



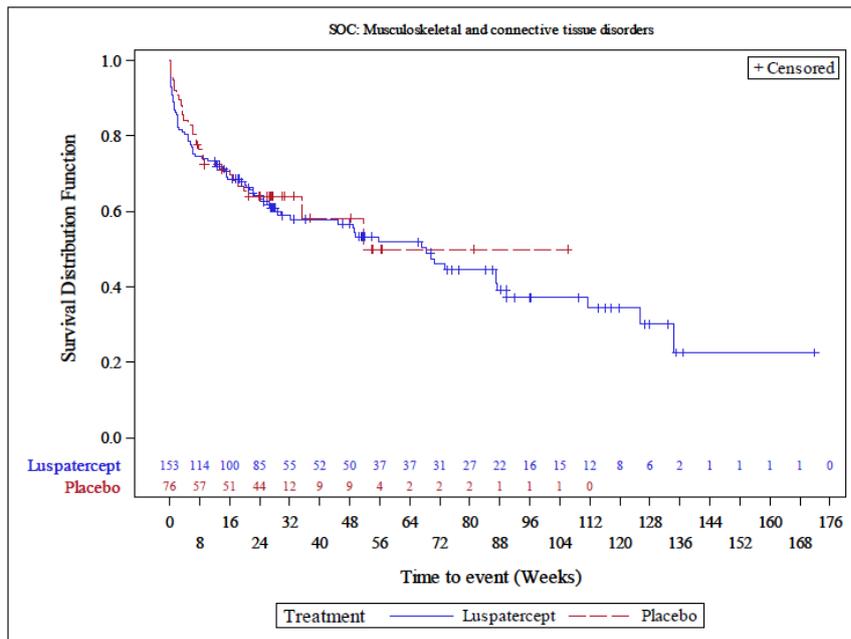
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-86 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Kopfschmerzen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



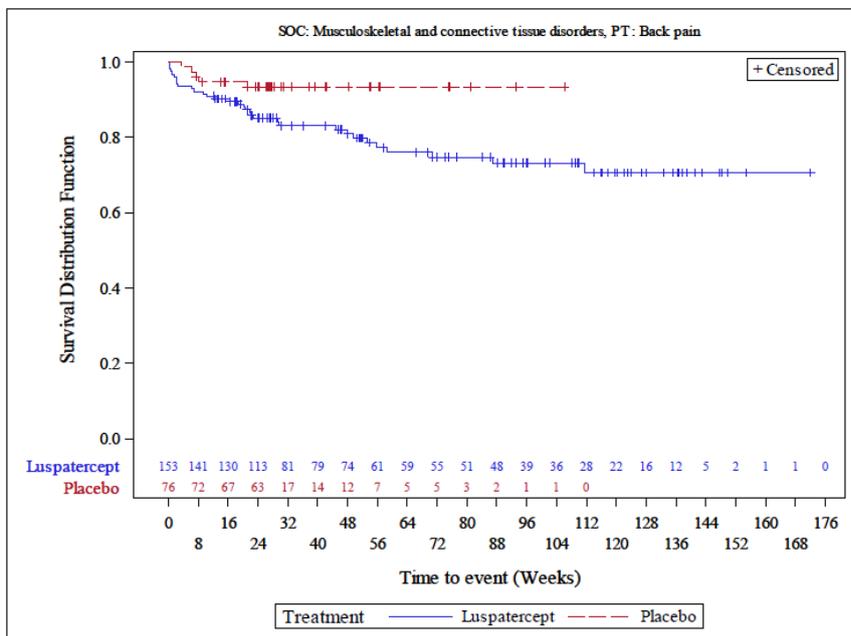
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-87 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Synkope aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



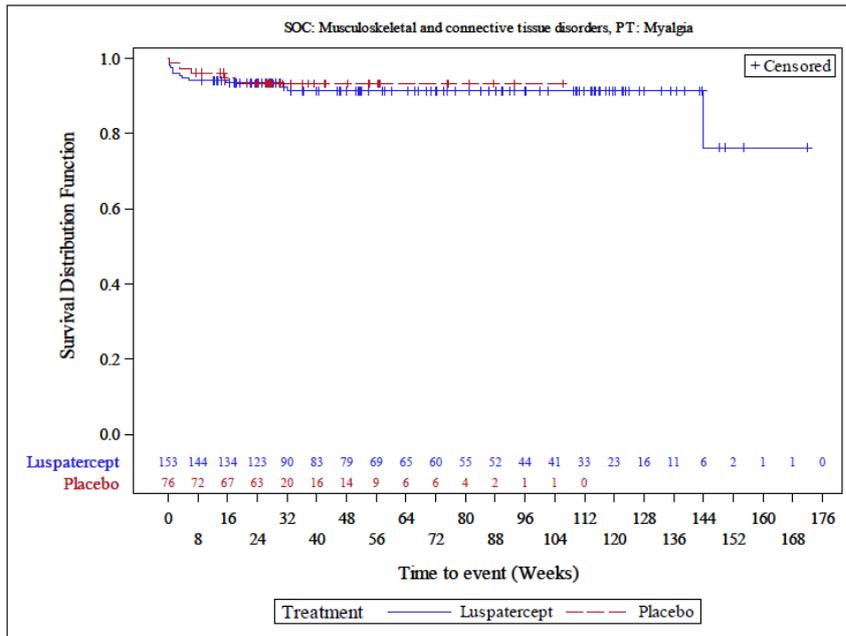
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-88 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



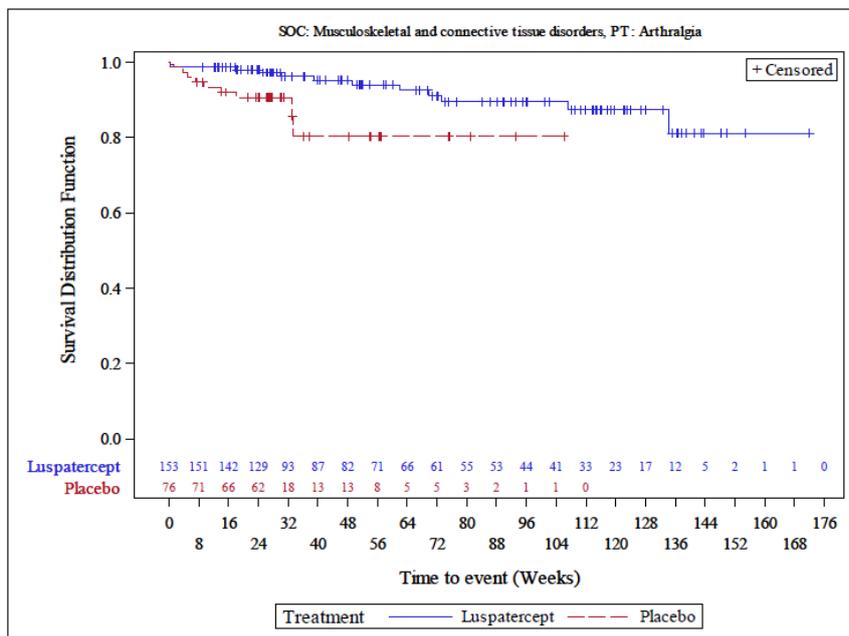
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-89 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Rückenschmerzen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



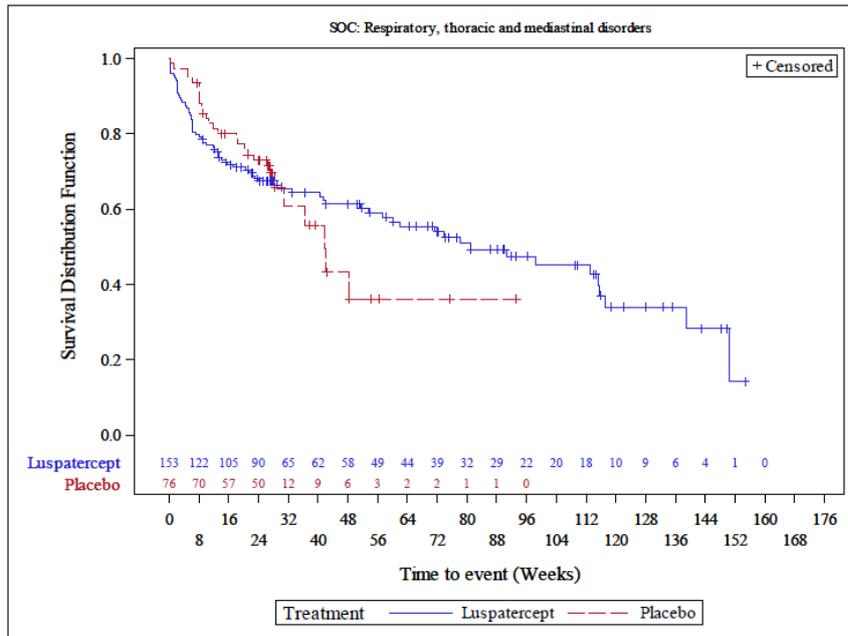
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-90 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Myalgie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



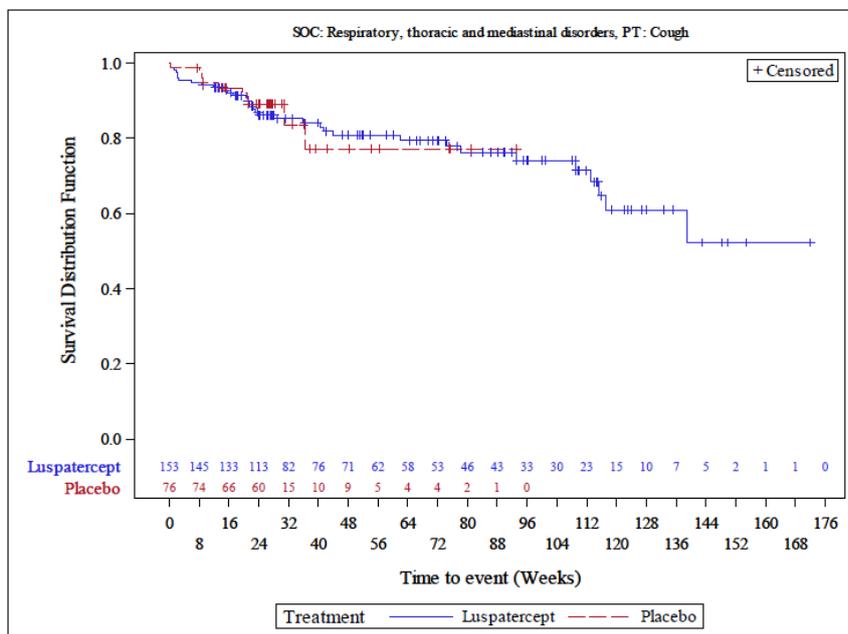
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-91 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Arthralgie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



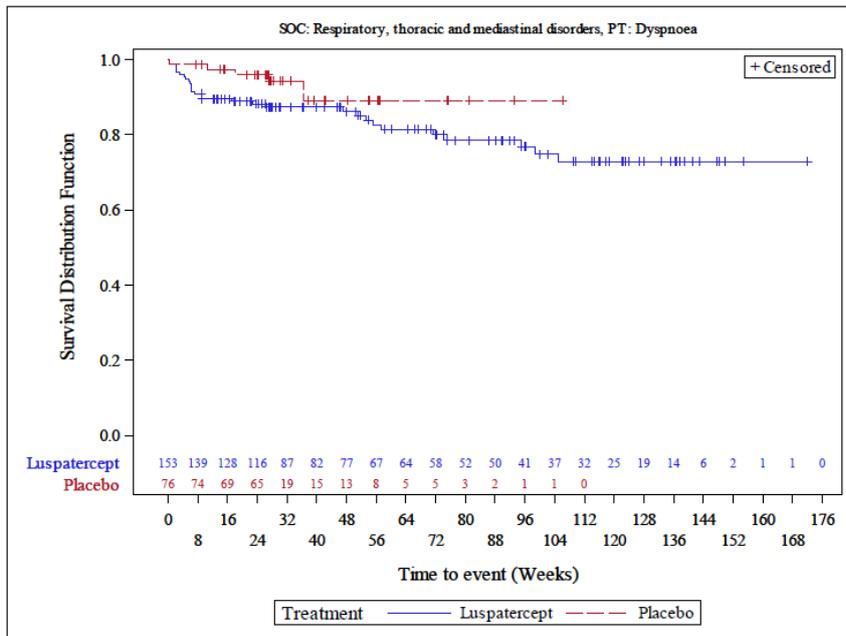
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-92 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



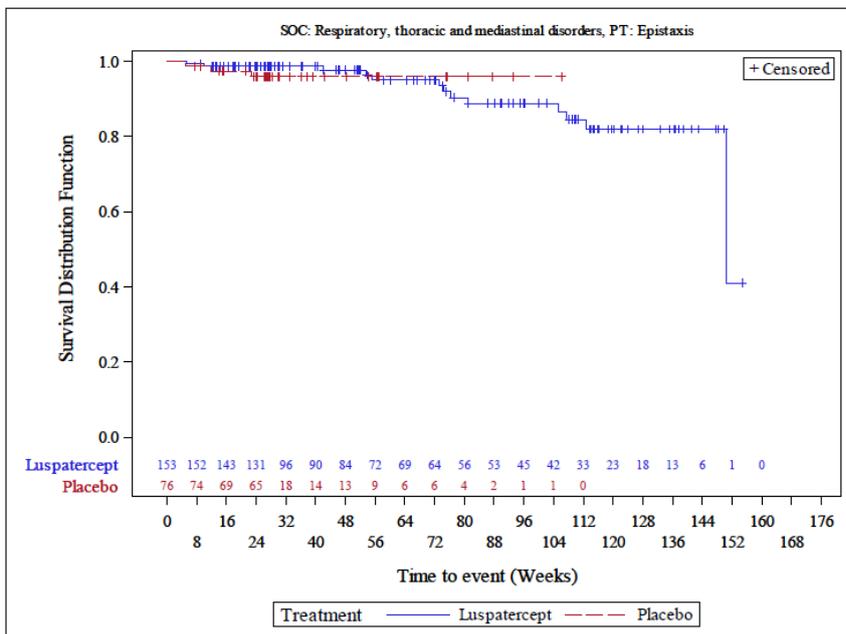
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-93 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Husten aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



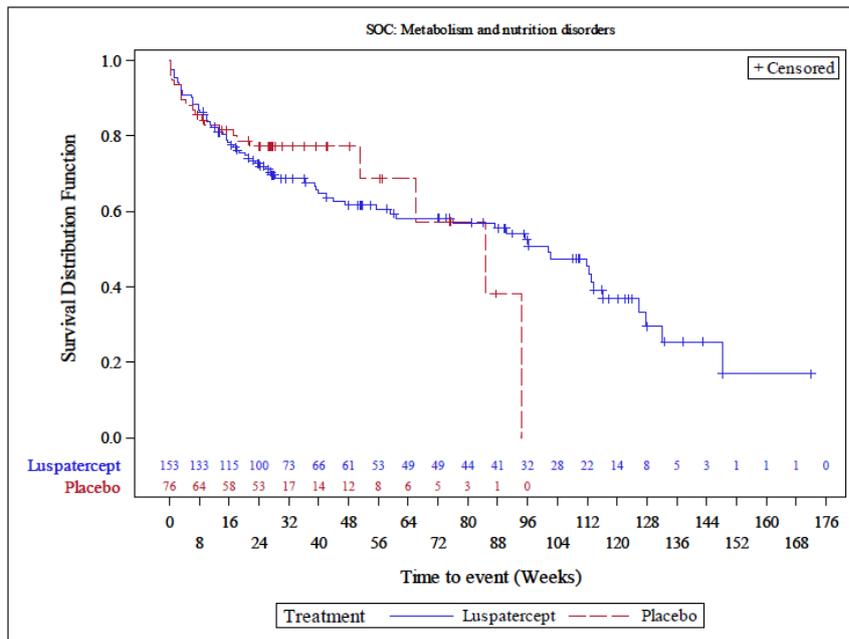
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-94 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Dyspnoe aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



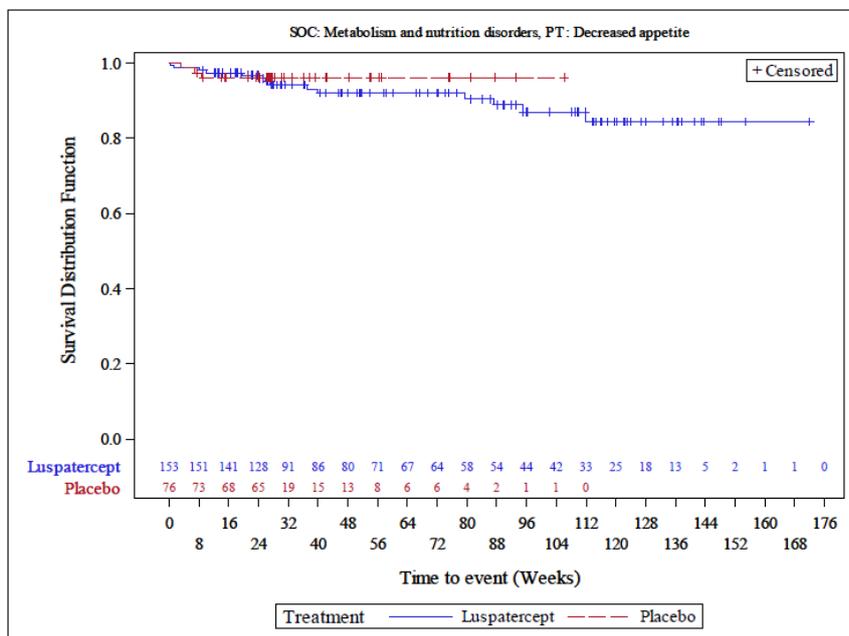
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-95 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Epistaxis aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



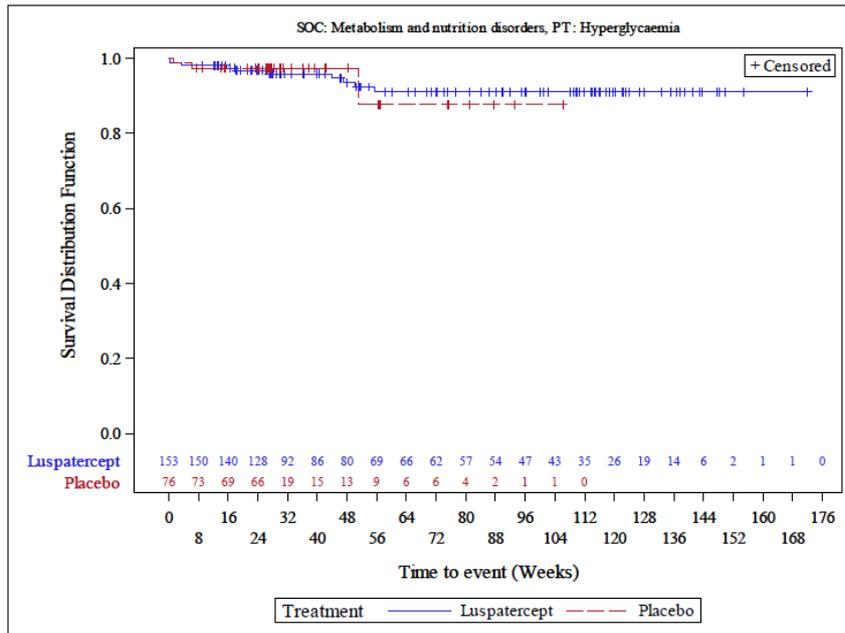
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-96 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Stoffwechsel- und Ernährungsstörungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



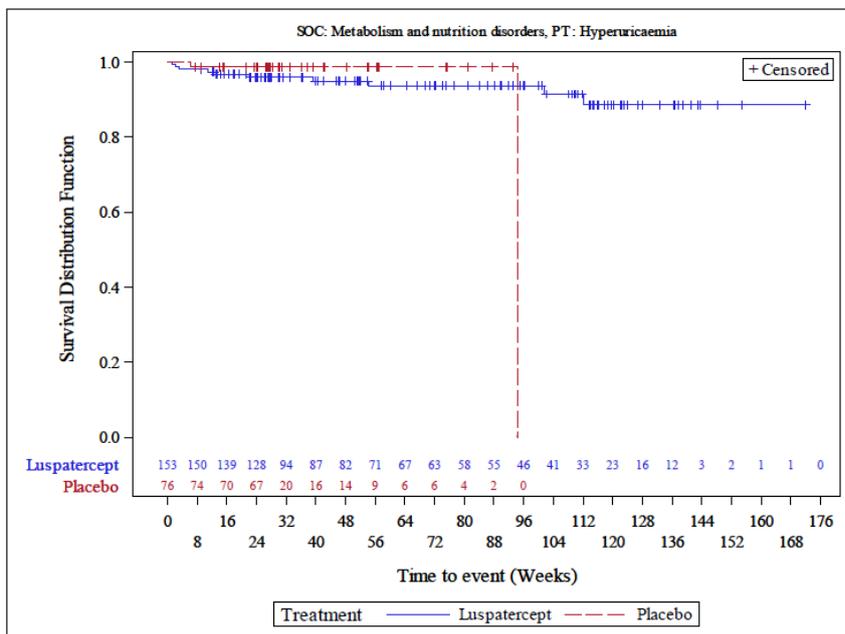
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-97 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT verminderter Appetit aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



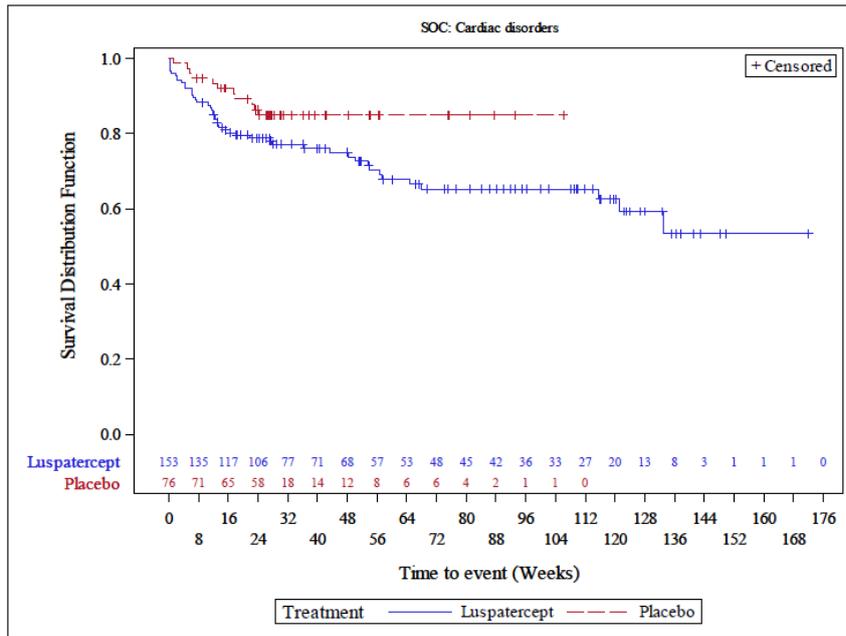
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-98 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Hyperglykämie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



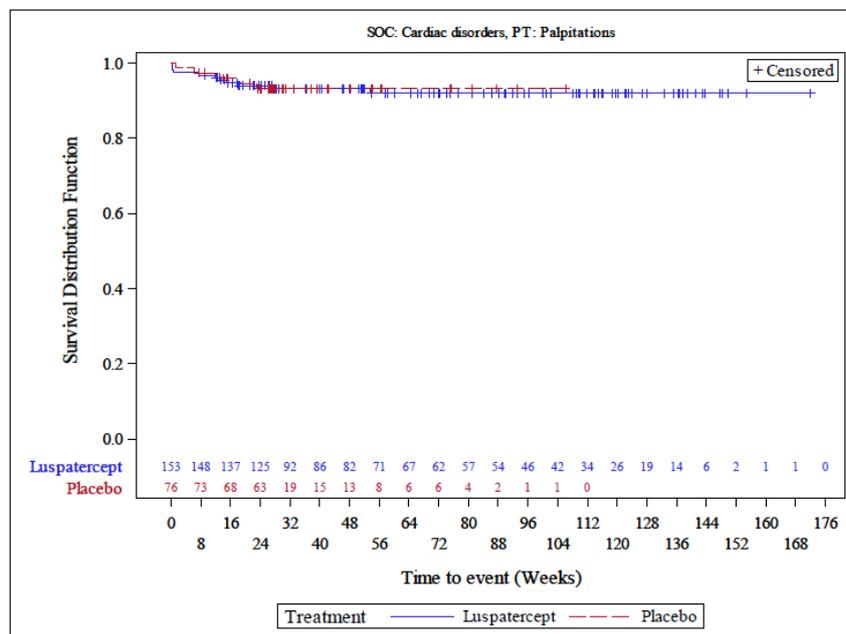
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-99 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Hyperurikämie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



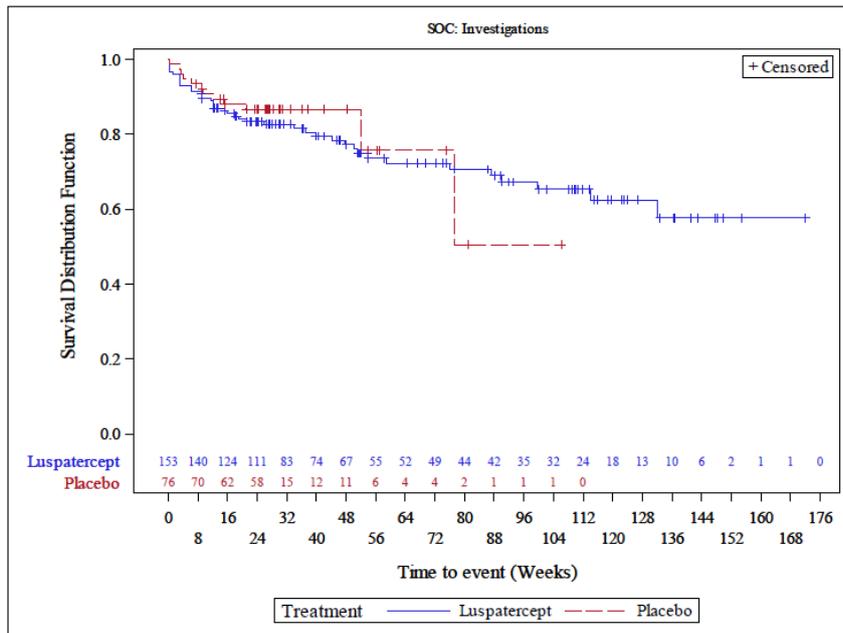
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-100 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Herzerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



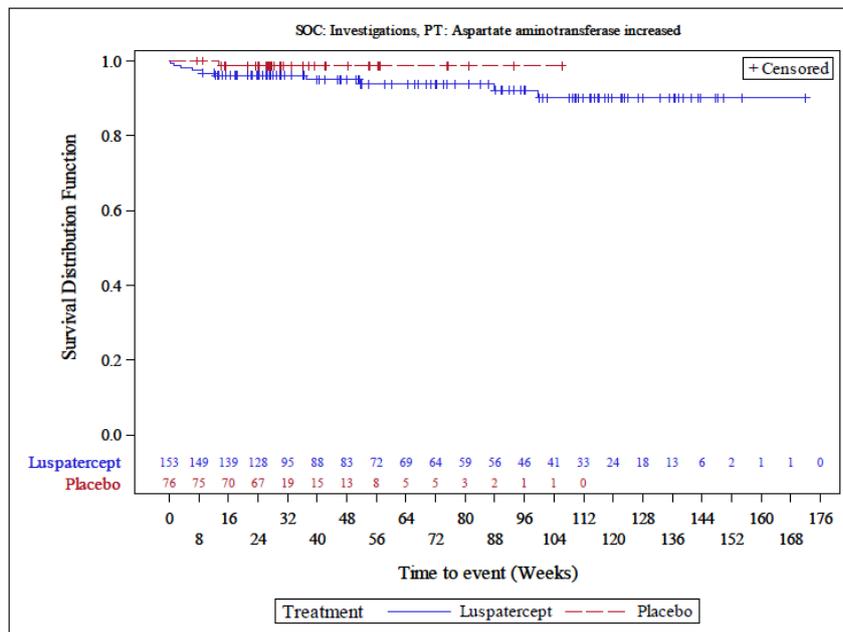
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-101 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Palpitationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



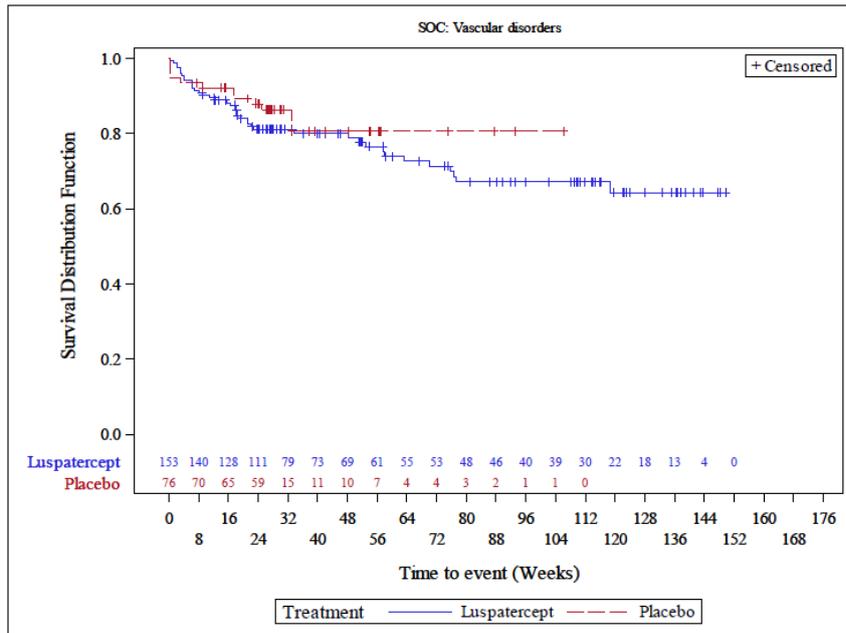
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-102 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Untersuchungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



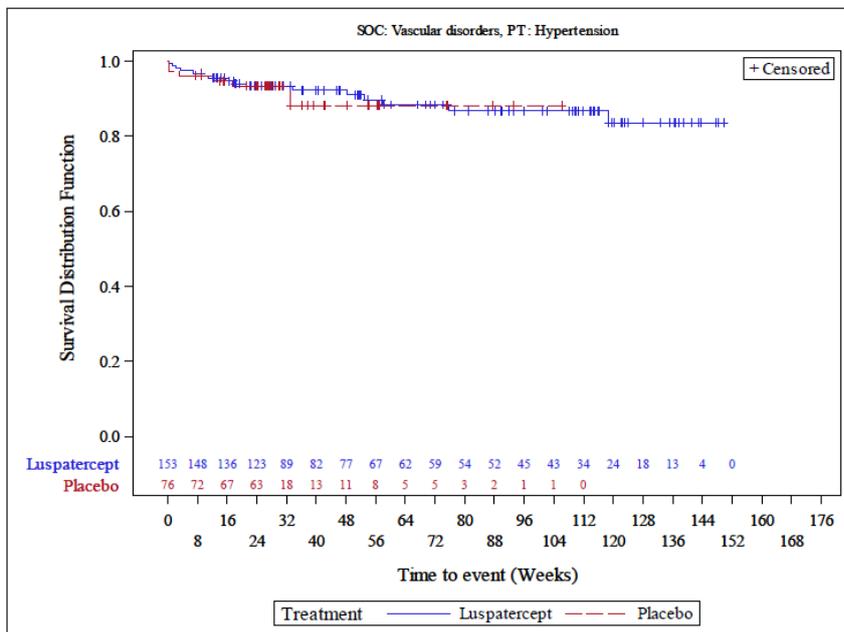
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-103 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Aspartataminotransferase erhöht aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



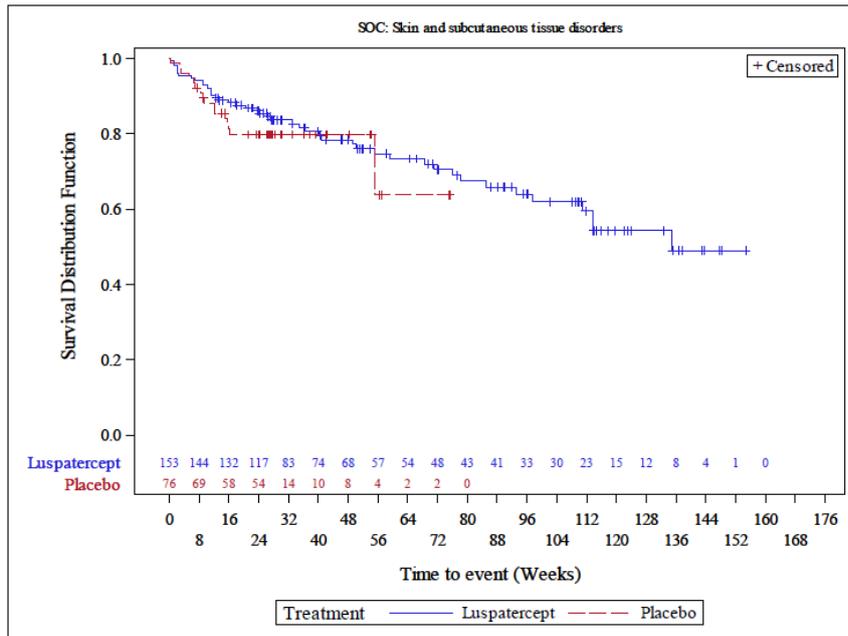
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-104 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Gefäßerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



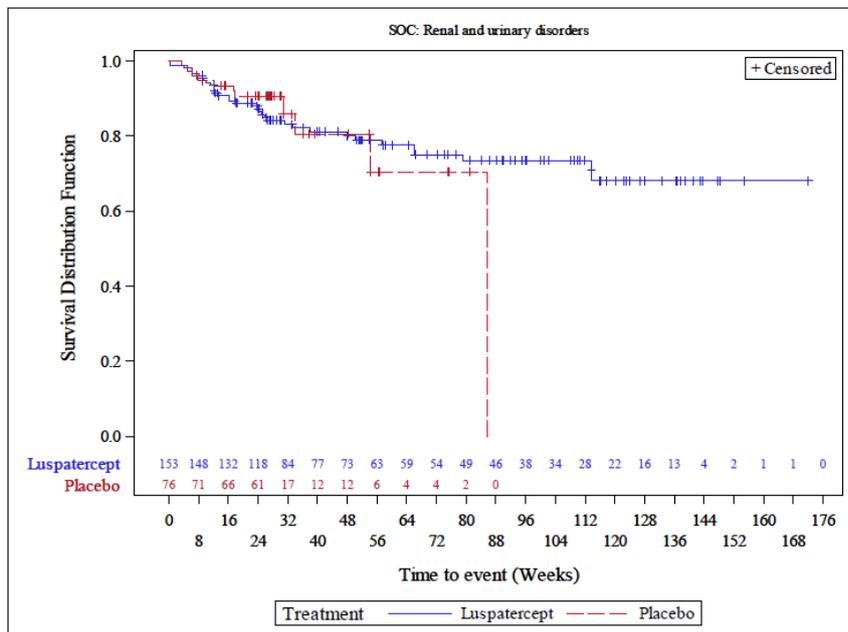
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-105 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Hypertonie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



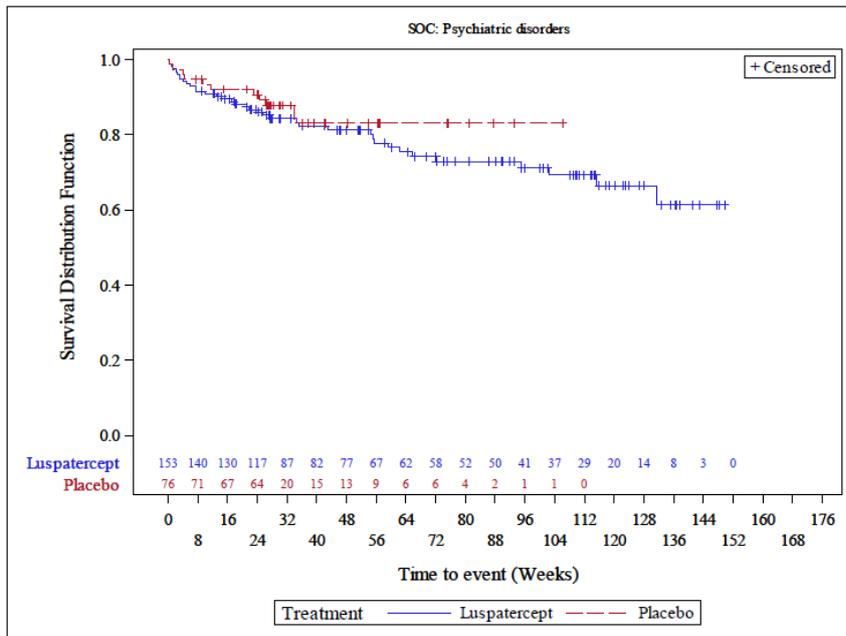
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-106 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen der Haut und des Unterhautgewebes aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



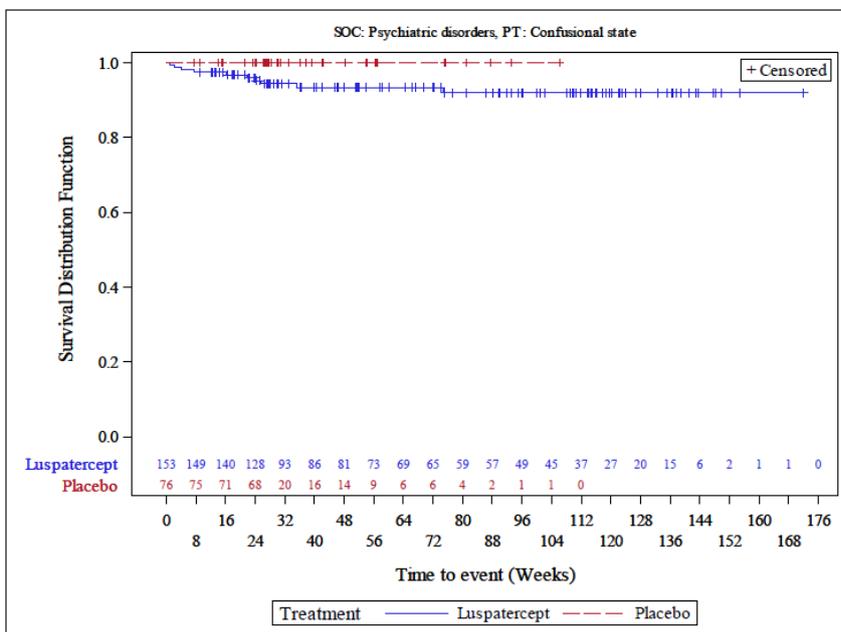
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-107 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen der Niere und Harnwege aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



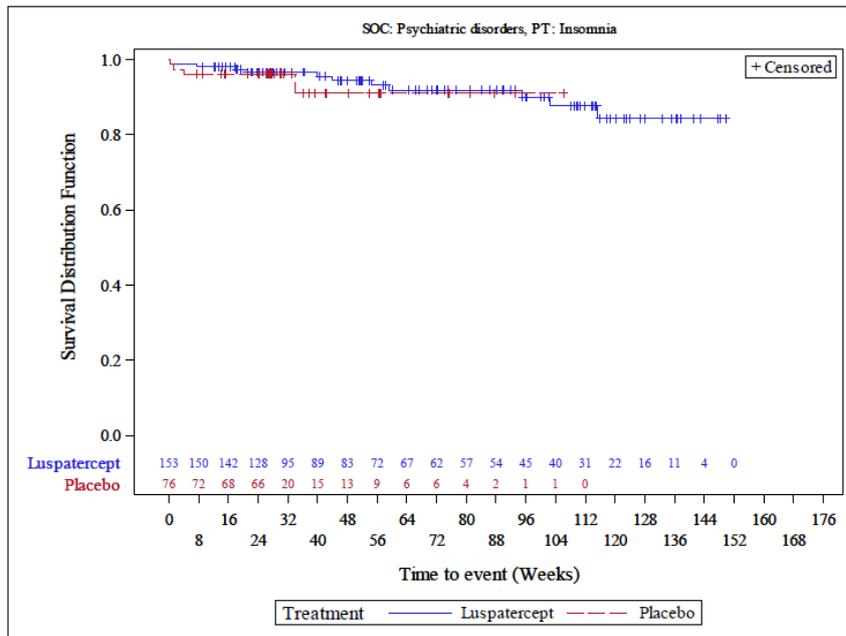
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-108 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Psychiatrische Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



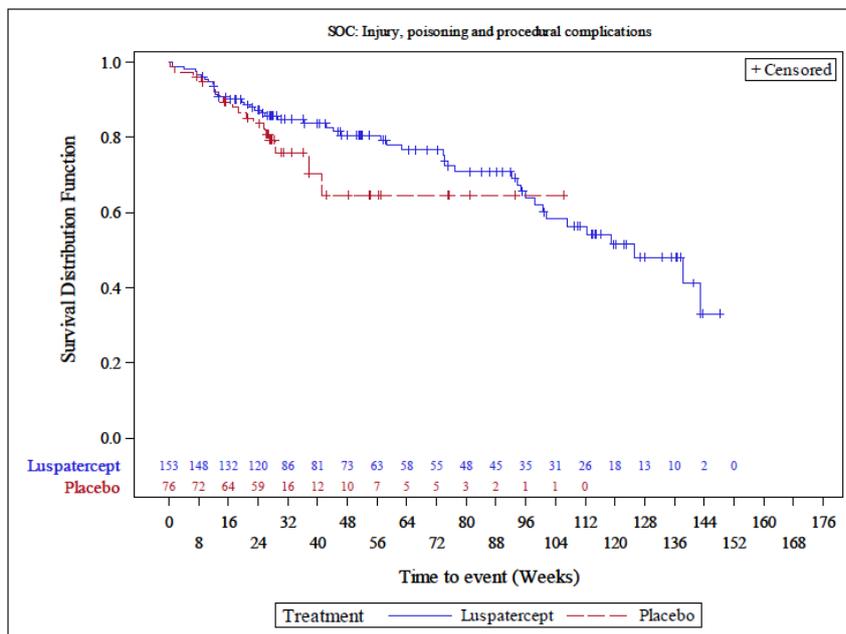
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-109 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Verwirrtheitszustand aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



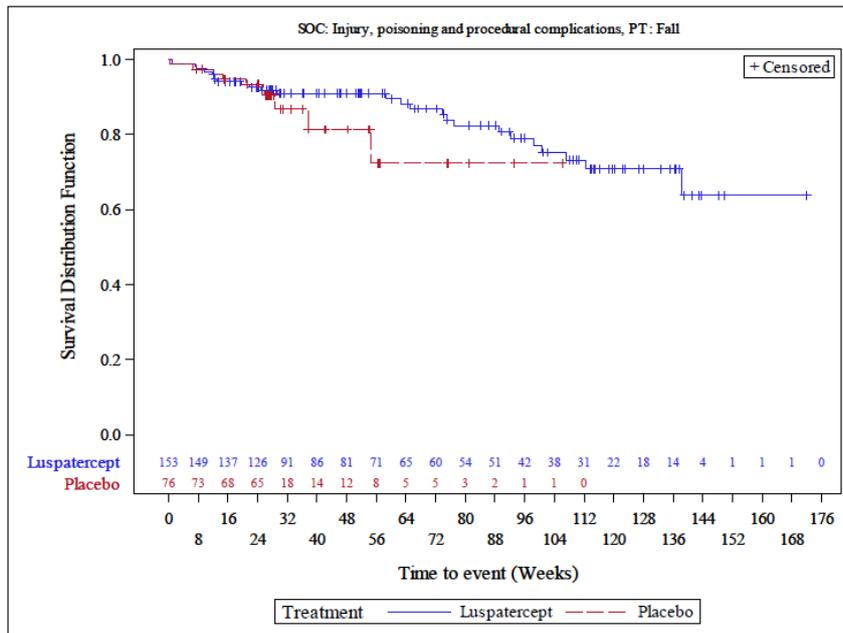
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-110 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Schlaflosigkeit aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



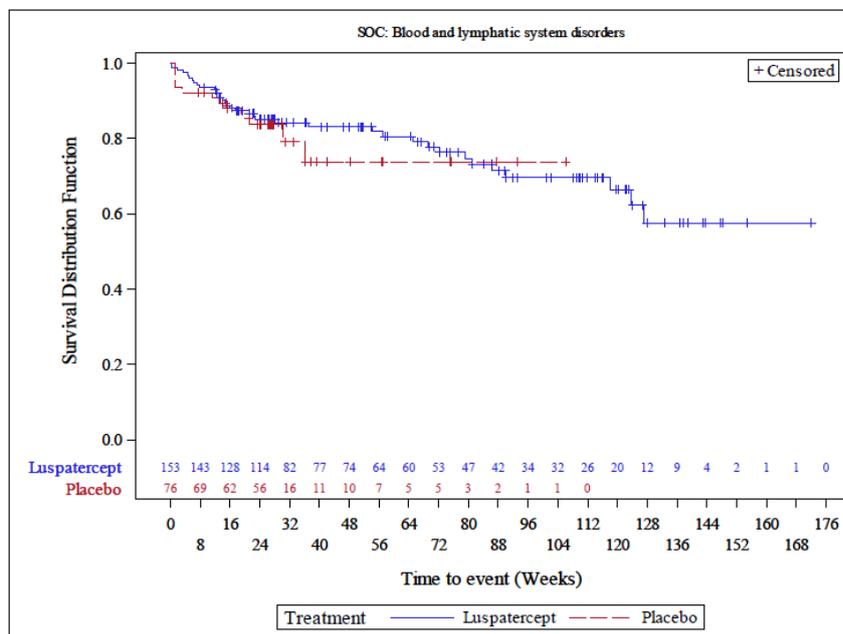
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-111 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



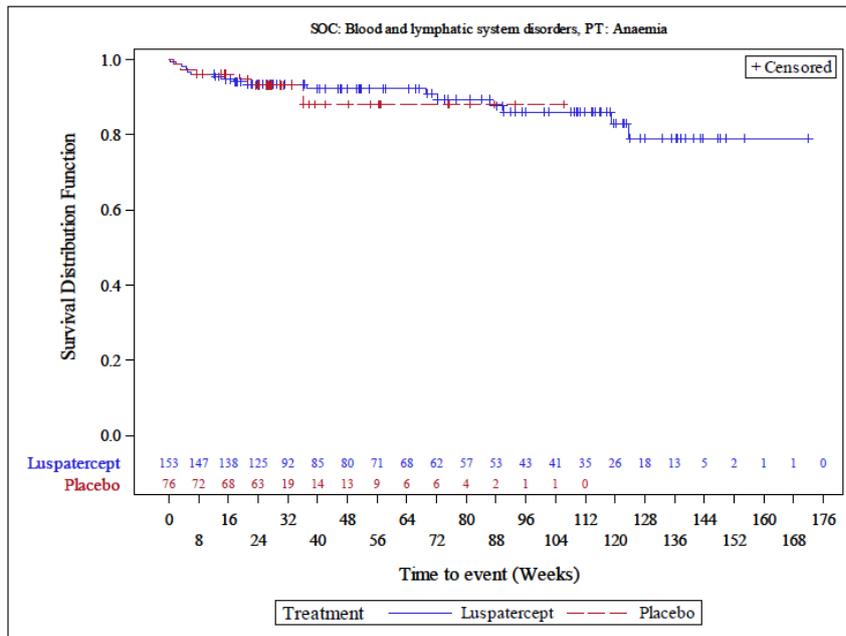
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-112 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Sturz aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



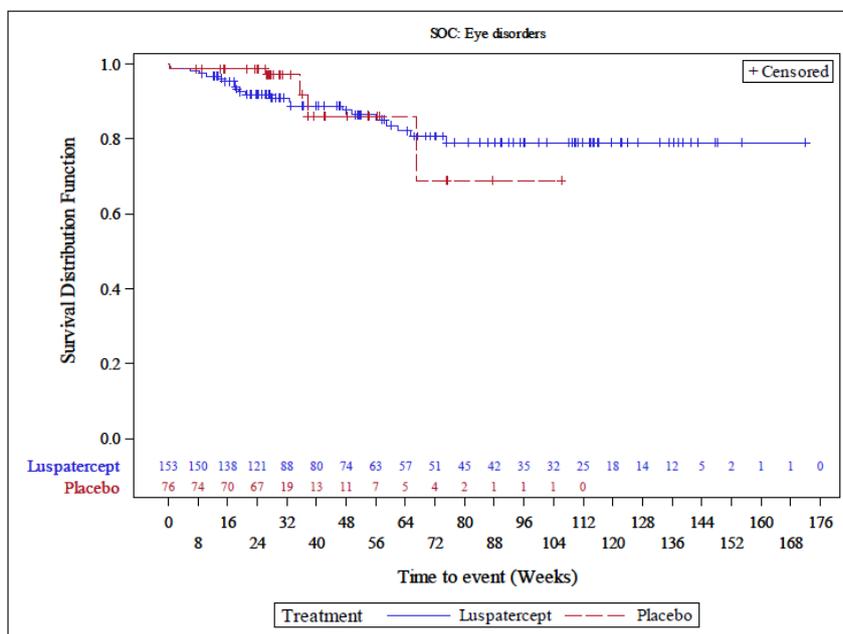
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-113 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen des Blutes und des Lymphsystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



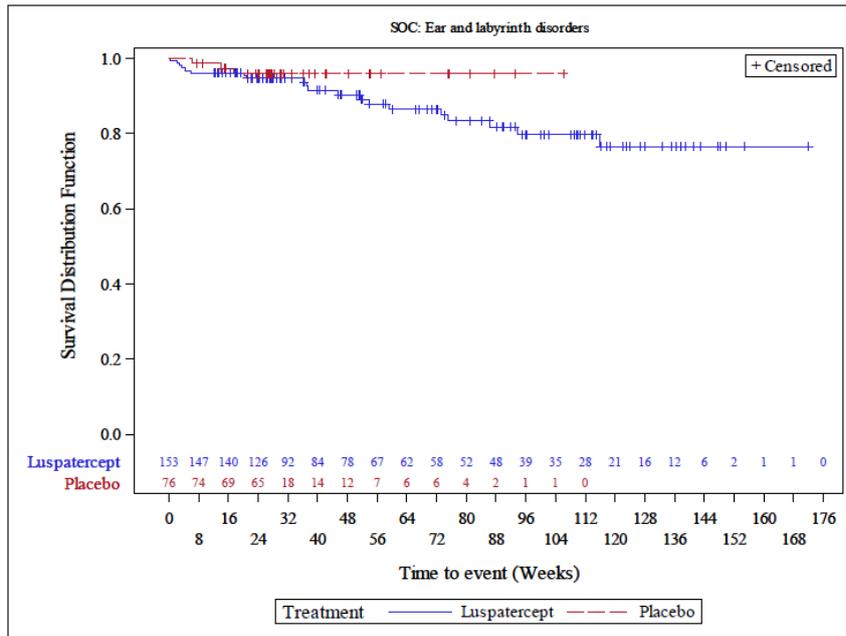
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-114 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Anämie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



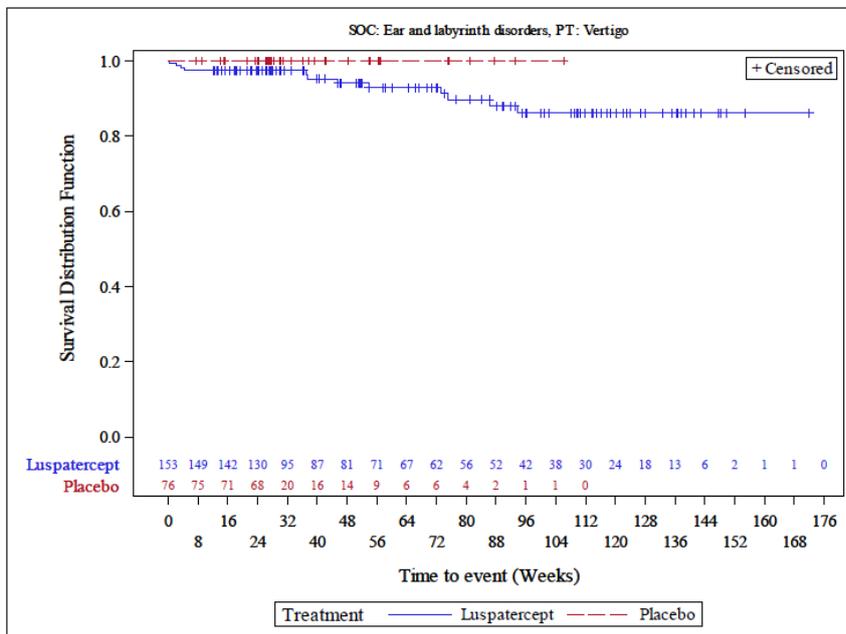
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-115 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Augenerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



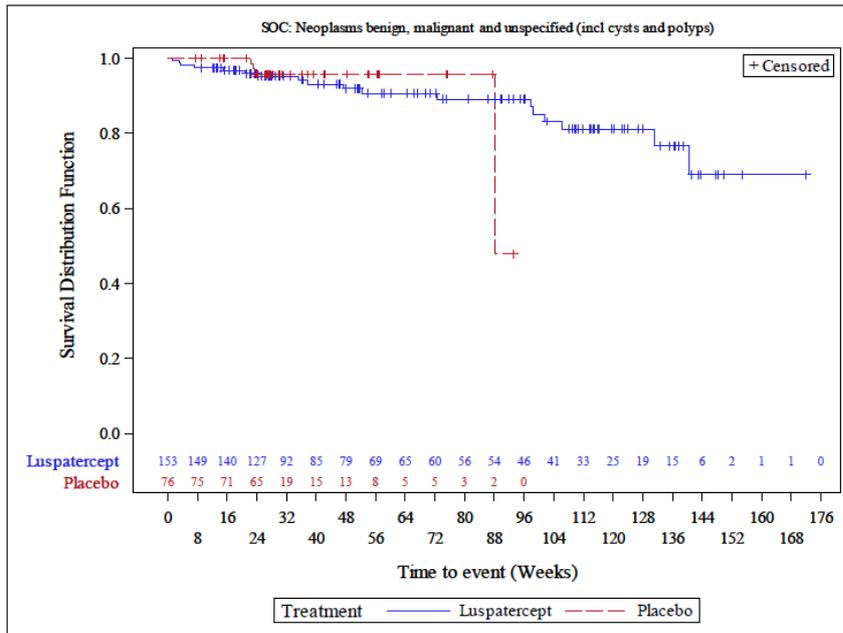
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-116 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen des Ohrs und des Labyrinths aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



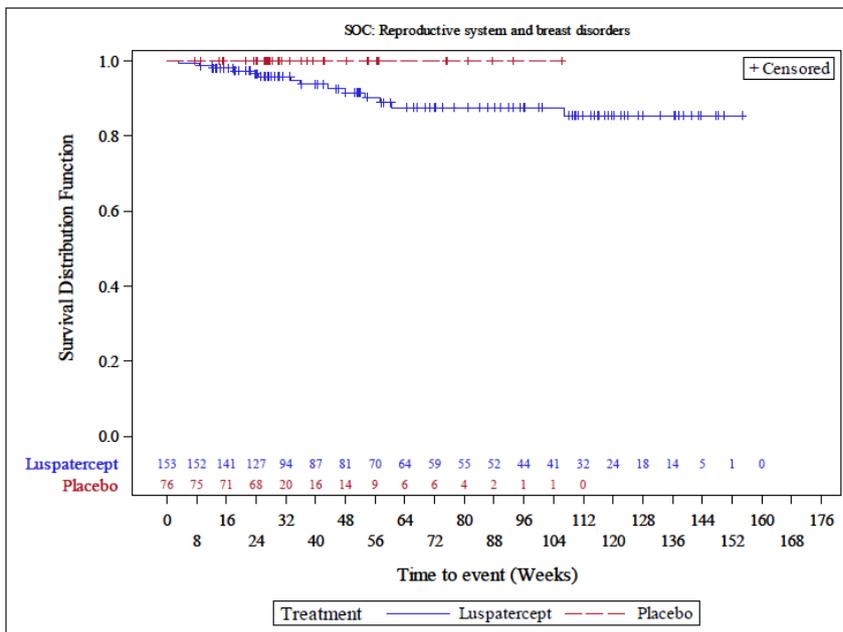
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-117 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE des PT Vertigo aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

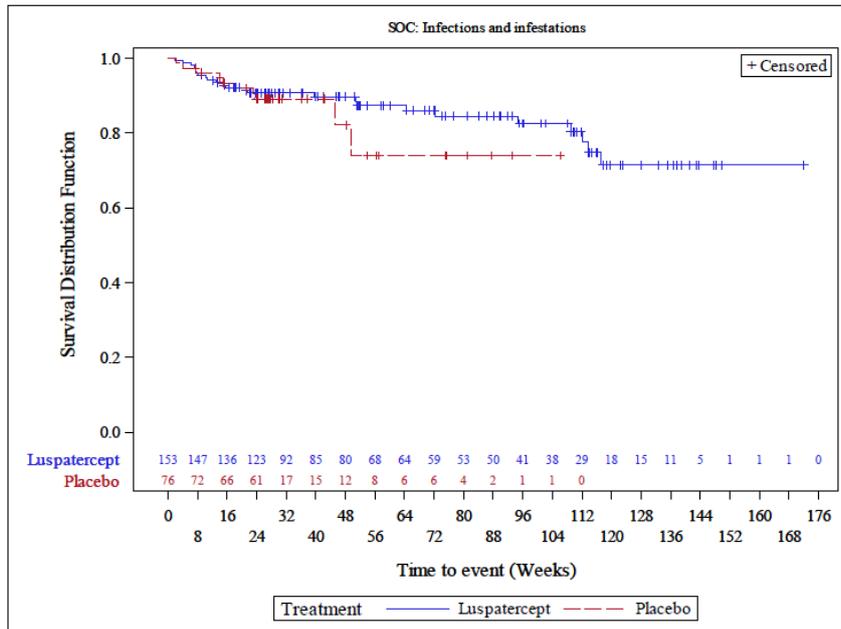
Abbildung 4-118 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Gutartige, bösartige und nicht spezifizierte Neubildungen (inkl. Zysten und Polypen) aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

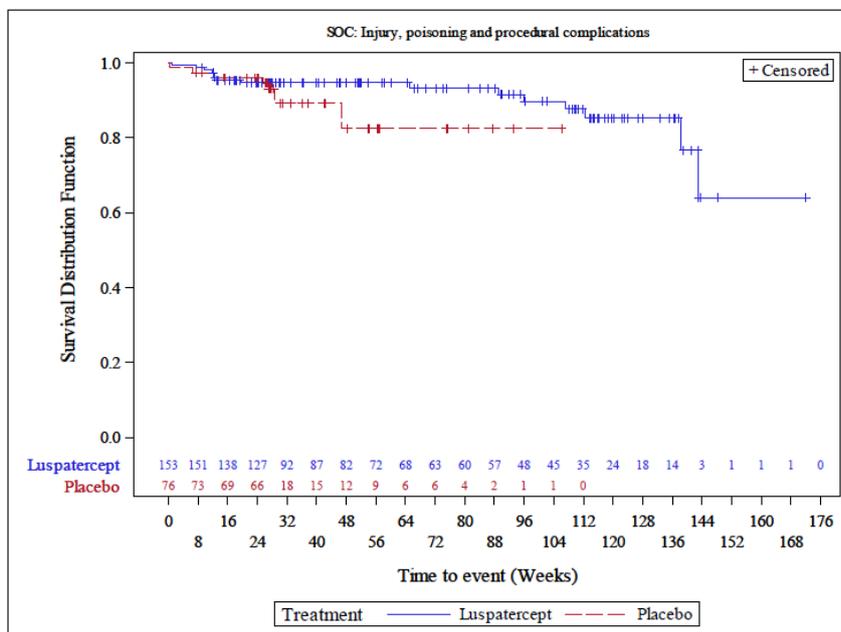
Abbildung 4-119 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten UE der SOC Erkrankungen der Geschlechtsorgane und der Brustdrüse aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

**Schwerwiegende UE**



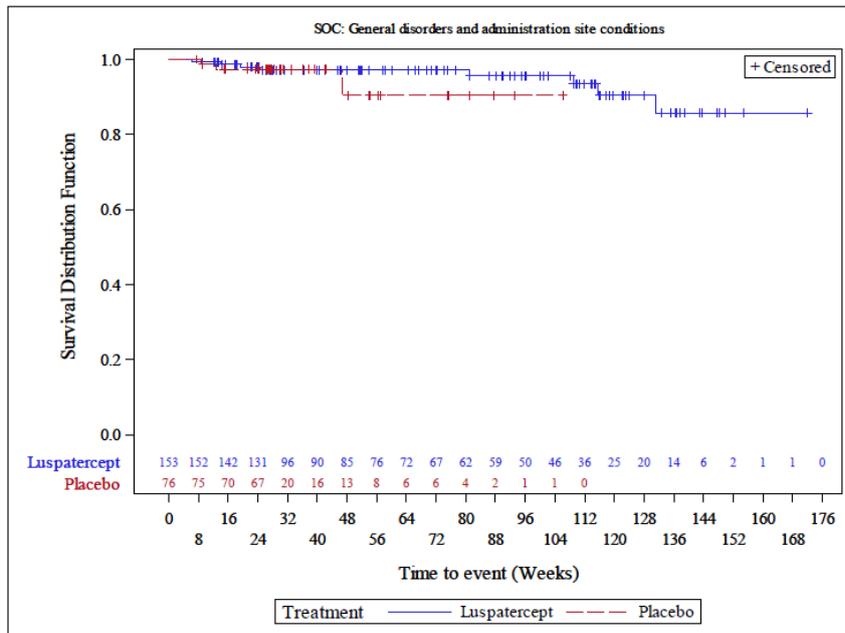
SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

Abbildung 4-120 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Infektionen und parasitäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



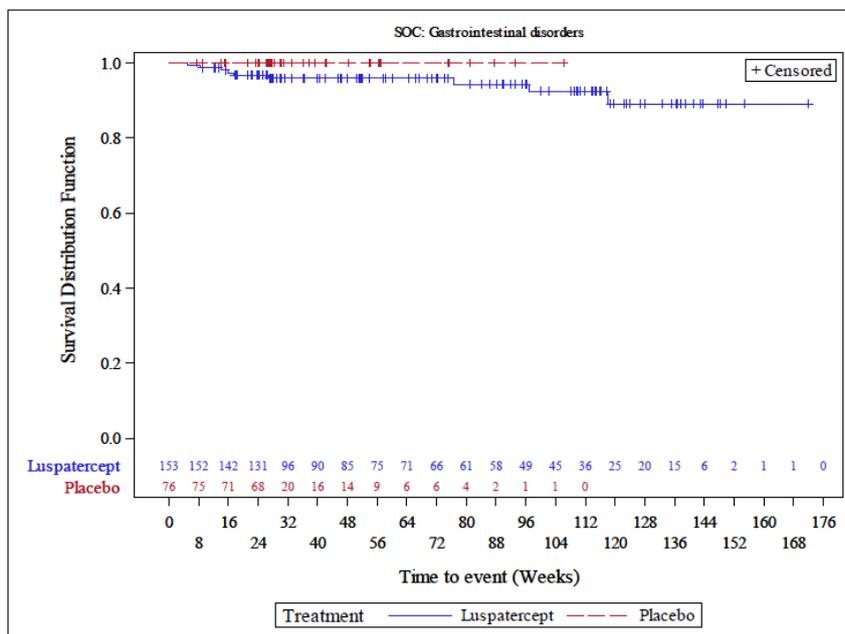
SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

Abbildung 4-121 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



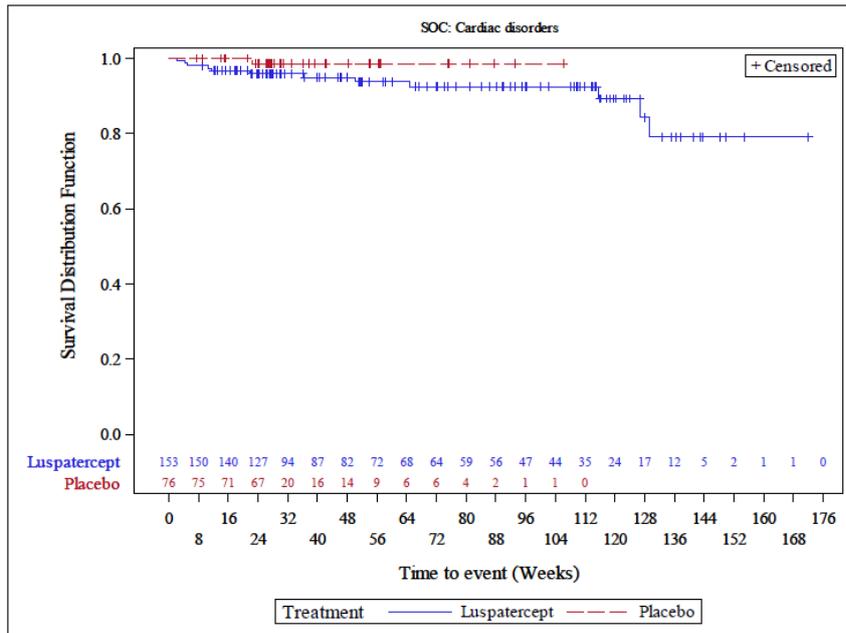
SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

Abbildung 4-122 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



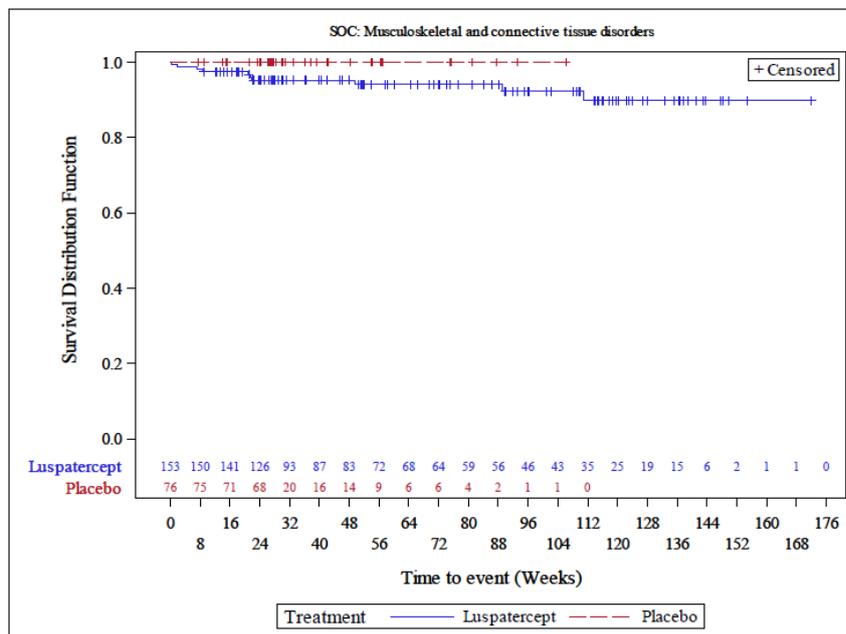
SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

Abbildung 4-123 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Gastrointestinale Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



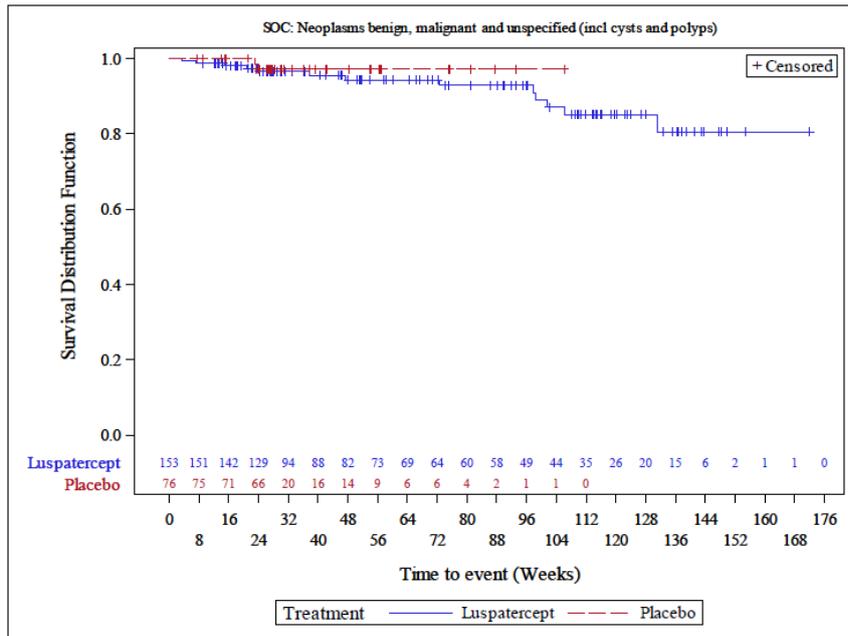
SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

Abbildung 4-124 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Herzerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

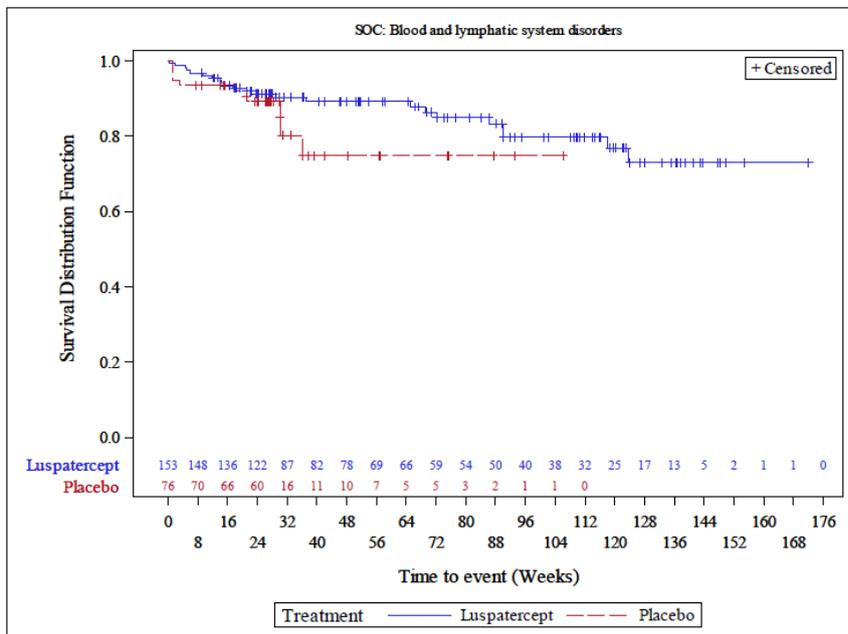
Abbildung 4-125 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

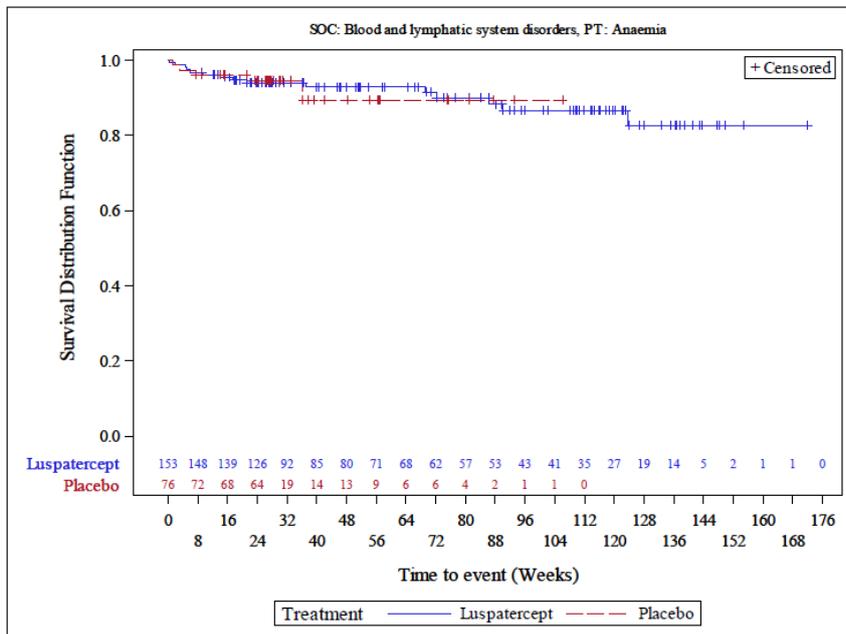
Abbildung 4-126 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen) aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

**Schwere UE (CTCAE ≥ 3)**



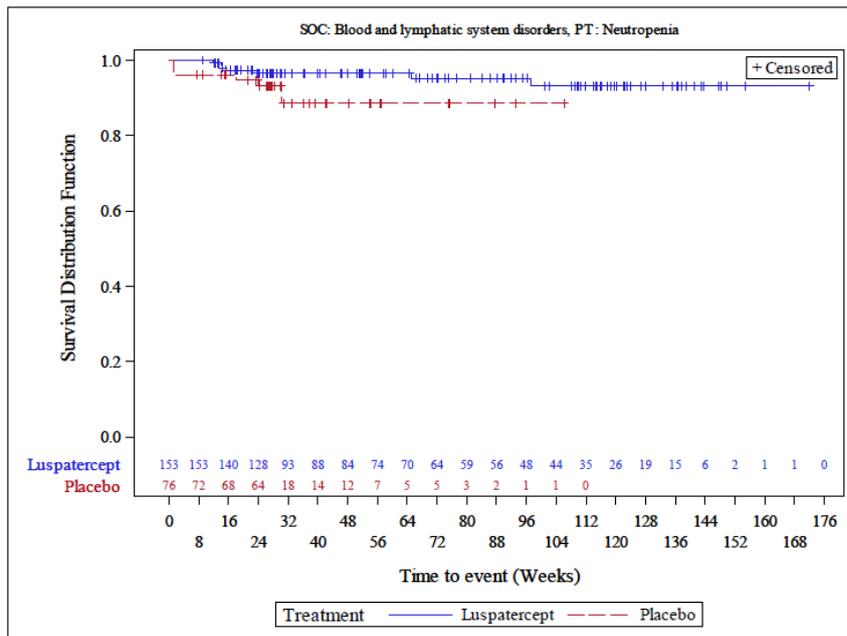
CTCAE = *Common Terminology Criteria for Adverse Events*; SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-127 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE ≥ 3) der SOC Erkrankungen des Blutes und des Lymphsystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



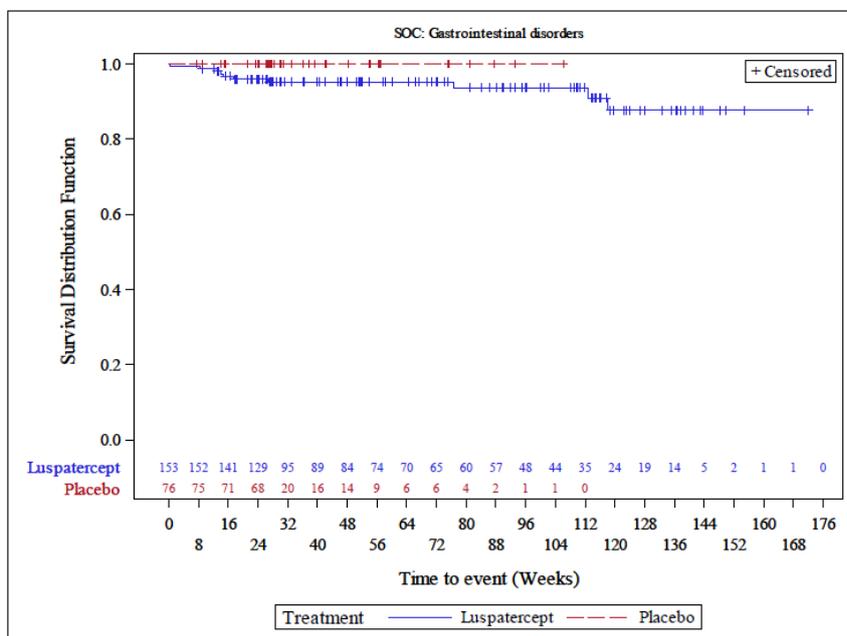
CTCAE = Common Terminology Criteria for Adverse Events; PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-128 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) des PT Anämie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



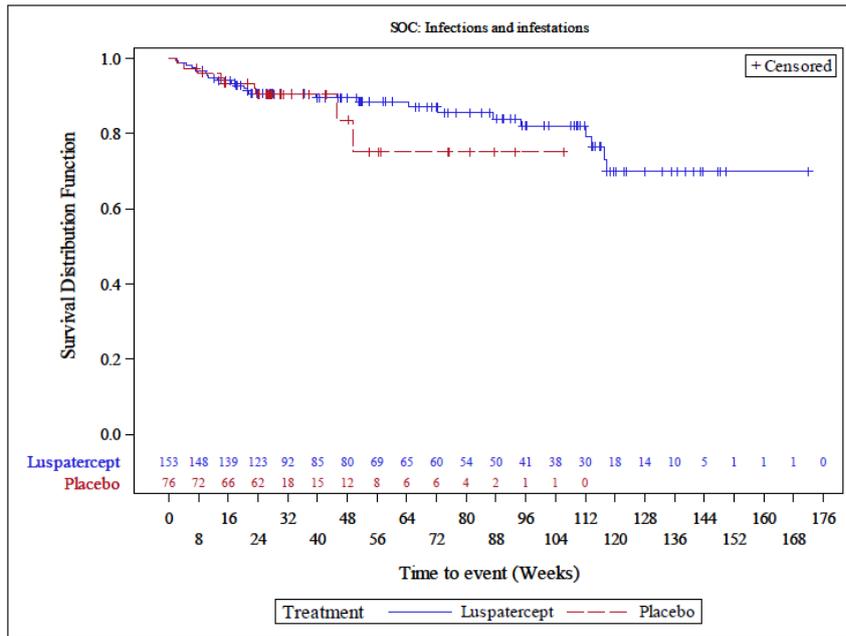
CTCAE = Common Terminology Criteria for Adverse Events; PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-129 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) des PT Neutropenie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



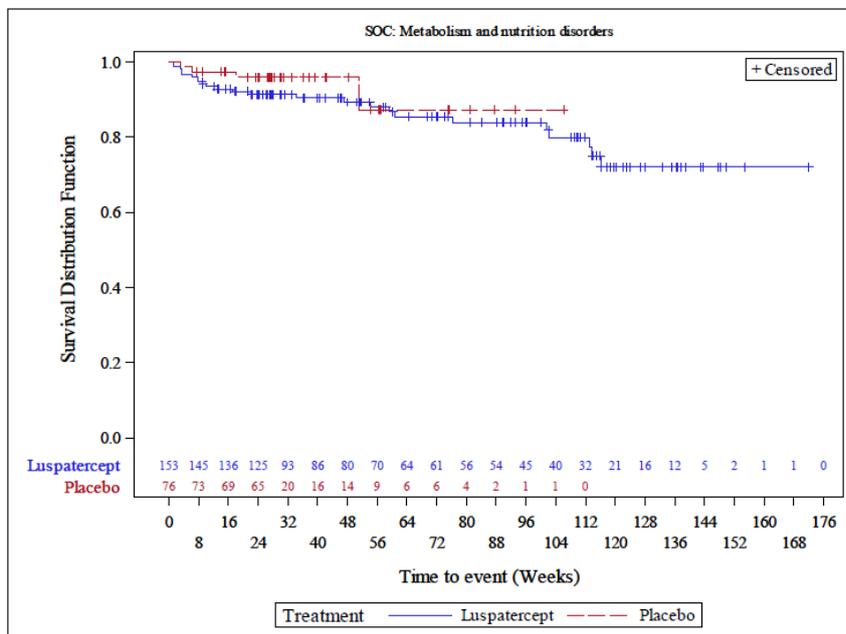
CTCAE = Common Terminology Criteria for Adverse Events; SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-130 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) der SOC Gastrointestinale Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



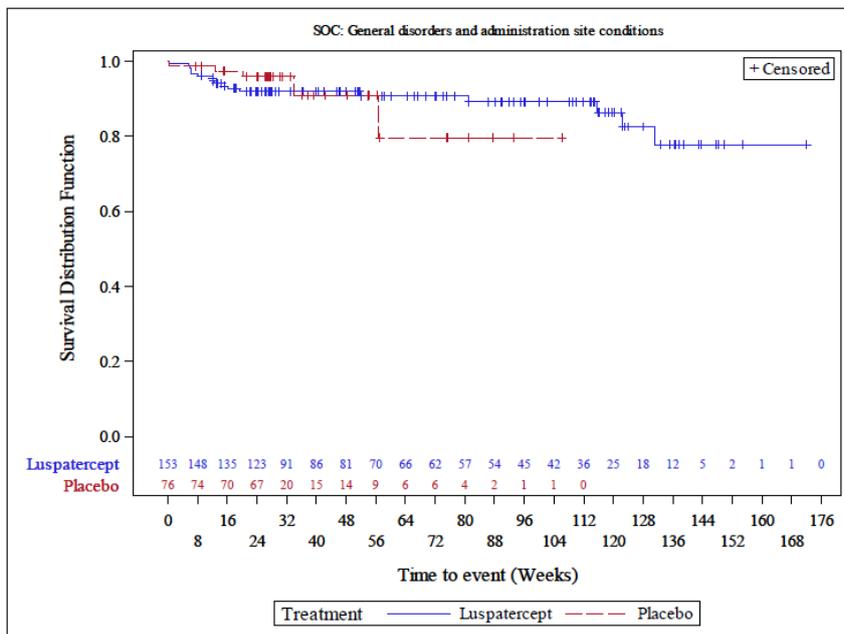
CTCAE = Common Terminology Criteria for Adverse Events; SOC = Systemorganklasse (system organ class); UE = unerwünschtes Ereignis

Abbildung 4-131 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) der SOC Infektionen und parasitäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



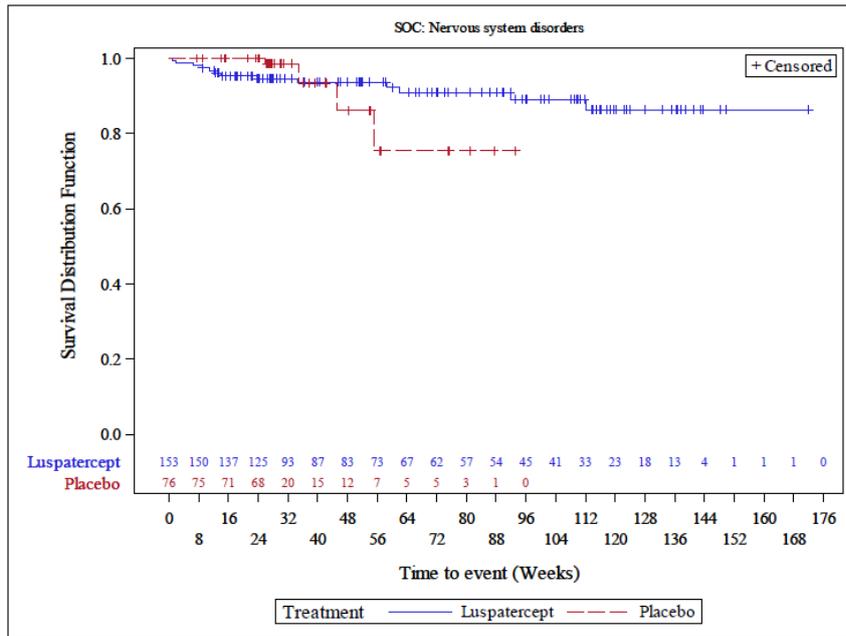
CTCAE = Common Terminology Criteria for Adverse Events; SOC = Systemorganklasse (system organ class); UE = unerwünschtes Ereignis

Abbildung 4-132 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) der SOC Stoffwechsel- und Ernährungsstörungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



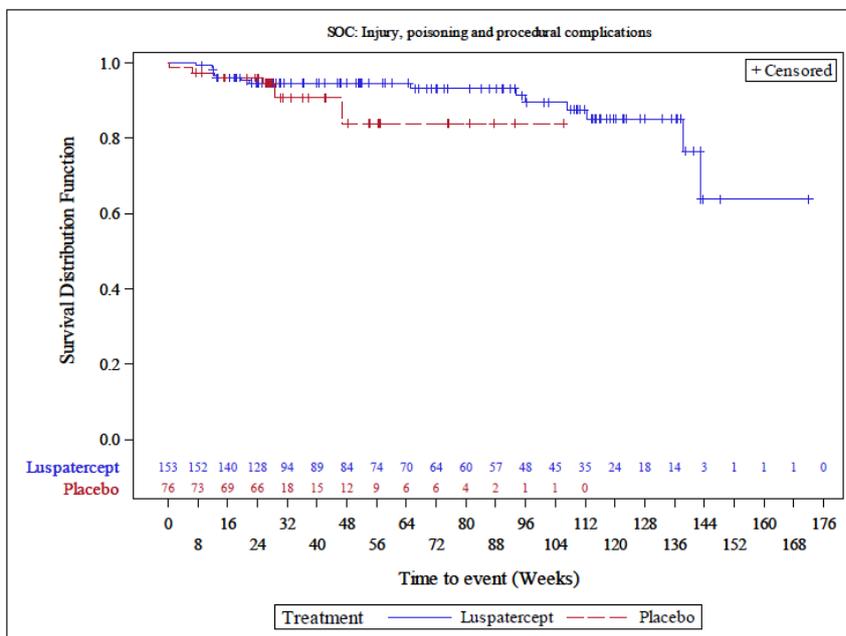
CTCAE = *Common Terminology Criteria for Adverse Events*; SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-133 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) der SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



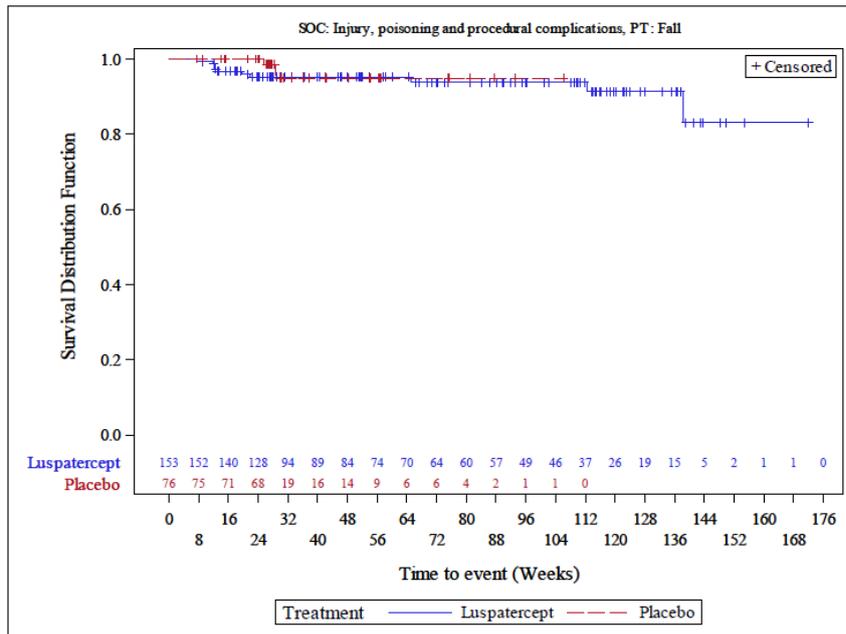
CTCAE = Common Terminology Criteria for Adverse Events; SOC = Systemorganklasse (system organ class); UE = unerwünschtes Ereignis

Abbildung 4-134 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) der SOC Erkrankungen des Nervensystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



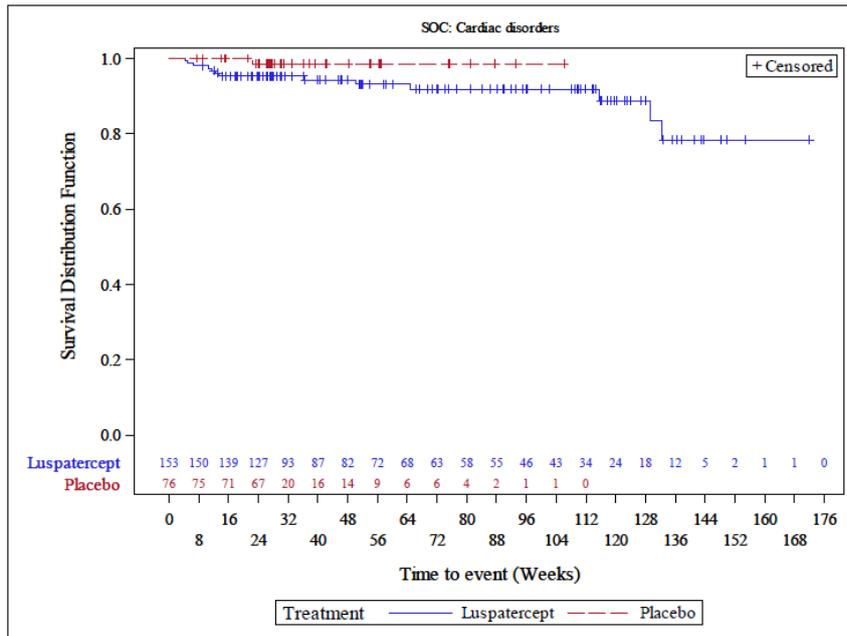
CTCAE = Common Terminology Criteria for Adverse Events; SOC = Systemorganklasse (system organ class); UE = unerwünschtes Ereignis

Abbildung 4-135 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) der SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



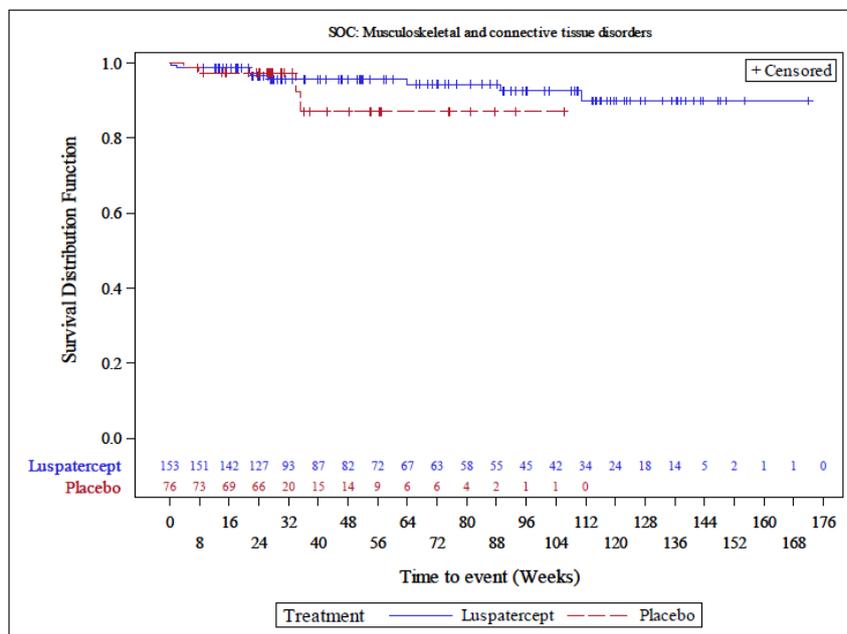
CTCAE = *Common Terminology Criteria for Adverse Events*; PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-136 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq$  3) des PT Sturz aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



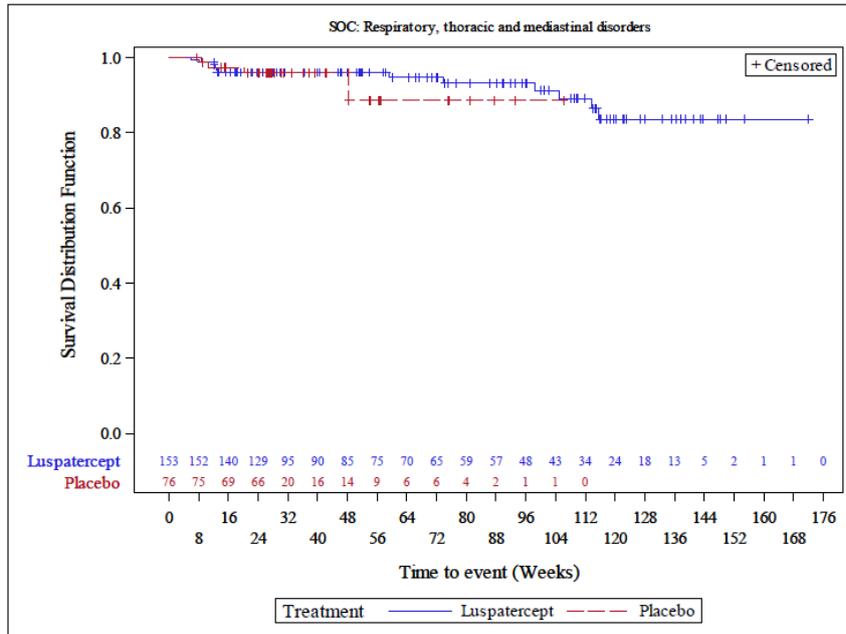
CTCAE = Common Terminology Criteria for Adverse Events; SOC = Systemorganklasse (system organ class); UE = unerwünschtes Ereignis

Abbildung 4-137 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) der SOC Herzerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



CTCAE = Common Terminology Criteria for Adverse Events; SOC = Systemorganklasse (system organ class); UE = unerwünschtes Ereignis

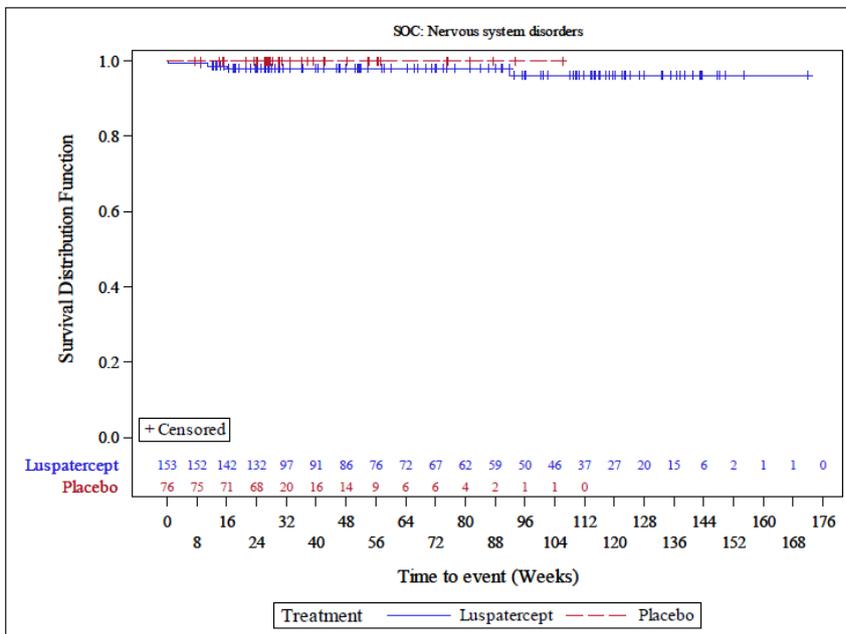
Abbildung 4-138 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten schweren UE (CTCAE  $\geq 3$ ) der SOC Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



SOC = Systemorganklasse (*system organ class*); SUE = schwerwiegendes UE; UE = unerwünschtes Ereignis

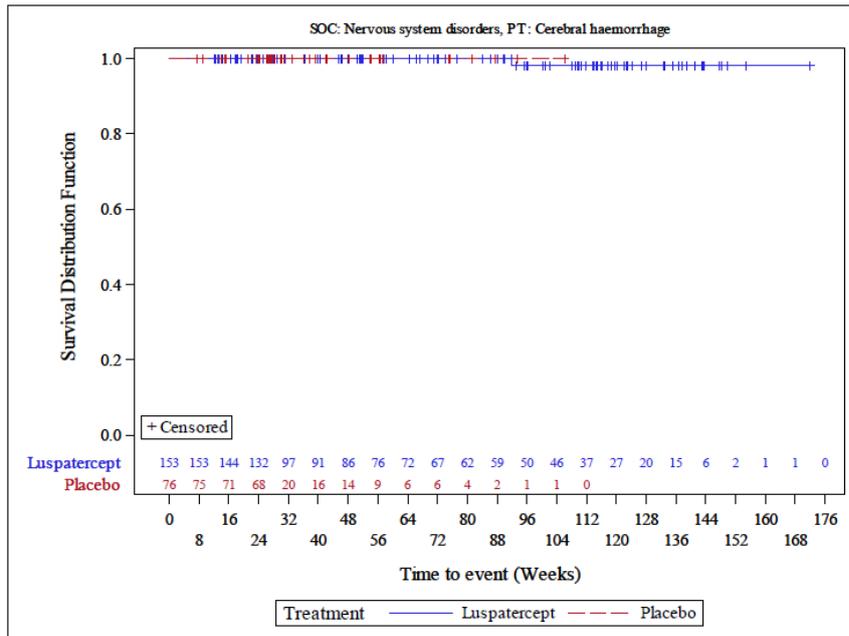
Abbildung 4-139 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum ersten SUE der SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

### Therapieabbrüche aufgrund von UE



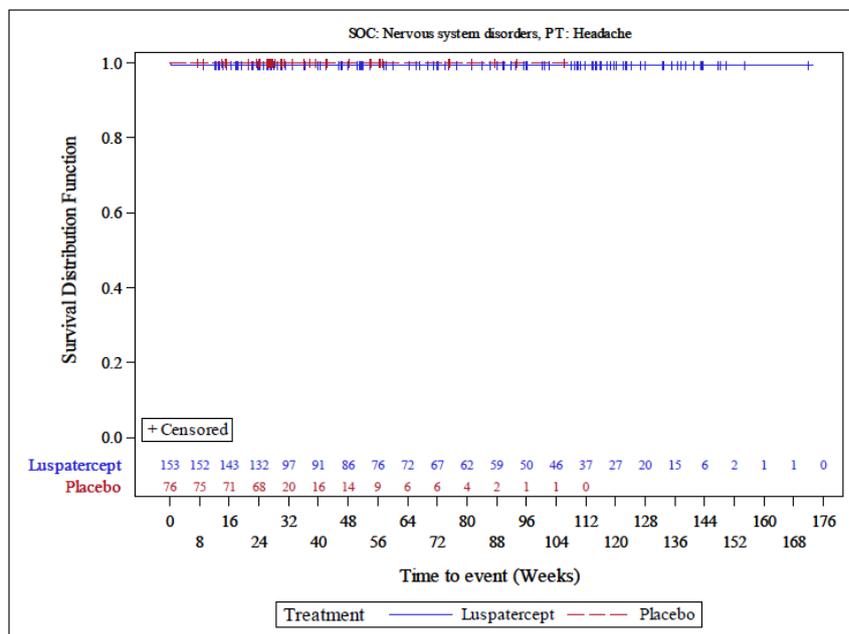
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-140 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Erkrankungen des Nervensystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



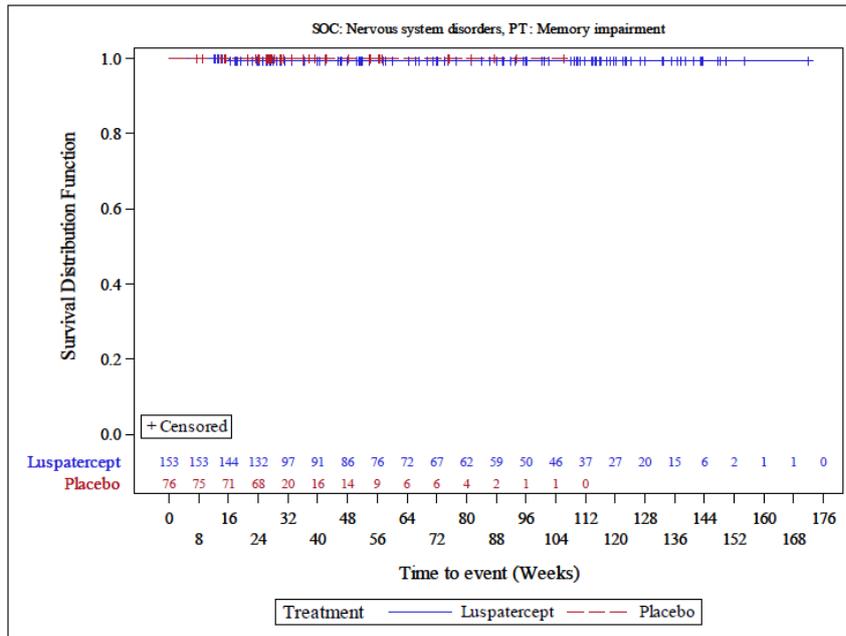
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-141 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Hirnblutung aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



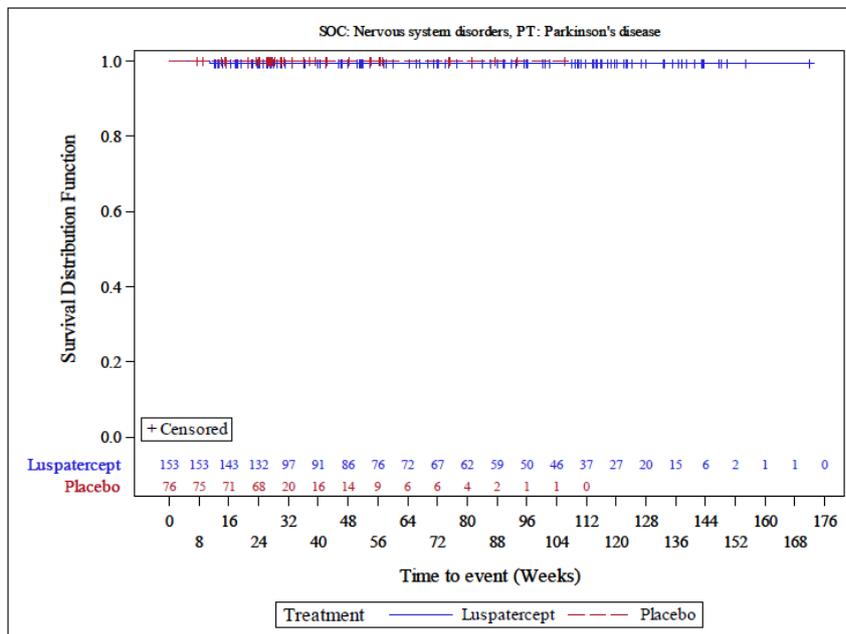
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-142 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Kopfschmerzen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



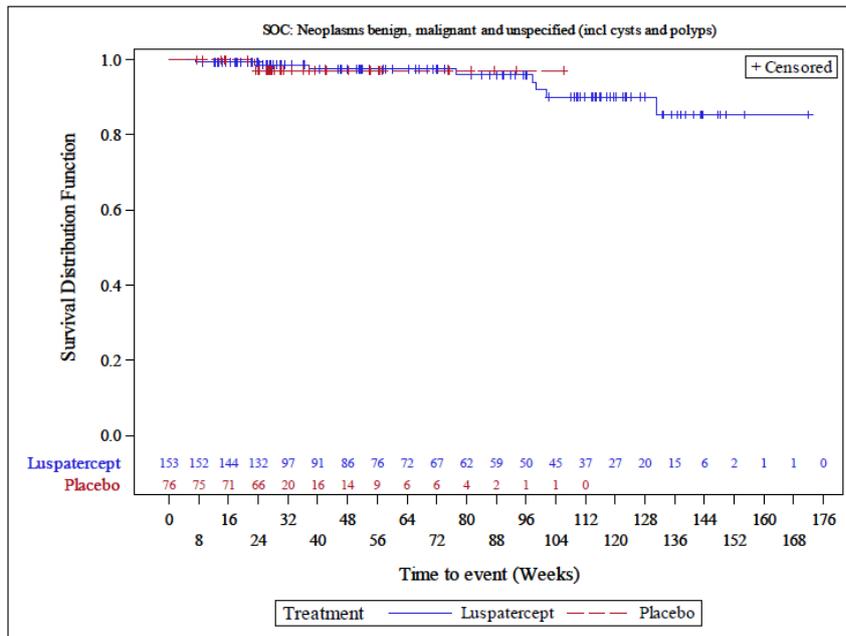
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-143 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Erinnerungsvermögen eingeschränkt aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



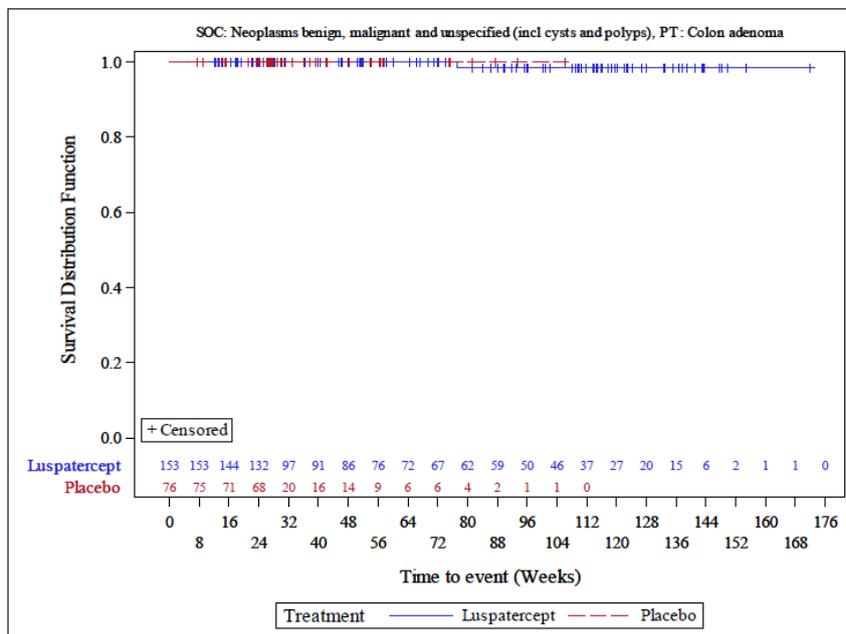
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-144 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Morbus Parkinson aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



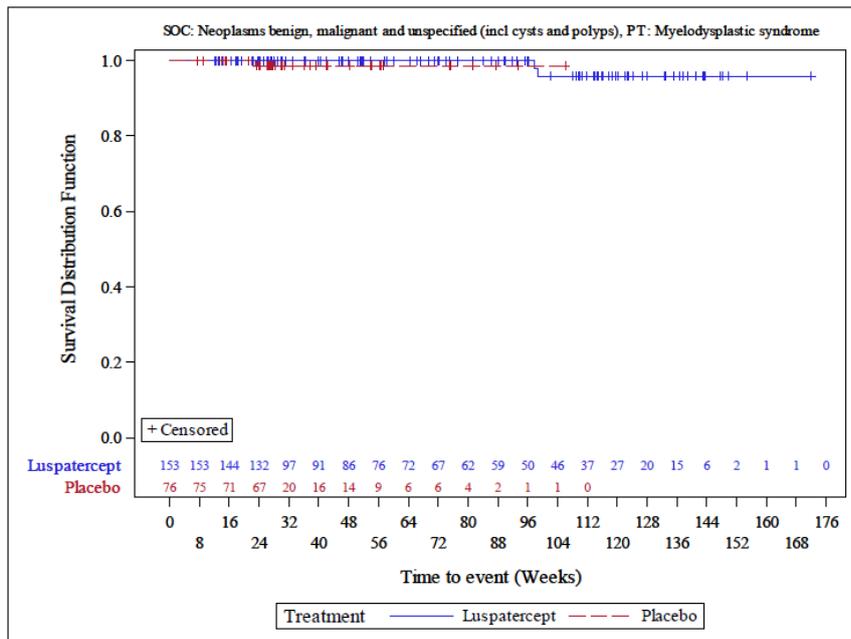
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-145 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Gutartige, bösartige und nicht spezifizierte Neubildungen (einschl. Zysten und Polypen) aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



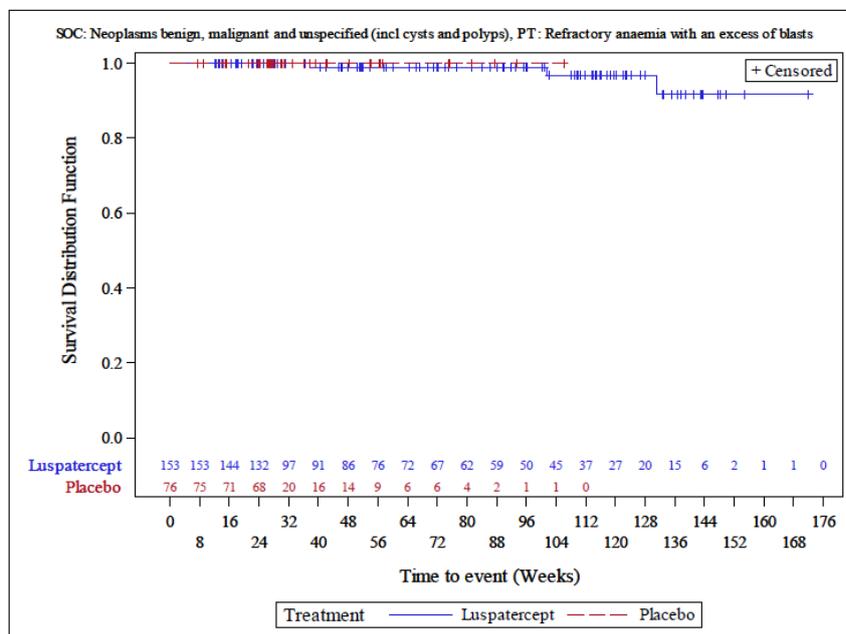
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-146 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Kolonadenom aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



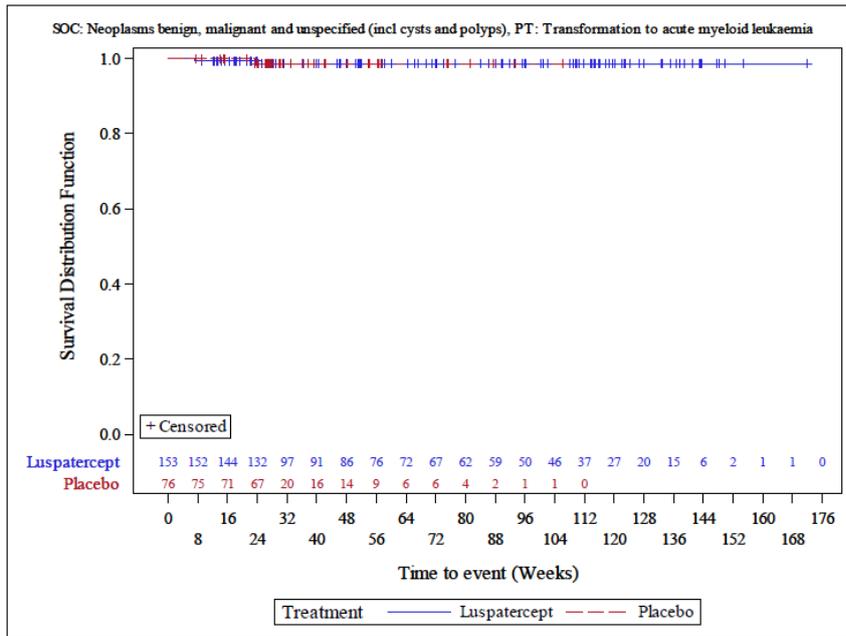
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-147 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Myelodysplastische Syndrome aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



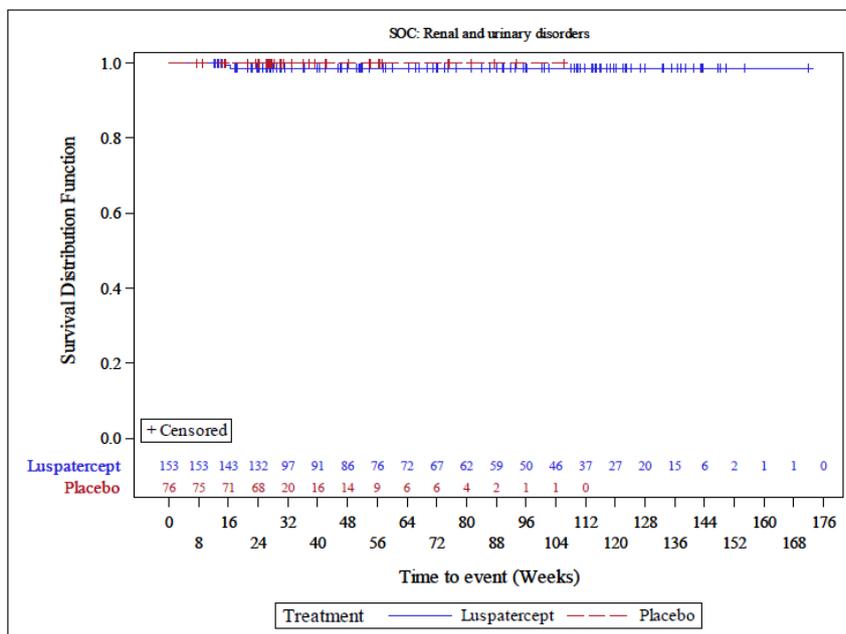
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-148 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Refraktäre Anämie mit Blastenüberschuss aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



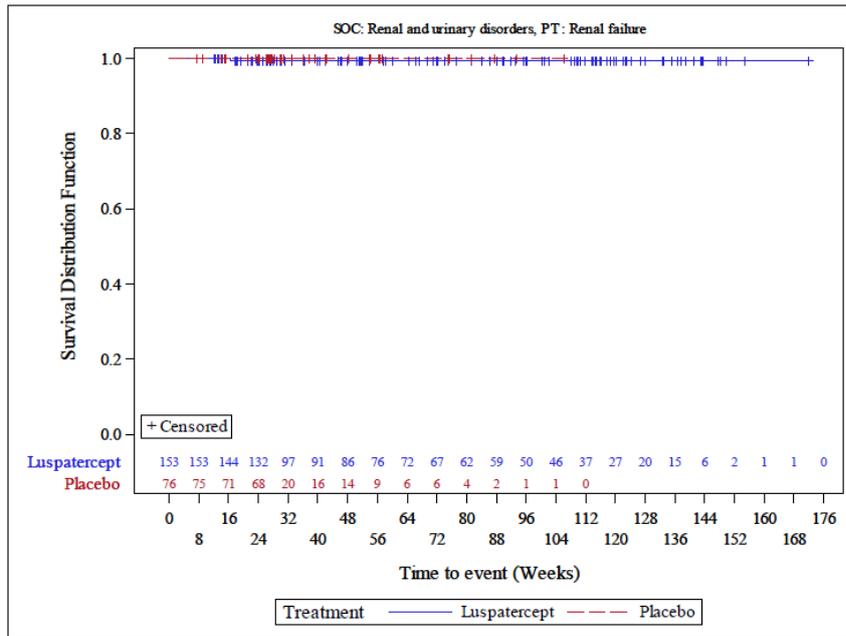
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-149 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Transformation zu einer AML aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



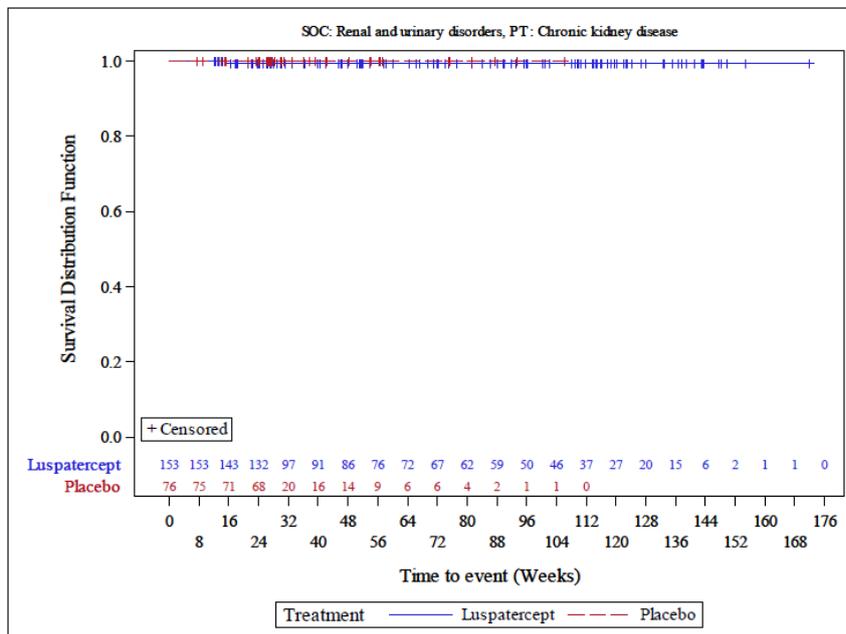
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-150 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Erkrankungen der Nieren und Harnwege aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



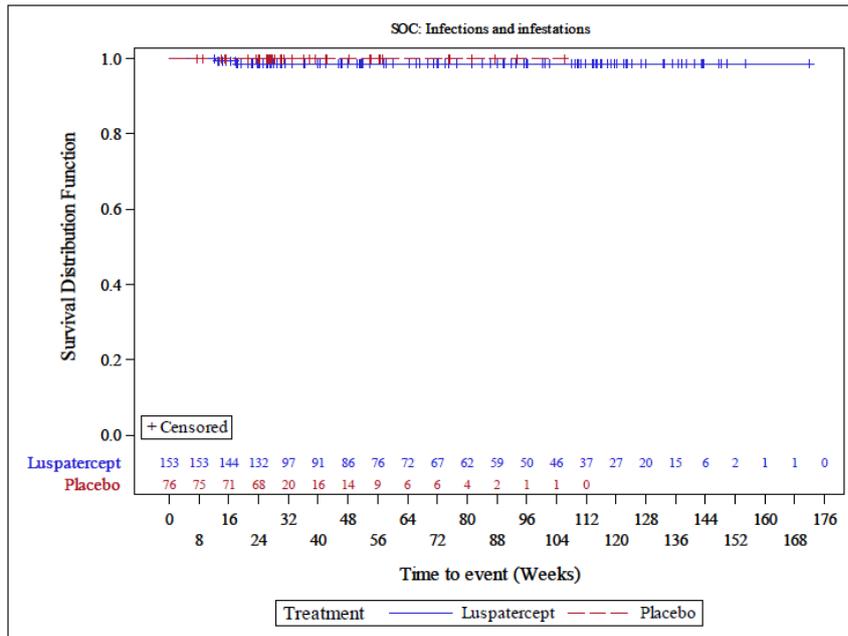
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-151 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Nierenversagen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



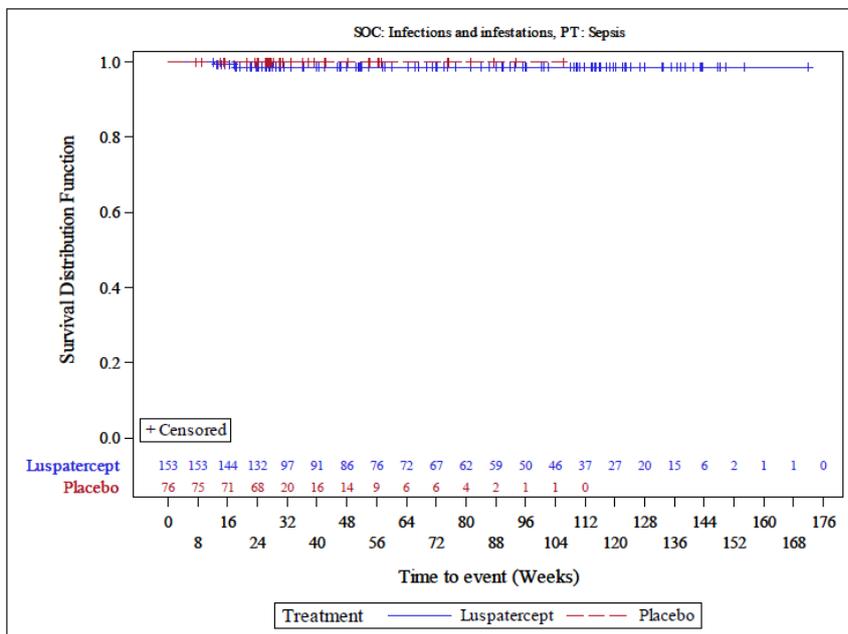
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-152 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT chronische Nierenerkrankung aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



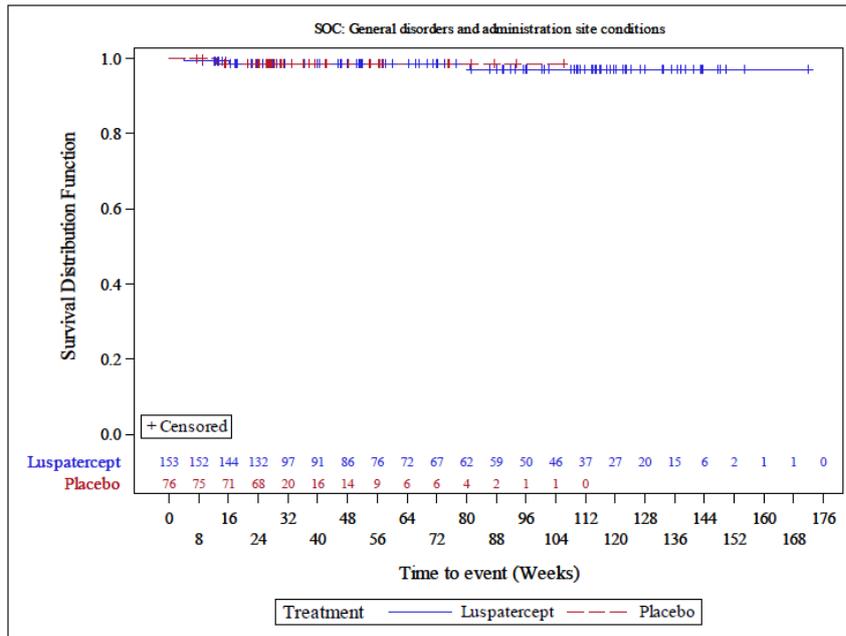
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-153 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Infektionen und parasitäre Erkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



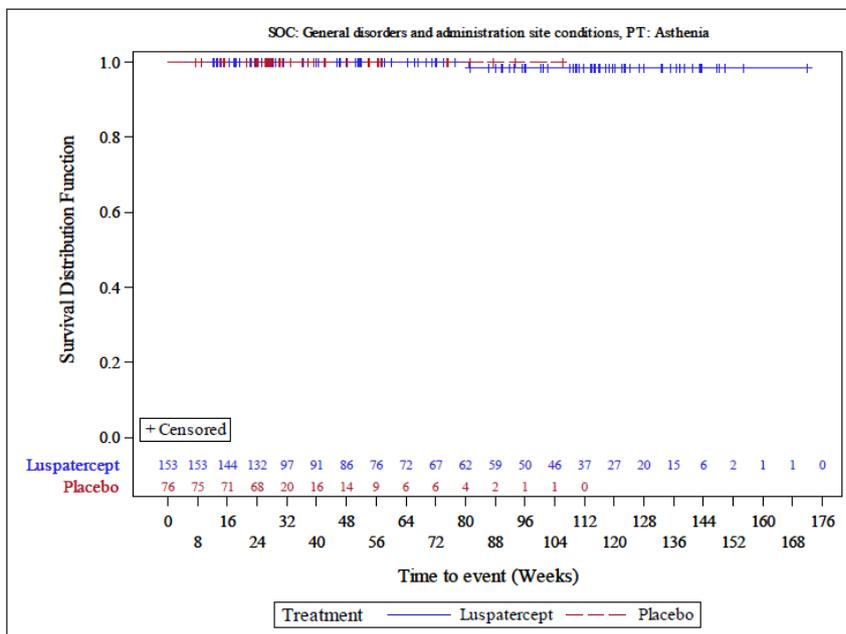
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-154 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Sepsis aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



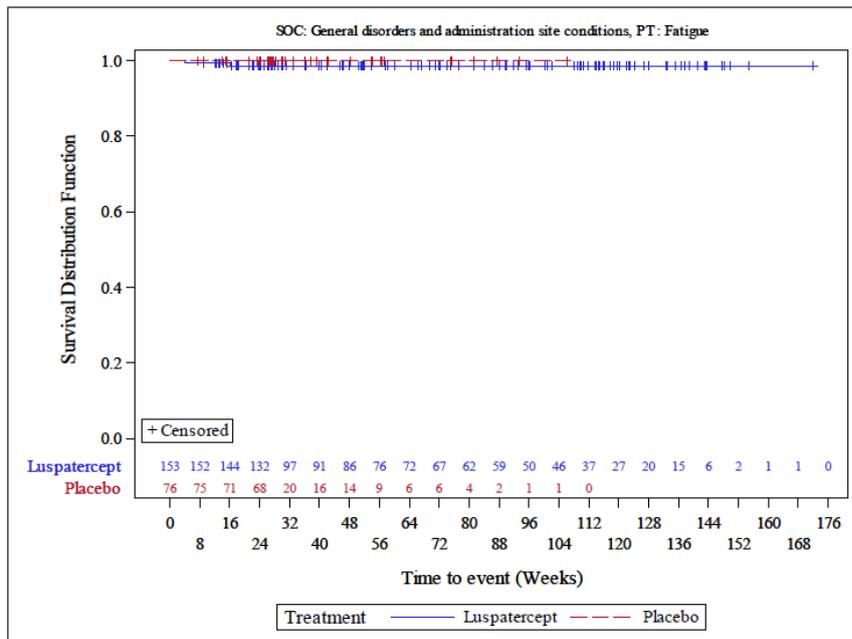
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-155 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Allgemeine Erkrankungen und Beschwerden am Verabreichungsort aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



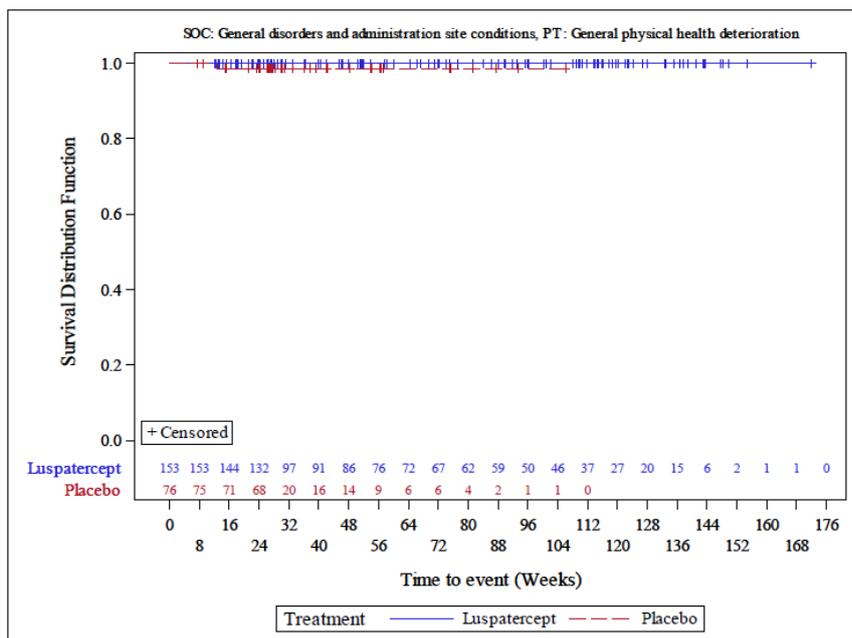
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-156 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Asthenie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



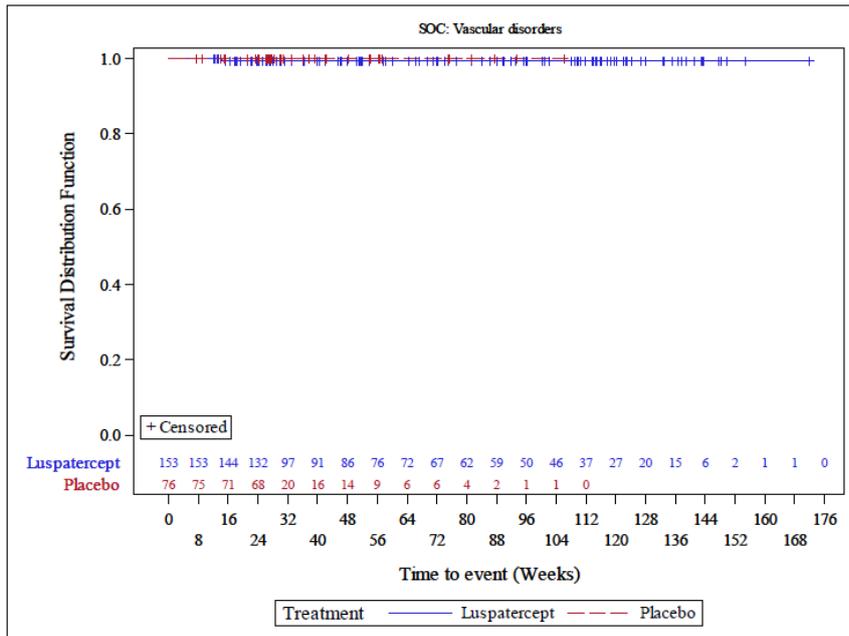
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-157 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Fatigue aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



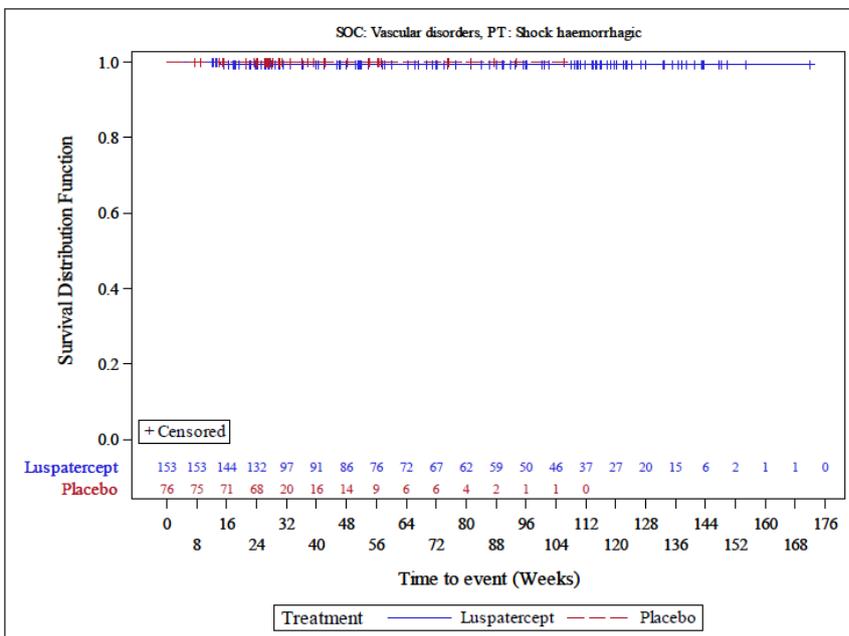
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-158 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Generelle Verschlechterung des physischen Gesundheitsstatus aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



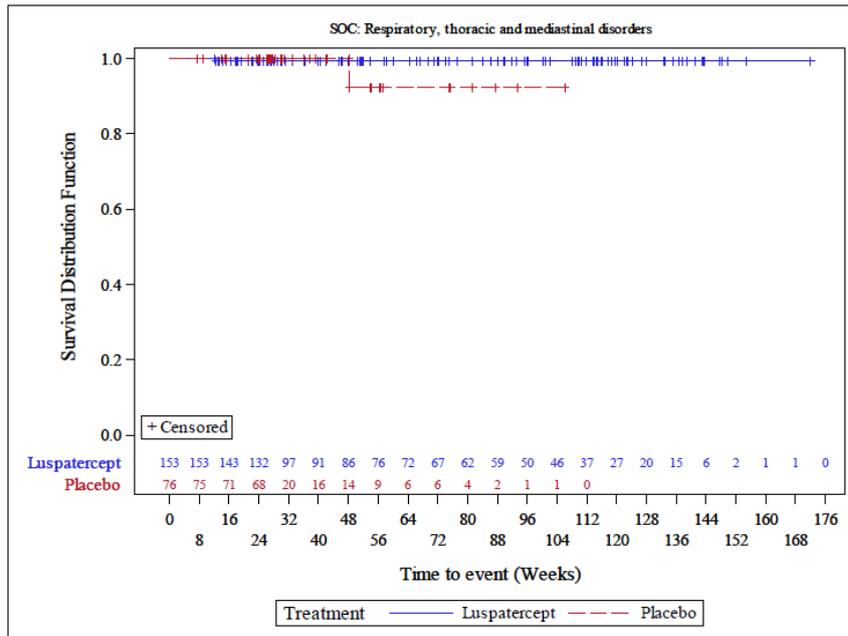
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-159 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Gefäßerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



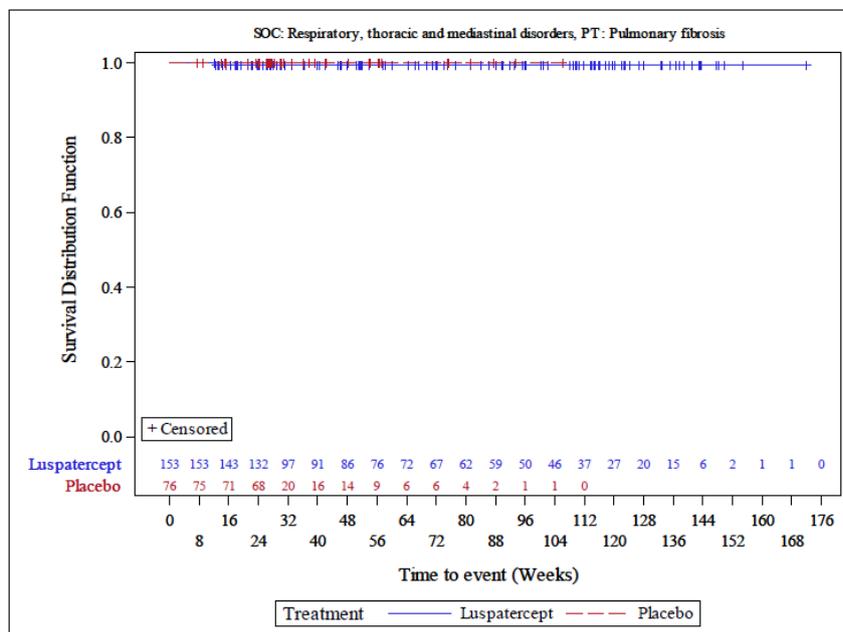
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-160 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Hämorrhagischer Schock aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



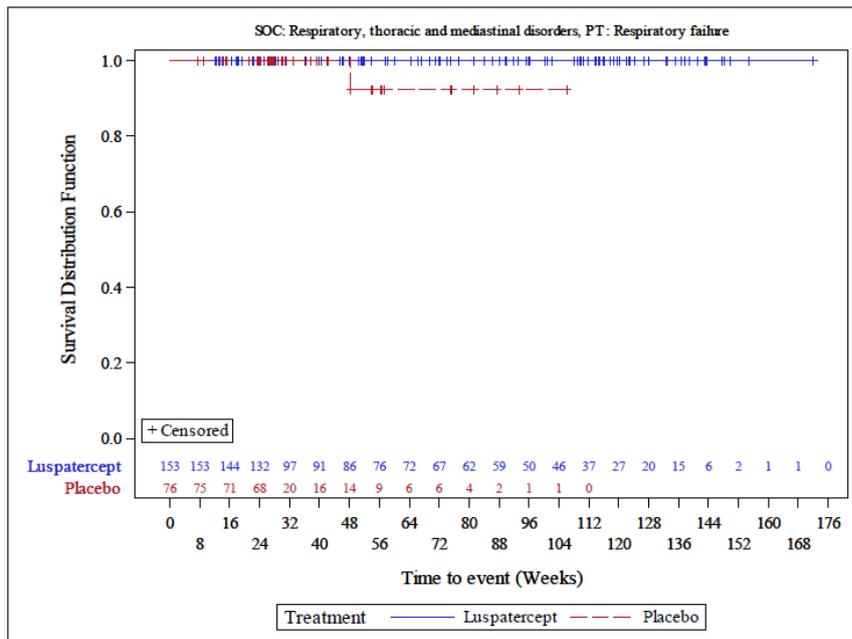
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-161 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Erkrankungen der Atemwege, des Brustraums und Mediastinums aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



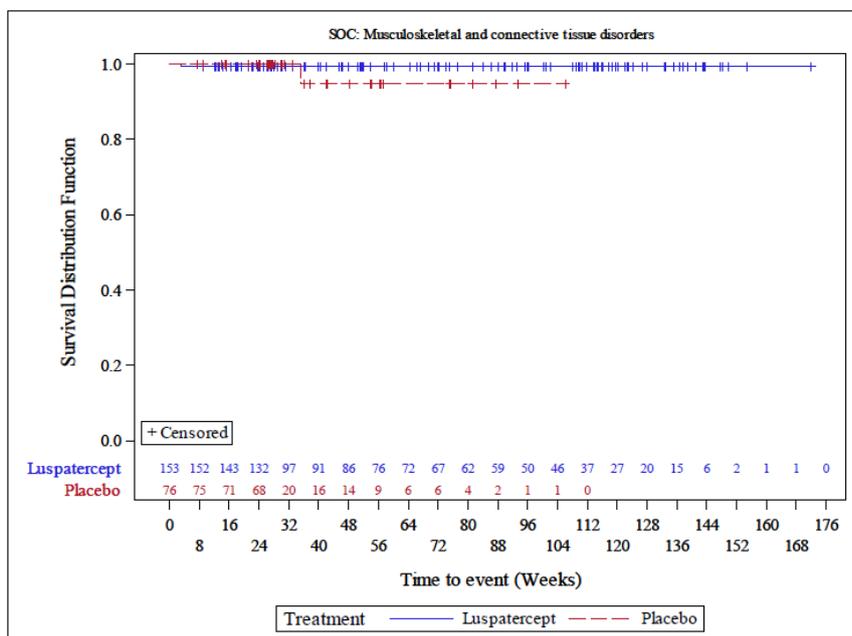
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-162 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Lungenfibrose aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



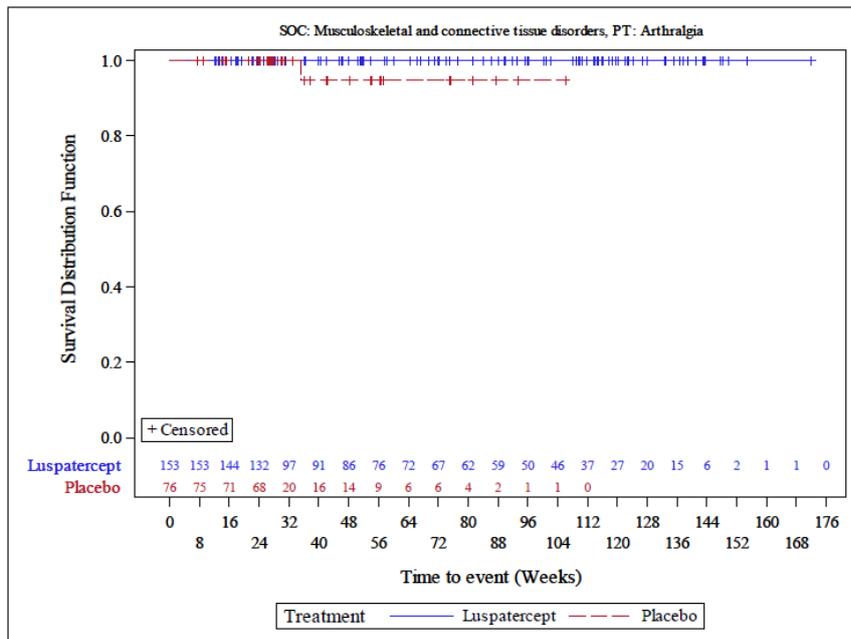
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-163 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Respiratorische Insuffizienz aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



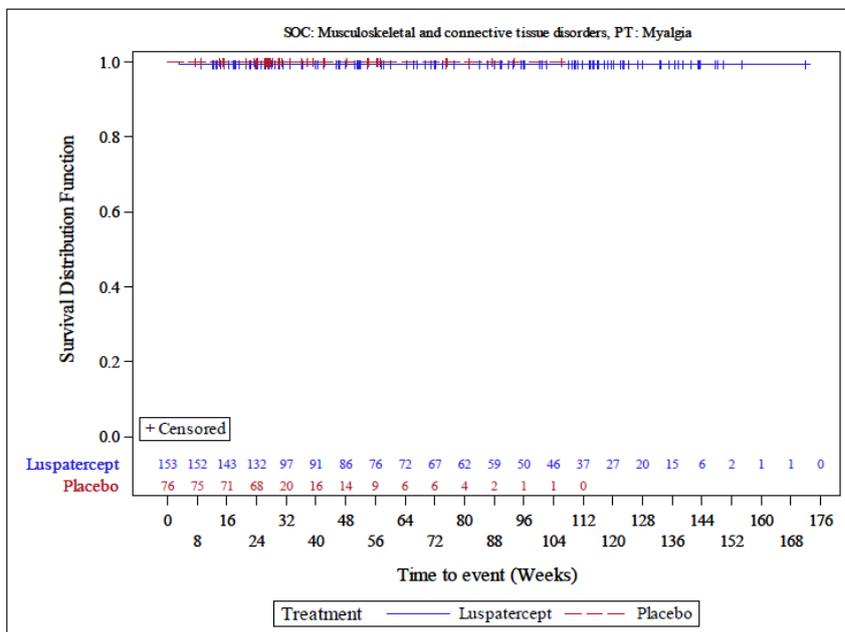
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-164 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



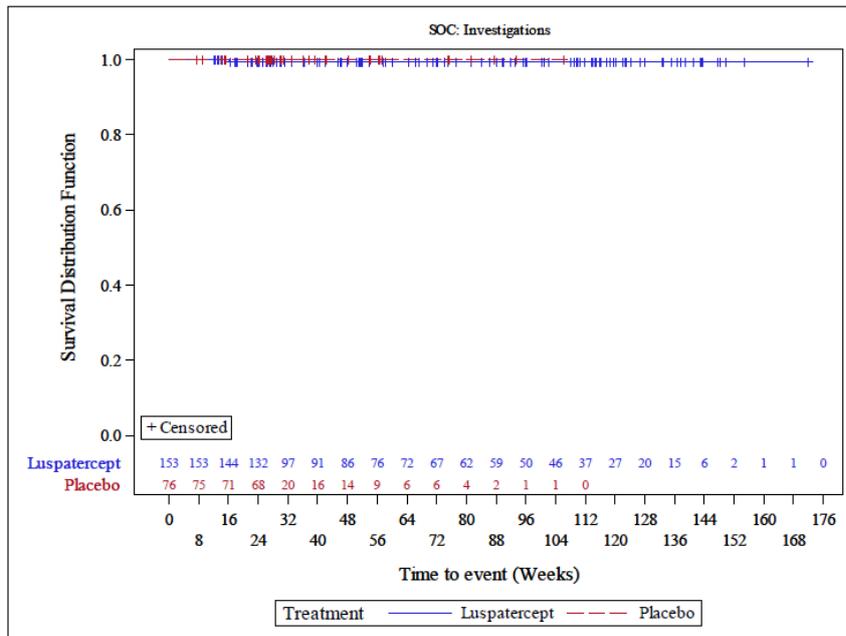
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-165 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Arthralgie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



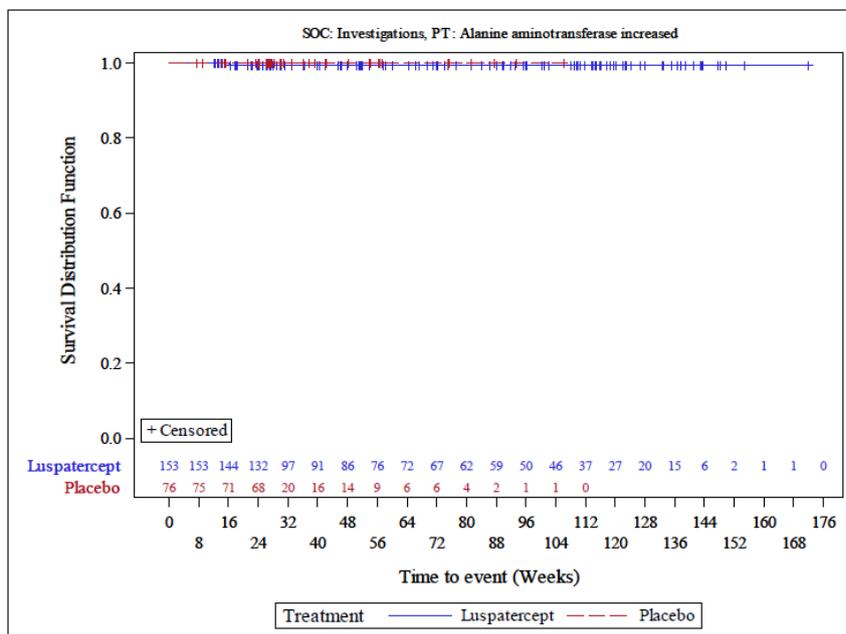
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-166 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Myalgie aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



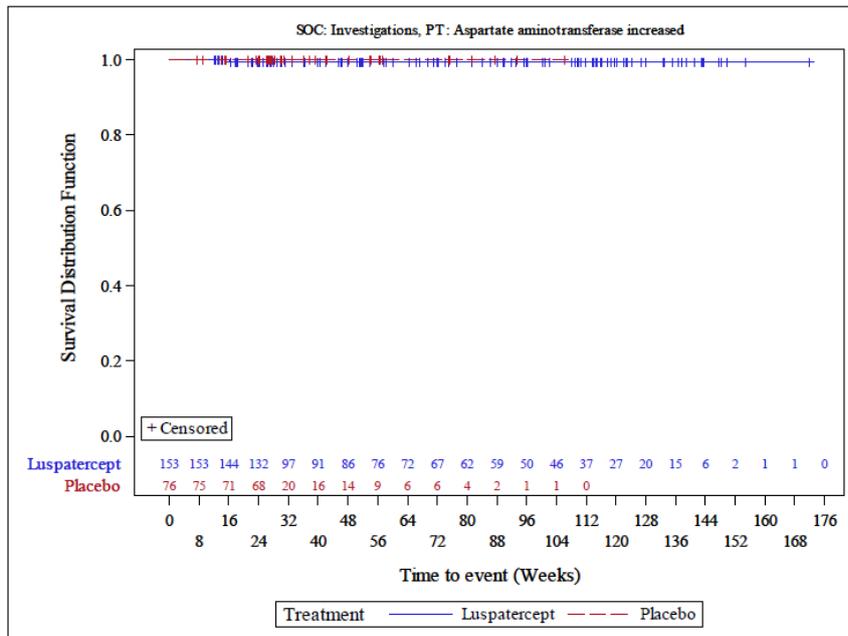
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-167 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Untersuchungen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



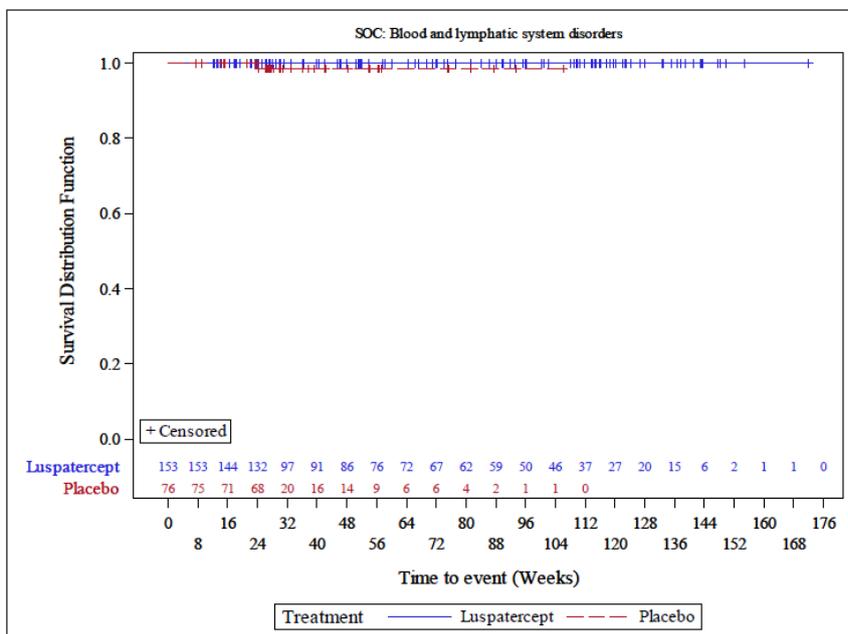
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-168 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Alaninaminotransferase erhöht aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



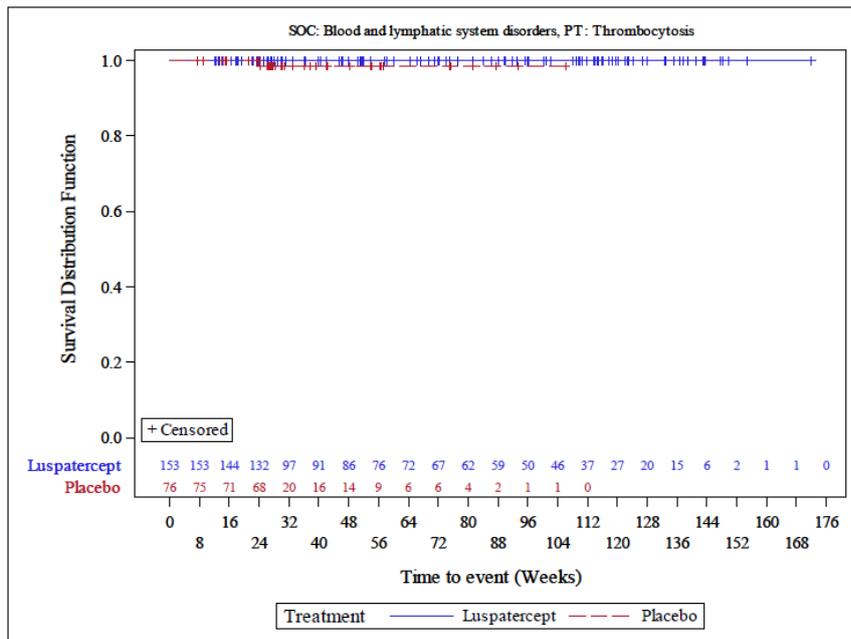
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-169 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Aspartataminotransferase erhöht aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



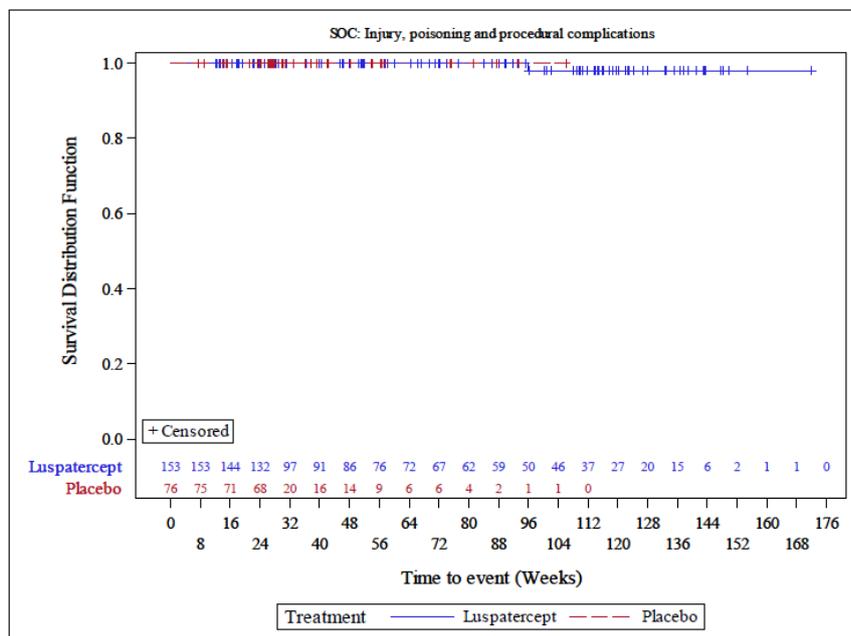
SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-170 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Erkrankungen des Blutes und des Lymphsystems aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



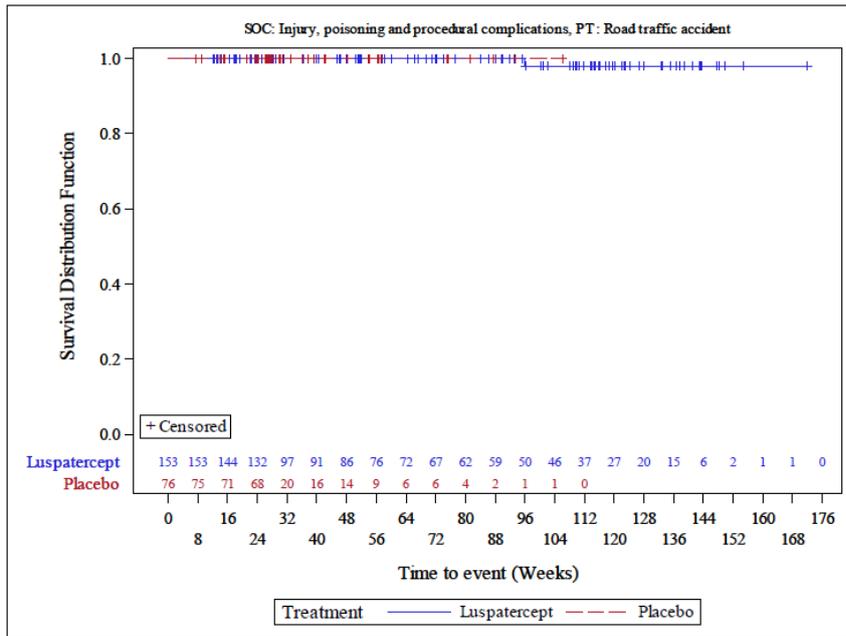
PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-171 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Thrombozytose aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



SOC = Systemorganklasse (*system organ class*); UE = unerwünschtes Ereignis

Abbildung 4-172 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE der SOC Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019



PT = bevorzugter Begriff (*preferred term*); UE = unerwünschtes Ereignis

Abbildung 4-173 (Anhang): Kaplan-Meier-Kurve der Zeit bis zum Therapieabbruch aufgrund von UE des PT Verkehrsunfall aus MEDALIST; Woche eins bis zum Datenschnitt vom 01.07.2019

## **Anhang 4-H: Ergänzende Subgruppenanalysen für statistisch nicht signifikante Interaktionstests**

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Overall Survival - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=153)			Placebo (N=76)			Hazard Ratio (95% CI)	p-Value	p-Value for Interaction
	n/ N (%)	Median	(95% CI)	n/ N (%)	Median	(95% CI)			
Age									0.584
<=64 years	0/ 29 ( 0.0)	NE ( NE , NE )		1/ 16 ( 6.3)	NE (67.6, NE )		NE		
65-74 years	7/ 72 ( 9.7)	NE ( NE , NE )		2/ 29 ( 6.9)	NE ( NE , NE )		1.42 ( 0.30, 6.85)	0.658	
>=75 years	5/ 52 ( 9.6)	NE ( NE , NE )		6/ 31 ( 19.4)	NE (61.6, NE )		0.52 ( 0.16, 1.69)	0.266	
Gender									0.544
female	4/ 59 ( 6.8)	NE ( NE , NE )		4/ 26 ( 15.4)	NE (67.6, NE )		0.48 ( 0.12, 1.93)	0.293	
male	8/ 94 ( 8.5)	NE ( NE , NE )		5/ 50 ( 10.0)	NE ( NE , NE )		0.82 ( 0.27, 2.52)	0.735	
Race									0.962
Black or African American	0/ 1 ( 0.0)	NE ( NE , NE )		0	NE ( NE , NE )		NE	NE	
White	9/107 ( 8.4)	NE ( NE , NE )		6/ 51 ( 11.8)	NE ( NE , NE )		0.72 ( 0.26, 2.03)	0.535	
Not Collected or Reported	3/ 44 ( 6.8)	NE ( NE , NE )		3/ 24 ( 12.5)	NE (67.6, NE )		0.55 ( 0.11, 2.70)	0.451	
Other	0/ 1 ( 0.0)	NE ( NE , NE )		0/ 1 ( 0.0)	NE ( NE , NE )		NE	NE	
Region									0.989
North America	0/ 31 ( 0.0)	NE ( NE , NE )		2/ 19 ( 10.5)	NE (60.7, NE )		NE	NE	
EMEA	12/122 ( 9.8)	NE ( NE , NE )		7/ 57 ( 12.3)	NE ( NE , NE )		0.81 ( 0.32, 2.07)	0.667	
MDS WHO classification									0.993
MDS RARS	1/ 7 ( 14.3)	NE (16.7, NE )		0/ 2 ( 0.0)	NE ( NE , NE )		NE	NE	
MDS RCMD	11/145 ( 7.6)	NE ( NE , NE )		9/ 74 ( 12.2)	NE ( NE , NE )		0.64 ( 0.26, 1.53)	0.308	
Other	0/ 1 ( 0.0)	NE ( NE , NE )		0	NE ( NE , NE )		NE	NE	
Time since initial diagnosis									0.997
<= 2 years	2/ 40 ( 5.0)	NE ( NE , NE )		0/ 19 ( 0.0)	NE ( NE , NE )		NE	NE	
>2 to 5 years	4/ 62 ( 6.5)	NE ( NE , NE )		4/ 34 ( 11.8)	NE ( NE , NE )		0.60 ( 0.15, 2.39)	0.461	
> 5 years	6/ 51 ( 11.8)	NE ( NE , NE )		5/ 23 ( 21.7)	NE (67.6, NE )		0.54 ( 0.17, 1.79)	0.309	
ECOG Performance Status									0.790
0	7/ 54 ( 13.0)	NE ( NE , NE )		2/ 33 ( 6.1)	NE ( NE , NE )		2.05 ( 0.43, 9.90)	0.360	
1	5/ 91 ( 5.5)	NE ( NE , NE )		2/ 32 ( 6.3)	NE ( NE , NE )		0.96 ( 0.19, 4.98)	0.965	
2	0/ 8 ( 0.0)	NE ( NE , NE )		5/ 11 ( 45.5)	60.7 (48.3, NE )		NE	NE	
RBC Transfusions/8 weeks over 16 weeks									0.229
< 6 units	3/ 87 ( 3.4)	NE ( NE , NE )		4/ 43 ( 9.3)	NE ( NE , NE )		0.33 ( 0.07, 1.49)	0.130	
>= 6 units	9/ 66 ( 13.6)	NE ( NE , NE )		5/ 33 ( 15.2)	NE ( NE , NE )		1.03 ( 0.35, 3.09)	0.953	
RBC Transfusions/8 weeks over 16 weeks II									0.815
< 4 units	0/ 46 ( 0.0)	NE ( NE , NE )		1/ 20 ( 5.0)	NE (60.7, NE )		NE	NE	
>= 4 and < 6 units	3/ 41 ( 7.3)	NE ( NE , NE )		3/ 23 ( 13.0)	NE (67.6, NE )		0.54 ( 0.11, 2.70)	0.450	
>= 6 units	9/ 66 ( 13.6)	NE ( NE , NE )		5/ 33 ( 15.2)	NE ( NE , NE )		1.03 ( 0.35, 3.09)	0.953	
IPSS-R risk									0.457
Very Low	2/ 18 ( 11.1)	NE ( NE , NE )		2/ 6 ( 33.3)	NE (23.6, NE )		0.35 ( 0.05, 2.47)	0.268	
Low	8/109 ( 7.3)	NE ( NE , NE )		4/ 57 ( 7.0)	NE ( NE , NE )		1.05 ( 0.32, 3.50)	0.932	
Intermediate	2/ 25 ( 8.0)	NE ( NE , NE )		3/ 13 ( 23.1)	NE (60.7, NE )		0.48 ( 0.08, 2.90)	0.415	
High	0/ 1 ( 0.0)	NE ( NE , NE )		0	NE ( NE , NE )		NE	NE	
Platelet count									0.947
< 100 x 10 <sup>9</sup> /L	2/ 8 ( 25.0)	NE (46.0, NE )		1/ 6 ( 16.7)	NE (48.3, NE )		1.51 ( 0.14, 16.70)	0.735	
100-400 x 10 <sup>9</sup> /L	10/128 ( 7.8)	NE ( NE , NE )		5/ 61 ( 8.2)	NE ( NE , NE )		1.01 ( 0.34, 2.94)	0.992	
> 400 x 10 <sup>9</sup> /L	0/ 17 ( 0.0)	NE ( NE , NE )		3/ 9 ( 33.3)	67.6 (14.1, NE )		NE	NE	
Renal status: Creatinine Clearance									0.998
40 - <60 mL/min	3/ 33 ( 9.1)	NE ( NE , NE )		2/ 12 ( 16.7)	NE (57.7, NE )		0.60 ( 0.10, 3.62)	0.575	
>= 60 mL/min	9/113 ( 8.0)	NE ( NE , NE )		7/ 62 ( 11.3)	NE ( NE , NE )		0.70 ( 0.26, 1.87)	0.472	
< 40 mL/min	0/ 7 ( 0.0)	NE ( NE , NE )		0/ 1 ( 0.0)	NE ( NE , NE )		NE	NE	

Confidence intervals (CI) for Kaplan-Meier median derived using the log-log transformation  
 Hazard Ratio (luspatercept over placebo) with 95% CI are estimated from Cox Proportional Hazard Model.  
 p-value from 2-sided log-rank test.  
 p-value for interaction from Cox Proportional Hazard Model with treatment, subgroup and treatment by subgroup interaction as covariates.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Overall Survival - Subgroup analysis

ITT Population

Subgroup Level	Luspatercept (N=153)			Placebo (N=76)			Hazard Ratio (95% CI)	p-Value	p-Value for Interaction
	n/ N (%)	Median (95% CI)		n/ N (%)	Median (95% CI)				
Serum EPO									0.943
< 100 U/L	3/ 51 ( 5.9)	NE ( NE , NE )		4/ 31 ( 12.9)	NE ( NE , NE )		0.41 ( 0.09, 1.85)	0.233	
100- < 200 U/L	5/ 37 ( 13.5)	NE ( NE , NE )		3/ 19 ( 15.8)	NE (60.7, NE )		0.82 ( 0.20, 3.42)	0.781	
200-500 U/L	2/ 43 ( 4.7)	NE ( NE , NE )		2/ 15 ( 13.3)	NE ( NE , NE )		0.42 ( 0.06, 3.05)	0.379	
> 500 U/L	2/ 21 ( 9.5)	NE ( NE , NE )		0/ 11 ( 0.0)	NE ( NE , NE )		NE	NE	
SF3B1 mutation status									0.174
Mutated	9/141 ( 6.4)	NE ( NE , NE )		8/ 65 ( 12.3)	NE ( NE , NE )		0.52 ( 0.20, 1.36)	0.177	
Non-mutated	3/ 12 ( 25.0)	NE (46.0, NE )		1/ 10 ( 10.0)	NE (43.3, NE )		2.89 ( 0.30, 27.92)	0.337	
Longest duration of prior ESA treatment									0.298
< 6 months	7/ 48 ( 14.6)	NE ( NE , NE )		1/ 20 ( 5.0)	NE ( NE , NE )		2.81 ( 0.35, 22.83)	0.313	
6-12 months	0/ 34 ( 0.0)	NE ( NE , NE )		1/ 9 ( 11.1)	NE (48.3, NE )		NE	NE	
>12-24 months	2/ 35 ( 5.7)	NE ( NE , NE )		2/ 21 ( 9.5)	NE ( NE , NE )		0.71 ( 0.10, 5.09)	0.734	
> 24 months	1/ 31 ( 3.2)	NE ( NE , NE )		4/ 20 ( 20.0)	NE (60.7, NE )		0.13 ( 0.01, 1.14)	0.030	

Confidence intervals (CI) for Kaplan-Meier median derived using the log-log transformation

Hazard Ratio (luspatercept over placebo) with 95% CI are estimated from Cox Proportional Hazard Model.

p-value from 2-sided log-rank test.

p-value for interaction from Cox Proportional Hazard Model with treatment, subgroup and treatment by subgroup interaction as covariates.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 RBC-TI >=8 weeks during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									
<=64 years	17/ 29 ( 58.6)	3/ 16 ( 18.8)			6.14 ( 1.43, 26.35)	0.011	3.13 ( 1.08, 9.07)	39.87 (13.66, 66.08)	0.837
65-74 years	23/ 72 ( 31.9)	4/ 29 ( 13.8)			2.93 ( 0.91, 9.41)	0.064	2.32 ( 0.88, 6.11)	18.15 ( 1.61, 34.69)	
>=75 years	18/ 52 ( 34.6)	3/ 31 ( 9.7)			4.94 ( 1.32, 18.51)	0.012	3.58 ( 1.15, 11.17)	24.94 ( 8.34, 41.54)	
Gender									
female	26/ 59 ( 44.1)	6/ 26 ( 23.1)			2.63 ( 0.92, 7.48)	0.067	1.91 ( 0.89, 4.08)	20.99 ( 0.43, 41.55)	0.194
male	32/ 94 ( 34.0)	4/ 50 ( 8.0)			5.94 ( 1.96, 17.96)	<.001	4.26 ( 1.60, 11.35)	26.04 (13.86, 38.22)	
Race									
Black or African American	1/ 1 (100.0)	0			NE	NE	NE	NE	0.392
White	40/107 ( 37.4)	8/ 51 ( 15.7)			3.21 ( 1.37, 7.51)	0.006	2.38 ( 1.21, 4.71)	21.70 ( 8.14, 35.25)	
Not Collected or Reported	17/ 44 ( 38.6)	2/ 24 ( 8.3)			6.93 ( 1.44, 33.28)	0.008	4.64 ( 1.17, 18.40)	30.30 (12.16, 48.45)	
Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)			NE	NE	NE	NE	
Region									
North America	17/ 31 ( 54.8)	4/ 19 ( 21.1)			4.55 ( 1.23, 16.88)	0.020	2.60 ( 1.03, 6.59)	33.79 ( 8.43, 59.14)	0.742
EMEA	41/122 ( 33.6)	6/ 57 ( 10.5)			4.30 ( 1.71, 10.86)	0.001	3.19 ( 1.44, 7.08)	23.08 (11.52, 34.64)	
MDS WHO classification									
MDS RARS	3/ 7 ( 42.9)	0/ 2 ( 0.0)			3.89 ( 0.14,109.99)	0.285	2.63 ( 0.19, 37.14)	42.86 ( 6.20, 79.52)	0.972
MDS RCMD	54/145 ( 37.2)	10/ 74 ( 13.5)			3.80 ( 1.80, 8.01)	<.001	2.76 ( 1.49, 5.09)	23.73 (12.66, 34.80)	
Other	1/ 1 (100.0)	0			NE	NE	NE	NE	
Time since initial diagnosis									
<= 2 years	14/ 40 ( 35.0)	3/ 19 ( 15.8)			2.87 ( 0.71, 11.57)	0.131	2.22 ( 0.72, 6.80)	19.21 (-2.86, 41.29)	0.595
>2 to 5 years	30/ 62 ( 48.4)	4/ 34 ( 11.8)			7.03 ( 2.21, 22.34)	<.001	4.11 ( 1.58, 10.70)	36.62 (20.13, 53.12)	
> 5 years	14/ 51 ( 27.5)	3/ 23 ( 13.0)			2.52 ( 0.65, 9.83)	0.176	2.10 ( 0.67, 6.62)	14.41 (-4.02, 32.83)	
ECOG Performance Status									
0	17/ 54 ( 31.5)	5/ 33 ( 15.2)			2.57 ( 0.85, 7.82)	0.091	2.08 ( 0.85, 5.10)	16.33 (-1.08, 33.74)	0.709
1	38/ 91 ( 41.8)	4/ 32 ( 12.5)			5.02 ( 1.63, 15.50)	0.003	3.34 ( 1.29, 8.62)	29.26 (13.96, 44.55)	
2	3/ 8 ( 37.5)	1/ 11 ( 9.1)			6.00 ( 0.49, 73.45)	0.144	4.13 ( 0.52, 32.75)	28.41 (-9.19, 66.01)	
RBC Transfusions/8 weeks over 16 weeks									
< 6 units	52/ 87 ( 59.8)	9/ 43 ( 20.9)			5.61 ( 2.40, 13.14)	<.001	2.86 ( 1.56, 5.23)	38.84 (22.90, 54.78)	0.964
>= 6 units	6/ 66 ( 9.1)	1/ 33 ( 3.0)			3.20 ( 0.37, 27.75)	0.270	3.00 ( 0.38, 23.90)	6.06 (-3.01, 15.13)	
RBC Transfusions/8 weeks over 16 weeks II									
< 4 units	37/ 46 ( 80.4)	8/ 20 ( 40.0)			6.17 ( 1.95, 19.55)	0.001	2.01 ( 1.15, 3.50)	40.43 (16.10, 64.77)	0.303
>= 4 and < 6 units	15/ 41 ( 36.6)	1/ 23 ( 4.3)			12.69 ( 1.55,103.89)	0.005	8.41 ( 1.19, 59.67)	32.24 (15.30, 49.17)	
>= 6 units	6/ 66 ( 9.1)	1/ 33 ( 3.0)			3.20 ( 0.37, 27.75)	0.270	3.00 ( 0.38, 23.90)	6.06 (-3.01, 15.13)	
IPSS-R risk									
Very Low	7/ 18 ( 38.9)	1/ 6 ( 16.7)			3.18 ( 0.30, 33.26)	0.328	2.33 ( 0.36, 15.30)	22.22 (-15.1, 59.59)	0.800
Low	41/109 ( 37.6)	8/ 57 ( 14.0)			3.69 ( 1.59, 8.57)	0.002	2.68 ( 1.35, 5.32)	23.58 (10.77, 36.39)	
Intermediate	10/ 25 ( 40.0)	1/ 13 ( 7.7)			8.00 ( 0.89, 71.58)	0.040	5.20 ( 0.74, 36.31)	32.31 ( 8.25, 56.36)	
High	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Platelet count									
< 100 x 10 <sup>9</sup> /L	2/ 8 ( 25.0)	1/ 6 ( 16.7)			1.67 ( 0.11, 24.26)	0.717	1.50 ( 0.17, 12.94)	8.33 (-34.0, 50.64)	0.470
100-400 x 10 <sup>9</sup> /L	42/128 ( 32.8)	8/ 61 ( 13.1)			3.24 ( 1.41, 7.42)	0.004	2.50 ( 1.25, 5.00)	19.70 ( 7.95, 31.44)	
> 400 x 10 <sup>9</sup> /L	14/ 17 ( 82.4)	1/ 9 ( 11.1)			37.33 ( 3.31,421.59)	<.001	7.41 ( 1.15, 47.66)	71.24 (43.86, 98.63)	
Renal status: Creatinine Clearance									
40 - <60 mL/min	11/ 33 ( 33.3)	1/ 12 ( 8.3)			5.50 ( 0.63, 48.23)	0.097	4.00 ( 0.58, 27.77)	25.00 ( 2.57, 47.43)	0.878
>= 60 mL/min	44/113 ( 38.9)	8/ 62 ( 12.9)			4.30 ( 1.87, 9.90)	<.001	3.02 ( 1.52, 6.00)	26.03 (13.77, 38.30)	
< 40 mL/min	3/ 7 ( 42.9)	0/ 1 ( 0.0)			2.33 ( 0.07, 76.67)	0.439	1.75 ( 0.14, 21.88)	42.86 ( 6.20, 79.52)	
Serum EPO									0.653

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

RBC-TI >=8 weeks during week 1-24 - Subgroup analysis

ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
< 100 U/L	23/ 51 ( 45.1)	7/ 31 ( 22.6)			2.82 ( 1.03, 7.71)	0.041	2.00 ( 0.97, 4.10)	22.52 ( 2.44, 42.60)	
100- < 200 U/L	14/ 37 ( 37.8)	2/ 19 ( 10.5)			5.17 ( 1.04, 25.85)	0.034	3.59 ( 0.91, 14.21)	27.31 ( 6.46, 48.16)	
200-500 U/L	17/ 43 ( 39.5)	1/ 15 ( 6.7)			9.15 ( 1.10, 76.18)	0.019	5.93 ( 0.86, 40.83)	32.87 ( 13.56, 52.18)	
> 500 U/L	3/ 21 ( 14.3)	0/ 11 ( 0.0)			4.35 ( 0.21, 92.18)	0.195	3.82 ( 0.21, 67.92)	14.29 (-0.68, 29.25)	
SF3B1 mutation status									0.350
Mutated	53/141 ( 37.6)	10/ 65 ( 15.4)			3.31 ( 1.56, 7.05)	0.001	2.44 ( 1.33, 4.49)	22.20 (10.34, 34.07)	
Non-mutated	5/ 12 ( 41.7)	0/ 10 ( 0.0)			15.40 ( 0.73,322.88)	0.023	9.31 ( 0.58,150.25)	41.67 (13.77, 69.56)	
Longest duration of prior ESA treatment									0.753
< 6 months	16/ 48 ( 33.3)	1/ 20 ( 5.0)			9.50 ( 1.17, 77.46)	0.015	6.67 ( 0.95, 46.94)	28.33 (11.93, 44.74)	
6-12 months	15/ 34 ( 44.1)	1/ 9 ( 11.1)			6.32 ( 0.71, 56.22)	0.072	3.97 ( 0.60, 26.18)	33.01 ( 6.55, 59.47)	
>12-24 months	15/ 35 ( 42.9)	4/ 21 ( 19.0)			3.19 ( 0.89, 11.45)	0.071	2.25 ( 0.86, 5.88)	23.81 ( 0.34, 47.28)	
> 24 months	12/ 31 ( 38.7)	3/ 20 ( 15.0)			3.58 ( 0.86, 14.87)	0.072	2.58 ( 0.83, 8.02)	23.71 ( 0.50, 46.92)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 RBC-TI >=12 weeks during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									0.551
<=64 years	16/ 29 ( 55.2)	2/ 16 ( 12.5)			8.62 ( 1.65, 44.97)	0.006	4.41 ( 1.16, 16.81)	42.67 (18.38, 66.97)	
65-74 years	16/ 72 ( 22.2)	3/ 29 ( 10.3)			2.48 ( 0.66, 9.25)	0.169	2.15 ( 0.68, 6.82)	11.88 (-2.79, 26.54)	
>=75 years	11/ 52 ( 21.2)	1/ 31 ( 3.2)			8.05 ( 0.99, 65.77)	0.026	6.56 ( 0.89, 48.37)	17.93 ( 5.20, 30.65)	
Gender									0.689
female	20/ 59 ( 33.9)	3/ 26 ( 11.5)			3.93 ( 1.05, 14.69)	0.034	2.94 ( 0.96, 9.03)	22.36 ( 5.13, 39.58)	
male	23/ 94 ( 24.5)	3/ 50 ( 6.0)			5.08 ( 1.44, 17.86)	0.006	4.08 ( 1.29, 12.92)	18.47 ( 7.57, 29.37)	
Race									0.599
Black or African American	1/ 1 (100.0)	0			NE	NE	NE	NE	
White	32/107 ( 29.9)	5/ 51 ( 9.8)			3.93 ( 1.43, 10.79)	0.005	3.05 ( 1.26, 7.37)	20.10 ( 8.19, 32.01)	
Not Collected or Reported	10/ 44 ( 22.7)	1/ 24 ( 4.2)			6.76 ( 0.81, 56.51)	0.049	5.45 ( 0.74, 40.08)	18.56 ( 3.82, 33.30)	
Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)			NE	NE	NE	NE	
Region									0.480
North America	13/ 31 ( 41.9)	3/ 19 ( 15.8)			3.85 ( 0.93, 16.01)	0.057	2.66 ( 0.87, 8.12)	26.15 ( 2.26, 50.03)	
EMEA	30/122 ( 24.6)	3/ 57 ( 5.3)			5.87 ( 1.71, 20.15)	0.002	4.67 ( 1.49, 14.67)	19.33 ( 9.74, 28.92)	
MDS WHO classification									0.669
MDS RARS	2/ 7 ( 28.6)	0/ 2 ( 0.0)			2.27 ( 0.08, 67.05)	0.419	1.88 ( 0.12, 28.78)	28.57 (-4.89, 62.04)	
MDS RCMD	41/145 ( 28.3)	6/ 74 ( 8.1)			4.47 ( 1.80, 11.10)	<.001	3.49 ( 1.55, 7.84)	20.17 (10.55, 29.78)	
Other	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Time since initial diagnosis									0.353
<= 2 years	7/ 40 ( 17.5)	1/ 19 ( 5.3)			3.82 ( 0.43, 33.53)	0.203	3.33 ( 0.44, 25.14)	12.24 (-3.24, 27.71)	
>2 to 5 years	24/ 62 ( 38.7)	2/ 34 ( 5.9)			10.11 ( 2.22, 46.07)	<.001	6.58 ( 1.65, 26.17)	32.83 (18.35, 47.30)	
> 5 years	12/ 51 ( 23.5)	3/ 23 ( 13.0)			2.05 ( 0.52, 8.11)	0.302	1.80 ( 0.56, 5.79)	10.49 (-7.54, 28.51)	
ECOG Performance Status									0.835
0	16/ 54 ( 29.6)	2/ 33 ( 6.1)			6.53 ( 1.39, 30.58)	0.009	4.89 ( 1.20, 19.92)	23.57 ( 8.92, 38.22)	
1	25/ 91 ( 27.5)	3/ 32 ( 9.4)			3.66 ( 1.02, 13.10)	0.036	2.93 ( 0.95, 9.05)	18.10 ( 4.46, 31.74)	
2	2/ 8 ( 25.0)	1/ 11 ( 9.1)			3.33 ( 0.25, 45.11)	0.361	2.75 ( 0.30, 25.35)	15.91 (-18.6, 50.39)	
RBC Transfusions/8 weeks over 16 weeks									0.578
< 6 units	39/ 87 ( 44.8)	5/ 43 ( 11.6)			6.18 ( 2.22, 17.19)	<.001	3.86 ( 1.64, 9.08)	33.20 (19.02, 47.38)	
>= 6 units	4/ 66 ( 6.1)	1/ 33 ( 3.0)			2.06 ( 0.22, 19.25)	0.518	2.00 ( 0.23, 17.19)	3.03 (-5.18, 11.24)	
RBC Transfusions/8 weeks over 16 weeks II									0.780
< 4 units	29/ 46 ( 63.0)	4/ 20 ( 20.0)			6.82 ( 1.96, 23.78)	0.001	3.15 ( 1.28, 7.78)	43.04 (20.64, 65.45)	
>= 4 and < 6 units	10/ 41 ( 24.4)	1/ 23 ( 4.3)			7.10 ( 0.85, 59.54)	0.043	5.61 ( 0.77, 41.09)	20.04 ( 4.48, 35.61)	
>= 6 units	4/ 66 ( 6.1)	1/ 33 ( 3.0)			2.06 ( 0.22, 19.25)	0.518	2.00 ( 0.23, 17.19)	3.03 (-5.18, 11.24)	
IPSS-R risk									0.703
Very Low	5/ 18 ( 27.8)	1/ 6 ( 16.7)			1.92 ( 0.18, 20.82)	0.594	1.67 ( 0.24, 11.58)	11.11 (-25.2, 47.41)	
Low	32/109 ( 29.4)	4/ 57 ( 7.0)			5.51 ( 1.84, 16.49)	<.001	4.18 ( 1.56, 11.25)	22.34 (11.52, 33.16)	
Intermediate	6/ 25 ( 24.0)	1/ 13 ( 7.7)			3.79 ( 0.40, 35.49)	0.225	3.12 ( 0.42, 23.24)	16.31 (-5.83, 38.45)	
High	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Platelet count									0.428
< 100 x 10 <sup>9</sup> /L	1/ 8 ( 12.5)	1/ 6 ( 16.7)			0.71 ( 0.04, 14.35)	0.832	0.75 ( 0.06, 9.72)	-4.17 (-41.8, 33.44)	
100-400 x 10 <sup>9</sup> /L	31/128 ( 24.2)	4/ 61 ( 6.6)			4.55 ( 1.53, 13.56)	0.004	3.69 ( 1.36, 10.00)	17.66 ( 7.98, 27.34)	
> 400 x 10 <sup>9</sup> /L	11/ 17 ( 64.7)	1/ 9 ( 11.1)			14.67 ( 1.46, 146.96)	0.011	5.82 ( 0.89, 38.20)	53.59 (22.97, 84.22)	
Renal status: Creatinine Clearance									0.727
40 - <60 mL/min	7/ 33 ( 21.2)	1/ 12 ( 8.3)			2.96 ( 0.32, 27.02)	0.323	2.55 ( 0.35, 18.59)	12.88 (-8.08, 33.83)	
>= 60 mL/min	33/113 ( 29.2)	4/ 62 ( 6.5)			5.98 ( 2.01, 17.81)	<.001	4.53 ( 1.68, 12.19)	22.75 (12.38, 33.13)	
< 40 mL/min	3/ 7 ( 42.9)	0/ 1 ( 0.0)			2.33 ( 0.07, 76.67)	0.439	1.75 ( 0.14, 21.88)	42.86 ( 6.20, 79.52)	
Serum EPO									0.691

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

RBC-TI >=12 weeks during week 1-24 - Subgroup analysis

ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
< 100 U/L	18/ 51 ( 35.3)	5/ 31 ( 16.1)	2.84 ( 0.93, 8.66)	0.063	2.19 ( 0.90, 5.30)	19.17 ( 0.74, 37.59)			
100- < 200 U/L	11/ 37 ( 29.7)	1/ 19 ( 5.3)	7.62 ( 0.90, 64.31)	0.036	5.65 ( 0.79, 40.54)	24.47 ( 6.64, 42.29)			
200-500 U/L	10/ 43 ( 23.3)	0/ 15 ( 0.0)	9.72 ( 0.53,176.63)	0.042	7.64 ( 0.47,122.94)	23.26 (10.63, 35.88)			
> 500 U/L	3/ 21 ( 14.3)	0/ 11 ( 0.0)	4.35 ( 0.21, 92.18)	0.195	3.82 ( 0.21, 67.92)	14.29 (-0.68, 29.25)			
SF3B1 mutation status								0.529	
Mutated	39/141 ( 27.7)	6/ 65 ( 9.2)	3.76 ( 1.50, 9.41)	0.003	3.00 ( 1.34, 6.72)	18.43 ( 8.23, 28.63)			
Non-mutated	4/ 12 ( 33.3)	0/ 10 ( 0.0)	11.12 ( 0.52,236.75)	0.049	7.62 ( 0.46,126.40)	33.33 ( 6.66, 60.01)			
Longest duration of prior ESA treatment								0.860	
< 6 months	11/ 48 ( 22.9)	0/ 20 ( 0.0)	12.57 ( 0.70,224.47)	0.020	9.86 ( 0.61,159.67)	22.92 (11.03, 34.81)			
6-12 months	13/ 34 ( 38.2)	1/ 9 ( 11.1)	4.95 ( 0.55, 44.29)	0.127	3.44 ( 0.52, 22.93)	27.12 ( 0.89, 53.36)			
>12-24 months	11/ 35 ( 31.4)	2/ 21 ( 9.5)	4.35 ( 0.86, 22.05)	0.062	3.30 ( 0.81, 13.46)	21.90 ( 2.05, 41.76)			
> 24 months	8/ 31 ( 25.8)	2/ 20 ( 10.0)	3.13 ( 0.59, 16.59)	0.169	2.58 ( 0.61, 10.93)	15.81 (-4.45, 36.06)			

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 RBC-TI >=16 weeks during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									
<=64 years	10/ 29 ( 34.5)	1/ 16 ( 6.3)			7.89 ( 0.91, 68.75)	0.037	5.52 ( 0.77, 39.28)	28.23 ( 7.26, 49.21)	0.632
65-74 years	12/ 72 ( 16.7)	2/ 29 ( 6.9)			2.70 ( 0.56, 12.90)	0.201	2.42 ( 0.58, 10.13)	9.77 (-2.85, 22.39)	
>=75 years	7/ 52 ( 13.5)	0/ 31 ( 0.0)			10.38 ( 0.57,188.45)	0.034	9.06 ( 0.54,153.30)	13.46 ( 4.18, 22.74)	
Gender									0.992
female	11/ 59 ( 18.6)	1/ 26 ( 3.8)			5.73 ( 0.70, 46.95)	0.073	4.85 ( 0.66, 35.62)	14.80 ( 2.41, 27.18)	
male	18/ 94 ( 19.1)	2/ 50 ( 4.0)			5.68 ( 1.26, 25.60)	0.013	4.79 ( 1.16, 19.81)	15.15 ( 5.52, 24.78)	
Race									0.551
Black or African American	1/ 1 (100.0)	0			NE	NE	NE	NE	
White	21/107 ( 19.6)	3/ 51 ( 5.9)			3.91 ( 1.11, 13.78)	0.025	3.34 ( 1.04, 10.67)	13.74 ( 3.83, 23.66)	
Not Collected or Reported	7/ 44 ( 15.9)	0/ 24 ( 0.0)			9.80 ( 0.54,179.47)	0.041	8.33 ( 0.50,139.90)	15.91 ( 5.10, 26.72)	
Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)			NE	NE	NE	NE	
Region									0.369
North America	10/ 31 ( 32.3)	2/ 19 ( 10.5)			4.05 ( 0.78, 21.02)	0.084	3.06 ( 0.75, 12.51)	21.73 ( 0.26, 43.21)	
EMEA	19/122 ( 15.6)	1/ 57 ( 1.8)			10.33 ( 1.35, 79.21)	0.006	8.88 ( 1.22, 64.69)	13.82 ( 6.54, 21.10)	
MDS WHO classification									0.364
MDS RARS	1/ 7 ( 14.3)	0/ 2 ( 0.0)			1.15 ( 0.03, 38.88)	0.593	1.13 ( 0.06, 20.71)	14.29 (-11.6, 40.21)	
MDS RCMD	28/145 ( 19.3)	3/ 74 ( 4.1)			5.66 ( 1.66, 19.31)	0.002	4.76 ( 1.50, 15.15)	15.26 ( 7.42, 23.10)	
Other	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Time since initial diagnosis									0.109
<= 2 years	5/ 40 ( 12.5)	1/ 19 ( 5.3)			2.57 ( 0.28, 23.70)	0.394	2.38 ( 0.30, 18.94)	7.24 (-7.11, 21.58)	
>2 to 5 years	19/ 62 ( 30.6)	0/ 34 ( 0.0)			30.93 ( 1.80,530.71)	<.001	21.67 ( 1.35,348.04)	30.65 (19.17, 42.12)	
> 5 years	5/ 51 ( 9.8)	2/ 23 ( 8.7)			1.14 ( 0.20, 6.37)	0.881	1.13 ( 0.24, 5.39)	1.11 (-13.0, 15.22)	
ECOG Performance Status									0.590
0	7/ 54 ( 13.0)	2/ 33 ( 6.1)			2.31 ( 0.45, 11.85)	0.308	2.14 ( 0.47, 9.69)	6.90 (-5.20, 19.01)	
1	21/ 91 ( 23.1)	1/ 32 ( 3.1)			9.30 ( 1.20, 72.25)	0.012	7.38 ( 1.03, 52.70)	19.95 ( 9.40, 30.50)	
2	1/ 8 ( 12.5)	0/ 11 ( 0.0)			4.60 ( 0.16,128.54)	0.241	4.00 ( 0.18, 87.18)	12.50 (-10.4, 35.42)	
RBC Transfusions/8 weeks over 16 weeks									0.383
< 6 units	25/ 87 ( 28.7)	2/ 43 ( 4.7)			8.27 ( 1.86, 36.80)	0.002	6.18 ( 1.53, 24.88)	24.08 (12.68, 35.49)	
>= 6 units	4/ 66 ( 6.1)	1/ 33 ( 3.0)			2.06 ( 0.22, 19.25)	0.518	2.00 ( 0.23, 17.19)	3.03 (-5.18, 11.24)	
RBC Transfusions/8 weeks over 16 weeks II									0.747
< 4 units	19/ 46 ( 41.3)	2/ 20 ( 10.0)			6.33 ( 1.31, 30.58)	0.013	4.13 ( 1.06, 16.08)	31.30 (11.93, 50.68)	
>= 4 and < 6 units	6/ 41 ( 14.6)	0/ 23 ( 0.0)			8.61 ( 0.46,160.08)	0.056	7.43 ( 0.44,126.18)	14.63 ( 3.82, 25.45)	
>= 6 units	4/ 66 ( 6.1)	1/ 33 ( 3.0)			2.06 ( 0.22, 19.25)	0.518	2.00 ( 0.23, 17.19)	3.03 (-5.18, 11.24)	
IPSS-R risk									0.212
Very Low	2/ 18 ( 11.1)	1/ 6 ( 16.7)			0.63 ( 0.05, 8.43)	0.727	0.67 ( 0.07, 6.11)	-5.56 (-38.7, 27.61)	
Low	25/109 ( 22.9)	2/ 57 ( 3.5)			8.18 ( 1.86, 35.95)	0.001	6.54 ( 1.61, 26.62)	19.43 (10.20, 28.65)	
Intermediate	2/ 25 ( 8.0)	0/ 13 ( 0.0)			2.87 ( 0.13, 64.35)	0.301	2.69 ( 0.14, 52.27)	8.00 (-2.63, 18.63)	
High	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Platelet count									0.341
< 100 x 10 <sup>9</sup> /L	1/ 8 ( 12.5)	1/ 6 ( 16.7)			0.71 ( 0.04, 14.35)	0.832	0.75 ( 0.06, 9.72)	-4.17 (-41.8, 33.44)	
100-400 x 10 <sup>9</sup> /L	20/128 ( 15.6)	2/ 61 ( 3.3)			5.46 ( 1.23, 24.19)	0.014	4.77 ( 1.15, 19.74)	12.35 ( 4.63, 20.06)	
> 400 x 10 <sup>9</sup> /L	8/ 17 ( 47.1)	0/ 9 ( 0.0)			17.00 ( 0.85,338.25)	0.015	9.44 ( 0.61,147.06)	47.06 (23.33, 70.79)	
Renal status: Creatinine Clearance									0.573
40 - <60 mL/min	5/ 33 ( 15.2)	0/ 12 ( 0.0)			4.82 ( 0.25, 94.08)	0.157	4.21 ( 0.25, 70.82)	15.15 ( 2.92, 27.38)	
>= 60 mL/min	22/113 ( 19.5)	2/ 62 ( 3.2)			7.25 ( 1.64, 31.98)	0.003	6.04 ( 1.47, 24.82)	16.24 ( 7.72, 24.77)	
< 40 mL/min	2/ 7 ( 28.6)	0/ 1 ( 0.0)			1.36 ( 0.04, 46.65)	0.564	1.25 ( 0.09, 17.02)	28.57 (-4.89, 62.04)	
Serum EPO									0.875

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 RBC-TI >=16 weeks during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
< 100 U/L	13/ 51 ( 25.5)	3/ 31 ( 9.7)	3.19 ( 0.83, 12.28)	0.082	2.63 ( 0.81, 8.52)	15.81 (-0.04, 31.67)			
100- < 200 U/L	7/ 37 ( 18.9)	0/ 19 ( 0.0)	9.59 ( 0.52,177.56)	0.045	7.89 ( 0.47,131.27)	18.92 ( 6.30, 31.54)			
200-500 U/L	7/ 43 ( 16.3)	0/ 15 ( 0.0)	6.37 ( 0.34,118.55)	0.099	5.45 ( 0.33, 90.14)	16.28 ( 5.24, 27.31)			
> 500 U/L	2/ 21 ( 9.5)	0/ 11 ( 0.0)	2.95 ( 0.13, 66.95)	0.298	2.73 ( 0.14, 52.30)	9.52 (-3.03, 22.08)			
SF3B1 mutation status								0.990	
Mutated	27/141 ( 19.1)	3/ 65 ( 4.6)	4.89 ( 1.43, 16.78)	0.006	4.15 ( 1.31, 13.18)	14.53 ( 6.28, 22.79)			
Non-mutated	2/ 12 ( 16.7)	0/ 10 ( 0.0)	5.00 ( 0.21,117.21)	0.186	4.23 ( 0.23, 79.10)	16.67 (-4.42, 37.75)			
Longest duration of prior ESA treatment								0.809	
< 6 months	10/ 48 ( 20.8)	0/ 20 ( 0.0)	11.18 ( 0.62,200.64)	0.028	9.00 ( 0.55,146.62)	20.83 ( 9.34, 32.32)			
6-12 months	8/ 34 ( 23.5)	1/ 9 ( 11.1)	2.46 ( 0.27, 22.77)	0.421	2.12 ( 0.30, 14.81)	12.42 (-12.6, 37.42)			
>12-24 months	6/ 35 ( 17.1)	1/ 21 ( 4.8)	4.14 ( 0.46, 37.06)	0.179	3.60 ( 0.46, 27.87)	12.38 (-3.07, 27.84)			
> 24 months	5/ 31 ( 16.1)	0/ 20 ( 0.0)	8.51 ( 0.44,162.89)	0.061	7.22 ( 0.42,123.83)	16.13 ( 3.18, 29.08)			

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 RBC-TI >=24 weeks during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									
<=64 years	8/ 29 ( 27.6)	0/ 16 ( 0.0)	13.05 ( 0.70,242.73)	0.022	9.63 ( 0.59,156.75)	27.59 (11.32, 43.85)		0.816	
65-74 years	8/ 72 ( 11.1)	1/ 29 ( 3.4)	3.50 ( 0.42, 29.33)	0.224	3.22 ( 0.42, 24.62)	7.66 (-2.18, 17.50)			
>=75 years	4/ 52 ( 7.7)	0/ 31 ( 0.0)	5.85 ( 0.30,112.35)	0.116	5.43 ( 0.30, 97.64)	7.69 ( 0.45, 14.93)			
Gender									
female	8/ 59 ( 13.6)	0/ 26 ( 0.0)	8.75 ( 0.49,157.46)	0.050	7.65 ( 0.46,127.80)	13.56 ( 4.82, 22.30)		0.918	
male	12/ 94 ( 12.8)	1/ 50 ( 2.0)	7.17 ( 0.90, 56.85)	0.032	6.38 ( 0.85, 47.68)	10.77 ( 2.98, 18.55)			
Race									
Black or African American	0/ 1 ( 0.0)	0	NE	NE	NE	NE		0.813	
White	16/107 ( 15.0)	1/ 51 ( 2.0)	8.79 ( 1.13, 68.26)	0.014	7.63 ( 1.04, 55.93)	12.99 ( 5.24, 20.75)			
Not Collected or Reported	4/ 44 ( 9.1)	0/ 24 ( 0.0)	5.44 ( 0.28,105.54)	0.131	5.00 ( 0.28, 89.12)	9.09 ( 0.60, 17.59)			
Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE			
Region									
North America	8/ 31 ( 25.8)	0/ 19 ( 0.0)	14.11 ( 0.76,260.16)	0.017	10.63 ( 0.65,174.19)	25.81 (10.40, 41.21)		0.715	
EMEA	12/122 ( 9.8)	1/ 57 ( 1.8)	6.11 ( 0.77, 48.18)	0.053	5.61 ( 0.75, 42.08)	8.08 ( 1.79, 14.37)			
MDS WHO classification									
MDS RARS	1/ 7 ( 14.3)	0/ 2 ( 0.0)	1.15 ( 0.03, 38.88)	0.593	1.13 ( 0.06, 20.71)	14.29 (-11.6, 40.21)		0.215	
MDS RCMD	19/145 ( 13.1)	1/ 74 ( 1.4)	11.01 ( 1.44, 83.94)	0.004	9.70 ( 1.32, 71.03)	11.75 ( 5.66, 17.84)			
Other	0/ 1 ( 0.0)	0	NE	NE	NE	NE			
Time since initial diagnosis									
<= 2 years	5/ 40 ( 12.5)	0/ 19 ( 0.0)	6.04 ( 0.32,115.13)	0.110	5.37 ( 0.31, 92.33)	12.50 ( 2.25, 22.75)		0.505	
>2 to 5 years	11/ 62 ( 17.7)	0/ 34 ( 0.0)	15.41 ( 0.88,270.14)	0.009	12.78 ( 0.78,210.36)	17.74 ( 8.23, 27.25)			
> 5 years	4/ 51 ( 7.8)	1/ 23 ( 4.3)	1.87 ( 0.20, 17.75)	0.582	1.80 ( 0.21, 15.26)	3.50 (-7.64, 14.63)			
ECOG Performance Status									
0	5/ 54 ( 9.3)	1/ 33 ( 3.0)	3.27 ( 0.36, 29.26)	0.269	3.06 ( 0.37, 25.02)	6.23 (-3.47, 15.92)		0.765	
1	14/ 91 ( 15.4)	0/ 32 ( 0.0)	12.16 ( 0.70,209.99)	0.019	10.40 ( 0.64,169.53)	15.38 ( 7.97, 22.80)			
2	1/ 8 ( 12.5)	0/ 11 ( 0.0)	4.60 ( 0.16,128.54)	0.241	4.00 ( 0.18, 87.18)	12.50 (-10.4, 35.42)			
RBC Transfusions/8 weeks over 16 weeks									
< 6 units	18/ 87 ( 20.7)	0/ 43 ( 0.0)	23.16 ( 1.36,394.18)	0.001	18.50 ( 1.14,299.88)	20.69 (12.18, 29.20)		0.081	
>= 6 units	2/ 66 ( 3.0)	1/ 33 ( 3.0)	1.00 ( 0.09, 11.45)	1.000	1.00 ( 0.09, 10.63)	0.00 (-7.16, 7.16)			
RBC Transfusions/8 weeks over 16 weeks II									
< 4 units	13/ 46 ( 28.3)	0/ 20 ( 0.0)	16.52 ( 0.93,293.07)	0.008	12.06 ( 0.75,193.56)	28.26 (15.25, 41.27)		0.327	
>= 4 and < 6 units	5/ 41 ( 12.2)	0/ 23 ( 0.0)	7.08 ( 0.37,134.11)	0.084	6.29 ( 0.36,108.81)	12.20 ( 2.18, 22.21)			
>= 6 units	2/ 66 ( 3.0)	1/ 33 ( 3.0)	1.00 ( 0.09, 11.45)	1.000	1.00 ( 0.09, 10.63)	0.00 (-7.16, 7.16)			
IPSS-R risk									
Very Low	1/ 18 ( 5.6)	0/ 6 ( 0.0)	1.11 ( 0.04, 30.97)	0.564	1.11 ( 0.05, 24.07)	5.56 (-5.03, 16.14)		0.398	
Low	18/109 ( 16.5)	1/ 57 ( 1.8)	11.08 ( 1.44, 85.27)	0.005	9.41 ( 1.29, 68.72)	14.76 ( 7.00, 22.52)			
Intermediate	1/ 25 ( 4.0)	0/ 13 ( 0.0)	1.65 ( 0.06, 43.44)	0.471	1.62 ( 0.07, 37.10)	4.00 (-3.68, 11.68)			
High	0/ 1 ( 0.0)	0	NE	NE	NE	NE			
Platelet count									
< 100 x 10 <sup>9</sup> /L	1/ 8 ( 12.5)	1/ 6 ( 16.7)	0.71 ( 0.04, 14.35)	0.832	0.75 ( 0.06, 9.72)	-4.17 (-41.8, 33.44)		0.247	
100-400 x 10 <sup>9</sup> /L	13/128 ( 10.2)	0/ 61 ( 0.0)	14.38 ( 0.84,245.96)	0.010	12.98 ( 0.78,214.76)	10.16 ( 4.92, 15.39)			
> 400 x 10 <sup>9</sup> /L	6/ 17 ( 35.3)	0/ 9 ( 0.0)	10.74 ( 0.53,216.14)	0.046	7.22 ( 0.45,115.33)	35.29 (12.58, 58.01)			
Renal status: Creatinine Clearance									
40 - <60 mL/min	3/ 33 ( 9.1)	0/ 12 ( 0.0)	2.87 ( 0.14, 59.70)	0.285	2.68 ( 0.15, 48.32)	9.09 (-0.72, 18.90)		0.335	
>= 60 mL/min	16/113 ( 14.2)	1/ 62 ( 1.6)	10.06 ( 1.30, 77.81)	0.008	8.78 ( 1.19, 64.63)	12.55 ( 5.39, 19.70)			
< 40 mL/min	1/ 7 ( 14.3)	0/ 1 ( 0.0)	0.69 ( 0.02, 26.90)	0.705	0.75 ( 0.05, 12.34)	14.29 (-11.6, 40.21)			
Serum EPO									0.974

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

RBC-TI >=24 weeks during week 1-24 - Subgroup analysis

ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
< 100 U/L	9/ 51 ( 17.6)	1/ 31 ( 3.2)	6.43 ( 0.77, 53.47)	0.054	5.47 ( 0.73, 41.12)	14.42 ( 2.25, 26.59)			
100- < 200 U/L	5/ 37 ( 13.5)	0/ 19 ( 0.0)	6.60 ( 0.35,125.97)	0.096	5.79 ( 0.34, 99.49)	13.51 ( 2.50, 24.53)			
200-500 U/L	4/ 43 ( 9.3)	0/ 15 ( 0.0)	3.53 ( 0.18, 69.55)	0.225	3.27 ( 0.19, 57.44)	9.30 ( 0.62, 17.98)			
> 500 U/L	2/ 21 ( 9.5)	0/ 11 ( 0.0)	2.95 ( 0.13, 66.95)	0.298	2.73 ( 0.14, 52.30)	9.52 (-3.03, 22.08)			
SF3B1 mutation status								NE	
Mutated	20/141 ( 14.2)	1/ 65 ( 1.5)	10.58 ( 1.39, 80.63)	0.005	9.22 ( 1.26, 67.23)	12.65 ( 6.16, 19.14)			
Non-mutated	0/ 12 ( 0.0)	0/ 10 ( 0.0)	NE	NE	NE	NE			
Longest duration of prior ESA treatment								0.590	
< 6 months	7/ 48 ( 14.6)	0/ 20 ( 0.0)	7.41 ( 0.40,136.18)	0.073	6.43 ( 0.38,107.50)	14.58 ( 4.60, 24.57)			
6-12 months	4/ 34 ( 11.8)	1/ 9 ( 11.1)	1.07 ( 0.10, 10.92)	0.957	1.06 ( 0.13, 8.34)	0.65 (-22.6, 23.87)			
>12-24 months	4/ 35 ( 11.4)	0/ 21 ( 0.0)	6.14 ( 0.31,120.06)	0.111	5.50 ( 0.31, 97.31)	11.43 ( 0.89, 21.97)			
> 24 months	5/ 31 ( 16.1)	0/ 20 ( 0.0)	8.51 ( 0.44,162.89)	0.061	7.22 ( 0.42,123.83)	16.13 ( 3.18, 29.08)			

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 mHI-E per IWG 2006 during any 8 week period during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									0.465
<=64 years	18/ 29 ( 62.1)	2/ 16 ( 12.5)			11.45 ( 2.18, 60.28)	0.002	4.97 ( 1.32, 18.72)	49.57 (25.60, 73.54)	
65-74 years	37/ 72 ( 51.4)	5/ 29 ( 17.2)			5.07 ( 1.74, 14.77)	0.002	2.98 ( 1.30, 6.82)	34.15 (16.20, 52.10)	
>=75 years	26/ 52 ( 50.0)	2/ 31 ( 6.5)			14.50 ( 3.13, 67.13)	<.001	7.75 ( 1.97, 30.43)	43.55 (27.44, 59.66)	
Gender									0.445
female	31/ 59 ( 52.5)	2/ 26 ( 7.7)			13.29 ( 2.88, 61.38)	<.001	6.83 ( 1.76, 26.44)	44.85 (28.50, 61.20)	
male	50/ 94 ( 53.2)	7/ 50 ( 14.0)			6.98 ( 2.85, 17.10)	<.001	3.80 ( 1.86, 7.75)	39.19 (25.25, 53.13)	
Race									0.486
Black or African American	1/ 1 (100.0)	0			NE	NE	NE	NE	
White	56/107 ( 52.3)	7/ 51 ( 13.7)			6.90 ( 2.85, 16.69)	<.001	3.81 ( 1.87, 7.77)	38.61 (25.24, 51.98)	
Not Collected or Reported	24/ 44 ( 54.5)	2/ 24 ( 8.3)			13.20 ( 2.76, 63.10)	<.001	6.55 ( 1.69, 25.35)	46.21 (27.81, 64.62)	
Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)			NE	NE	NE	NE	
Region									0.875
North America	16/ 31 ( 51.6)	2/ 19 ( 10.5)			9.07 ( 1.78, 46.08)	0.004	4.90 ( 1.27, 19.00)	41.09 (18.73, 63.44)	
EMEA	65/122 ( 53.3)	7/ 57 ( 12.3)			8.15 ( 3.42, 19.39)	<.001	4.34 ( 2.13, 8.85)	41.00 (28.71, 53.29)	
MDS WHO classification									0.858
MDS RARS	4/ 7 ( 57.1)	0/ 2 ( 0.0)			6.43 ( 0.23,181.82)	0.176	3.37 ( 0.25, 45.58)	57.14 (20.48, 93.80)	
MDS RCMD	76/145 ( 52.4)	9/ 74 ( 12.2)			7.95 ( 3.69, 17.17)	<.001	4.31 ( 2.29, 8.10)	40.25 (29.23, 51.28)	
Other	1/ 1 (100.0)	0			NE	NE	NE	NE	
Time since initial diagnosis									0.146
<= 2 years	16/ 40 ( 40.0)	0/ 19 ( 0.0)			26.27 ( 1.48,465.88)	0.001	16.10 ( 1.02,254.91)	40.00 (24.82, 55.18)	
>2 to 5 years	41/ 62 ( 66.1)	4/ 34 ( 11.8)			14.64 ( 4.55, 47.10)	<.001	5.62 ( 2.20, 14.36)	54.36 (38.36, 70.37)	
> 5 years	24/ 51 ( 47.1)	5/ 23 ( 21.7)			3.20 ( 1.03, 9.94)	0.040	2.16 ( 0.95, 4.96)	25.32 ( 3.60, 47.04)	
ECOG Performance Status									0.566
0	26/ 54 ( 48.1)	2/ 33 ( 6.1)			14.39 ( 3.13, 66.22)	<.001	7.94 ( 2.02, 31.31)	42.09 (26.47, 57.70)	
1	50/ 91 ( 54.9)	5/ 32 ( 15.6)			6.59 ( 2.33, 18.63)	<.001	3.52 ( 1.54, 8.04)	39.32 (23.11, 55.53)	
2	5/ 8 ( 62.5)	2/ 11 ( 18.2)			7.50 ( 0.92, 61.05)	0.054	3.44 ( 0.88, 13.44)	44.32 ( 3.76, 84.88)	
RBC Transfusions/8 weeks over 16 weeks									0.042
< 6 units	45/ 87 ( 51.7)	2/ 43 ( 4.7)			21.96 ( 5.00, 96.52)	<.001	11.12 ( 2.83, 43.70)	47.07 (34.83, 59.32)	
>= 6 units	36/ 66 ( 54.5)	7/ 33 ( 21.2)			4.46 ( 1.70, 11.70)	0.002	2.57 ( 1.29, 5.14)	33.33 (14.93, 51.74)	
RBC Transfusions/8 weeks over 16 weeks II									0.136
< 4 units	29/ 46 ( 63.0)	1/ 20 ( 5.0)			32.41 ( 3.98,264.18)	<.001	12.61 ( 1.84, 86.27)	58.04 (41.14, 74.95)	
>= 4 and < 6 units	16/ 41 ( 39.0)	1/ 23 ( 4.3)			14.08 ( 1.72,114.97)	0.003	8.98 ( 1.27, 63.38)	34.68 (17.58, 51.78)	
>= 6 units	36/ 66 ( 54.5)	7/ 33 ( 21.2)			4.46 ( 1.70, 11.70)	0.002	2.57 ( 1.29, 5.14)	33.33 (14.93, 51.74)	
IPSS-R risk									0.339
Very Low	12/ 18 ( 66.7)	2/ 6 ( 33.3)			4.00 ( 0.56, 28.40)	0.160	2.00 ( 0.62, 6.49)	33.33 (-10.2, 76.89)	
Low	55/109 ( 50.5)	5/ 57 ( 8.8)			10.59 ( 3.93, 28.55)	<.001	5.75 ( 2.44, 13.56)	41.69 (29.77, 53.60)	
Intermediate	14/ 25 ( 56.0)	2/ 13 ( 15.4)			7.00 ( 1.28, 38.36)	0.018	3.64 ( 0.97, 13.64)	40.62 (12.99, 68.24)	
High	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Platelet count									0.084
< 100 x 10 <sup>9</sup> /L	2/ 8 ( 25.0)	2/ 6 ( 33.3)			0.67 ( 0.06, 6.87)	0.742	0.75 ( 0.14, 3.90)	-8.33 (-56.5, 39.87)	
100-400 x 10 <sup>9</sup> /L	65/128 ( 50.8)	6/ 61 ( 9.8)			9.46 ( 3.80, 23.52)	<.001	5.16 ( 2.37, 11.25)	40.95 (29.51, 52.38)	
> 400 x 10 <sup>9</sup> /L	14/ 17 ( 82.4)	1/ 9 ( 11.1)			37.33 ( 3.31,421.59)	<.001	7.41 ( 1.15, 47.66)	71.24 (43.86, 98.63)	
Renal status: Creatinine Clearance									0.816
40 - <60 mL/min	17/ 33 ( 51.5)	1/ 12 ( 8.3)			11.69 ( 1.35,101.14)	0.010	6.18 ( 0.92, 41.56)	43.18 (20.05, 66.32)	
>= 60 mL/min	60/113 ( 53.1)	8/ 62 ( 12.9)			7.64 ( 3.33, 17.51)	<.001	4.12 ( 2.11, 8.04)	40.19 (27.77, 52.62)	
< 40 mL/min	4/ 7 ( 57.1)	0/ 1 ( 0.0)			3.86 ( 0.12,126.73)	0.317	2.25 ( 0.19, 26.79)	57.14 (20.48, 93.80)	
Serum EPO									0.988

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

mHI-E per IWG 2006 during any 8 week period during week 1-24 - Subgroup analysis

ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
< 100 U/L	33/ 51 ( 64.7)	4/ 31 ( 12.9)	12.38 ( 3.74, 40.96)	<.001	5.01 ( 1.97, 12.80)	51.80 (34.16, 69.45)			
100- < 200 U/L	16/ 37 ( 43.2)	2/ 19 ( 10.5)	6.48 ( 1.30, 32.17)	0.014	4.11 ( 1.05, 16.04)	32.72 (11.62, 53.82)			
200-500 U/L	22/ 43 ( 51.2)	2/ 15 ( 13.3)	6.81 ( 1.37, 33.87)	0.011	3.84 ( 1.02, 14.41)	37.83 (15.04, 60.61)			
> 500 U/L	9/ 21 ( 42.9)	1/ 11 ( 9.1)	7.50 ( 0.81, 69.74)	0.054	4.71 ( 0.68, 32.57)	33.77 ( 6.63, 60.91)			
SF3B1 mutation status								0.339	
Mutated	73/141 ( 51.8)	9/ 65 ( 13.8)	6.68 ( 3.07, 14.54)	<.001	3.74 ( 2.00, 7.00)	37.93 (26.16, 49.70)			
Non-mutated	8/ 12 ( 66.7)	0/ 10 ( 0.0)	39.67 ( 1.86,844.72)	0.002	14.38 ( 0.93,222.06)	66.67 (39.99, 93.34)			
Longest duration of prior ESA treatment								0.998	
< 6 months	25/ 48 ( 52.1)	2/ 20 ( 10.0)	9.78 ( 2.04, 46.87)	0.001	5.21 ( 1.36, 19.94)	42.08 (22.78, 61.39)			
6-12 months	18/ 34 ( 52.9)	1/ 9 ( 11.1)	9.00 ( 1.01, 80.03)	0.026	4.76 ( 0.73, 31.07)	41.83 (15.32, 68.35)			
>12-24 months	22/ 35 ( 62.9)	3/ 21 ( 14.3)	10.15 ( 2.50, 41.23)	<.001	4.40 ( 1.50, 12.93)	48.57 (26.66, 70.49)			
> 24 months	15/ 31 ( 48.4)	2/ 20 ( 10.0)	8.44 ( 1.67, 42.71)	0.005	4.84 ( 1.24, 18.93)	38.39 (16.42, 60.35)			

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 RBC Transfusion Reduction of >=4 Units/8 weeks (mHI-E per IWG 2006) for patients with >=4 Units/8 weeks at BL during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=107) n/ N (%)	Placebo (N=56) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
Age							
<=64 years	8/ 14 ( 57.1)	2/ 12 ( 16.7)	6.67 ( 1.05, 42.43)	0.038	3.43 ( 0.89, 13.15)	40.48 ( 7.06, 73.89)	0.503
65-74 years	26/ 57 ( 45.6)	4/ 20 ( 20.0)	3.35 ( 1.00, 11.29)	0.045	2.28 ( 0.91, 5.73)	25.61 ( 3.83, 47.40)	
>=75 years	18/ 36 ( 50.0)	2/ 24 ( 8.3)	11.00 ( 2.25, 53.84)	<.001	6.00 ( 1.53, 23.53)	41.67 (21.94, 61.39)	
Gender							
female	18/ 38 ( 47.4)	2/ 18 ( 11.1)	7.20 ( 1.45, 35.73)	0.009	4.26 ( 1.11, 16.43)	36.26 (14.74, 57.77)	0.693
male	34/ 69 ( 49.3)	6/ 38 ( 15.8)	5.18 ( 1.92, 13.96)	<.001	3.12 ( 1.44, 6.76)	33.49 (16.95, 50.03)	
Race							
White	33/ 70 ( 47.1)	6/ 37 ( 16.2)	4.61 ( 1.71, 12.43)	0.002	2.91 ( 1.34, 6.30)	30.93 (14.26, 47.59)	0.531
Not Collected or Reported	19/ 36 ( 52.8)	2/ 18 ( 11.1)	8.94 ( 1.79, 44.69)	0.003	4.75 ( 1.24, 18.19)	41.67 (19.83, 63.50)	
Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE	
Region							
North America	5/ 14 ( 35.7)	1/ 9 ( 11.1)	4.44 ( 0.42, 46.55)	0.200	3.21 ( 0.45, 23.21)	24.60 (-7.82, 57.03)	0.960
EMEA	47/ 93 ( 50.5)	7/ 47 ( 14.9)	5.84 ( 2.37, 14.36)	<.001	3.39 ( 1.66, 6.92)	35.64 (21.26, 50.03)	
MDS WHO classification							
MDS RARS	2/ 3 ( 66.7)	0/ 2 ( 0.0)	8.33 ( 0.22,320.38)	0.182	3.75 ( 0.27, 52.64)	66.67 (13.32,100.00)	0.911
MDS RCMD	49/103 ( 47.6)	8/ 54 ( 14.8)	5.22 ( 2.24, 12.14)	<.001	3.21 ( 1.64, 6.28)	32.76 (19.24, 46.28)	
Other	1/ 1 (100.0)	0	NE	NE	NE	NE	
Time since initial diagnosis							
<= 2 years	11/ 26 ( 42.3)	0/ 9 ( 0.0)	14.10 ( 0.74,267.81)	0.020	8.52 ( 0.55,131.52)	42.31 (23.32, 61.30)	0.572
>2 to 5 years	24/ 42 ( 57.1)	4/ 26 ( 15.4)	7.33 ( 2.15, 25.04)	<.001	3.71 ( 1.45, 9.50)	41.76 (21.35, 62.16)	
> 5 years	17/ 39 ( 43.6)	4/ 21 ( 19.0)	3.28 ( 0.93, 11.57)	0.059	2.29 ( 0.88, 5.92)	24.54 ( 1.65, 47.44)	
ECOG Performance Status							
0	17/ 39 ( 43.6)	2/ 23 ( 8.7)	8.11 ( 1.67, 39.49)	0.004	5.01 ( 1.27, 19.76)	34.89 (15.53, 54.25)	0.712
1	31/ 61 ( 50.8)	5/ 26 ( 19.2)	4.34 ( 1.45, 13.00)	0.006	2.64 ( 1.16, 6.03)	31.59 (11.92, 51.26)	
2	4/ 7 ( 57.1)	1/ 7 ( 14.3)	8.00 ( 0.60,106.94)	0.107	4.00 ( 0.58, 27.41)	42.86 (-2.04, 87.76)	
RBC Transfusions/8 weeks over 16 weeks							
< 6 units	16/ 41 ( 39.0)	1/ 23 ( 4.3)	14.08 ( 1.72,114.97)	0.003	8.98 ( 1.27, 63.38)	34.68 (17.58, 51.78)	0.216
>= 6 units	36/ 66 ( 54.5)	7/ 33 ( 21.2)	4.46 ( 1.70, 11.70)	0.002	2.57 ( 1.29, 5.14)	33.33 (14.93, 51.74)	
RBC Transfusions/8 weeks over 16 weeks II							
>= 4 and < 6 units	16/ 41 ( 39.0)	1/ 23 ( 4.3)	14.08 ( 1.72,114.97)	0.003	8.98 ( 1.27, 63.38)	34.68 (17.58, 51.78)	0.216
>= 6 units	36/ 66 ( 54.5)	7/ 33 ( 21.2)	4.46 ( 1.70, 11.70)	0.002	2.57 ( 1.29, 5.14)	33.33 (14.93, 51.74)	
IPSS-R risk							
Very Low	9/ 13 ( 69.2)	2/ 5 ( 40.0)	3.38 ( 0.40, 28.74)	0.268	1.73 ( 0.56, 5.37)	29.23 (-20.5, 78.96)	0.437
Low	33/ 76 ( 43.4)	5/ 41 ( 12.2)	5.53 ( 1.95, 15.63)	<.001	3.56 ( 1.51, 8.42)	31.23 (16.24, 46.21)	
Intermediate	10/ 17 ( 58.8)	1/ 10 ( 10.0)	12.86 ( 1.31,125.78)	0.014	5.88 ( 0.88, 39.39)	48.82 (18.94, 78.71)	
High	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
Platelet count							
< 100 x 10 <sup>9</sup> /L	2/ 6 ( 33.3)	2/ 5 ( 40.0)	0.75 ( 0.06, 8.83)	0.827	0.83 ( 0.18, 3.96)	-6.67 (-63.8, 50.49)	0.161
100-400 x 10 <sup>9</sup> /L	45/ 95 ( 47.4)	5/ 45 ( 11.1)	7.20 ( 2.61, 19.83)	<.001	4.26 ( 1.82, 10.01)	36.26 (22.65, 49.86)	
> 400 x 10 <sup>9</sup> /L	5/ 6 ( 83.3)	1/ 6 ( 16.7)	25.00 ( 1.20,520.73)	0.027	5.00 ( 0.81, 31.00)	66.67 (24.49,100.00)	
Renal status: Creatinine Clearance							
40 - <60 mL/min	14/ 28 ( 50.0)	1/ 12 ( 8.3)	11.00 ( 1.25, 97.02)	0.014	6.00 ( 0.89, 40.63)	41.67 (17.43, 65.91)	0.684
>= 60 mL/min	37/ 76 ( 48.7)	7/ 43 ( 16.3)	4.88 ( 1.93, 12.32)	<.001	2.99 ( 1.46, 6.12)	32.41 (16.66, 48.15)	
< 40 mL/min	1/ 3 ( 33.3)	0/ 1 ( 0.0)	1.80 ( 0.04, 79.42)	0.564	1.50 ( 0.10, 22.62)	33.33 (-20.0, 86.68)	
Serum EPO							
< 100 U/L	24/ 38 ( 63.2)	4/ 23 ( 17.4)	8.14 ( 2.30, 28.81)	<.001	3.63 ( 1.44, 9.14)	45.77 (23.97, 67.57)	0.983
100- < 200 U/L	9/ 25 ( 36.0)	1/ 12 ( 8.3)	6.19 ( 0.68, 56.07)	0.080	4.32 ( 0.62, 30.30)	27.67 ( 3.20, 52.13)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

RBC Transfusion Reduction of  $\geq 4$  Units/8 weeks (mHI-E per IWG 2006) for patients with  $\geq 4$  Units/8 weeks at BL during week 1-24 - Subgroup analysis

ITT Population

Subgroup Level	Luspatercept (N=107)		Placebo (N=56)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/	N (%)	n/	N (%)					
200-500 U/L	12/	26 ( 46.2)	2/	12 ( 16.7)	4.29 ( 0.78, 23.52)	0.084	2.77 ( 0.73, 10.49)	29.49 ( 1.00, 57.98)	
> 500 U/L	7/	18 ( 38.9)	1/	9 ( 11.1)	5.09 ( 0.52, 50.00)	0.144	3.50 ( 0.50, 24.27)	27.78 (-2.70, 58.25)	
SF3B1 mutation status									0.499
Mutated	48/	100 ( 48.0)	8/	50 ( 16.0)	4.85 ( 2.07, 11.36)	<.001	3.00 ( 1.54, 5.85)	32.00 (17.89, 46.11)	
Non-mutated	4/	7 ( 57.1)	0/	6 ( 0.0)	16.71 ( 0.68,409.09)	0.033	7.88 ( 0.51,121.96)	57.14 (20.48, 93.80)	
Longest duration of prior ESA treatment									0.879
< 6 months	17/	34 ( 50.0)	2/	13 ( 15.4)	5.50 ( 1.06, 28.63)	0.032	3.25 ( 0.87, 12.15)	34.62 ( 8.79, 60.44)	
6-12 months	11/	23 ( 47.8)	1/	7 ( 14.3)	5.50 ( 0.57, 53.21)	0.119	3.35 ( 0.52, 21.59)	33.54 ( 0.54, 66.54)	
>12-24 months	14/	24 ( 58.3)	3/	15 ( 20.0)	5.60 ( 1.25, 25.17)	0.020	2.92 ( 1.00, 8.48)	38.33 (10.07, 66.60)	
> 24 months	9/	21 ( 42.9)	1/	17 ( 5.9)	12.00 ( 1.33,108.02)	0.011	7.29 ( 1.02, 51.96)	36.97 (13.04, 60.91)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Mean hemoglobin increase of >=1.5 g/dL for 8 weeks (mHI-E per IWG 2006) for patients with <4 Units/8 weeks at BL in the absence of transfusions during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=46)		Placebo (N=20)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/	N (%)	n/	N (%)					
Age									
<=64 years	10/	15 ( 66.7)	0/	4 ( 0.0)	17.18 ( 0.78,380.84)	0.021	6.56 ( 0.46, 93.20)	66.67 (42.81, 90.52)	0.993
65-74 years	11/	15 ( 73.3)	1/	9 ( 11.1)	22.00 ( 2.05,236.05)	0.004	6.60 ( 1.01, 42.95)	62.22 (31.85, 92.59)	
>=75 years	8/	16 ( 50.0)	0/	7 ( 0.0)	15.00 ( 0.73,306.28)	0.023	8.00 ( 0.52,122.12)	50.00 (25.50, 74.50)	
Gender									
female	13/	21 ( 61.9)	0/	8 ( 0.0)	27.00 ( 1.37,530.76)	0.003	11.05 ( 0.73,166.68)	61.90 (41.13, 82.67)	0.828
male	16/	25 ( 64.0)	1/	12 ( 8.3)	19.56 ( 2.16,177.20)	0.002	7.68 ( 1.15, 51.32)	55.67 (31.20, 80.13)	
Race									
Black or African American	1/	1 (100.0)	0		NE	NE	NE	NE	0.992
White	23/	37 ( 62.2)	1/	14 ( 7.1)	21.36 ( 2.51,181.47)	<.001	8.70 ( 1.29, 58.50)	55.02 (34.37, 75.66)	
Not Collected or Reported	5/	8 ( 62.5)	0/	6 ( 0.0)	20.43 ( 0.86,487.97)	0.020	8.56 ( 0.56,130.05)	62.50 (28.95, 96.05)	
Region									
North America	11/	17 ( 64.7)	1/	10 ( 10.0)	16.50 ( 1.67,163.42)	0.007	6.47 ( 0.98, 42.93)	54.71 (25.35, 84.06)	0.648
EMEA	18/	29 ( 62.1)	0/	10 ( 0.0)	33.78 ( 1.80,633.27)	<.001	13.57 ( 0.89,206.51)	62.07 (44.41, 79.73)	
MDS WHO classification									
MDS RARS	2/	4 ( 50.0)	0		NE	NE	NE	NE	NE
MDS RCMD	27/	42 ( 64.3)	1/	20 ( 5.0)	34.20 ( 4.16,281.42)	<.001	12.86 ( 1.88, 88.01)	59.29 (41.93, 76.64)	
Time since initial diagnosis									
<= 2 years	5/	14 ( 35.7)	0/	10 ( 0.0)	12.16 ( 0.59,250.42)	0.038	8.07 ( 0.50,131.14)	35.71 (10.61, 60.81)	0.075
>2 to 5 years	17/	20 ( 85.0)	0/	8 ( 0.0)	85.00 ( 3.93,1839.1)	<.001	15.00 ( 1.01,223.31)	85.00 (69.35,100.00)	
> 5 years	7/	12 ( 58.3)	1/	2 ( 50.0)	1.40 ( 0.07, 28.12)	0.832	1.17 ( 0.27, 5.05)	8.33 (-66.4, 83.03)	
ECOG Performance Status									
0	9/	15 ( 60.0)	0/	10 ( 0.0)	30.69 ( 1.52,621.02)	0.003	13.06 ( 0.84,201.94)	60.00 (35.21, 84.79)	0.609
1	19/	30 ( 63.3)	0/	6 ( 0.0)	22.04 ( 1.13,428.52)	0.005	8.81 ( 0.60,129.04)	63.33 (46.09, 80.58)	
2	1/	1 (100.0)	1/	4 ( 25.0)	7.00 ( 0.17,291.34)	0.221	4.00 ( 0.73, 21.84)	75.00 (32.57,100.00)	
RBC Transfusions/8 weeks over 16 weeks									
< 6 units	29/	46 ( 63.0)	1/	20 ( 5.0)	32.41 ( 3.98,264.18)	<.001	12.61 ( 1.84, 86.27)	58.04 (41.14, 74.95)	NE
RBC Transfusions/8 weeks over 16 weeks II									
< 4 units	29/	46 ( 63.0)	1/	20 ( 5.0)	32.41 ( 3.98,264.18)	<.001	12.61 ( 1.84, 86.27)	58.04 (41.14, 74.95)	NE
IPSS-R risk									
Very Low	3/	5 ( 60.0)	0/	1 ( 0.0)	4.20 ( 0.12,151.97)	0.317	2.33 ( 0.19, 28.25)	60.00 (17.06,100.00)	0.125
Low	22/	33 ( 66.7)	0/	16 ( 0.0)	64.57 ( 3.55,1175.5)	<.001	22.50 ( 1.45,348.91)	66.67 (50.58, 82.75)	
Intermediate	4/	8 ( 50.0)	1/	3 ( 33.3)	2.00 ( 0.13, 31.98)	0.637	1.50 ( 0.26, 8.58)	16.67 (-46.9, 80.27)	
Platelet count									
< 100 x 10 <sup>9</sup> /L	0/	2 ( 0.0)	0/	1 ( 0.0)	NE	NE	NE	NE	0.796
100-400 x 10 <sup>9</sup> /L	20/	33 ( 60.6)	1/	16 ( 6.3)	23.08 ( 2.71,196.39)	<.001	9.70 ( 1.43, 65.98)	54.36 (33.90, 74.82)	
> 400 x 10 <sup>9</sup> /L	9/	11 ( 81.8)	0/	3 ( 0.0)	26.60 ( 1.01,702.94)	0.012	6.33 ( 0.47, 86.04)	81.82 (59.03,100.00)	
Renal status: Creatinine Clearance									
40 - <60 mL/min	3/	5 ( 60.0)	0		NE	NE	NE	NE	NE
>= 60 mL/min	23/	37 ( 62.2)	1/	19 ( 5.3)	29.57 ( 3.55,246.47)	<.001	11.81 ( 1.72, 80.90)	56.90 (38.32, 75.47)	
< 40 mL/min	3/	4 ( 75.0)	0		NE	NE	NE	NE	
Serum EPO									
< 100 U/L	9/	13 ( 69.2)	0/	8 ( 0.0)	35.89 ( 1.67,769.09)	0.002	12.21 ( 0.81,185.00)	69.23 (44.14, 94.32)	0.910
100- < 200 U/L	7/	12 ( 58.3)	1/	7 ( 14.3)	8.40 ( 0.76, 93.34)	0.068	4.08 ( 0.63, 26.67)	44.05 ( 5.97, 82.13)	
200-500 U/L	10/	17 ( 58.8)	0/	3 ( 0.0)	9.80 ( 0.44,219.25)	0.067	4.67 ( 0.34, 64.23)	58.82 (35.43, 82.22)	
> 500 U/L	2/	3 ( 66.7)	0/	2 ( 0.0)	8.33 ( 0.22,320.38)	0.182	3.75 ( 0.27, 52.64)	66.67 (13.32,100.00)	
SF3B1 mutation status									0.905

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Mean hemoglobin increase of >=1.5 g/dL for 8 weeks (mHI-E per IWG 2006) for patients with <4 Units/8 weeks at BL in the absence of transfusions during week 1-24 - Subgroup analysis

ITT Population

Subgroup Level	Luspatercept (N=46)		Placebo (N=20)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/	N (%)	n/	N (%)					
Mutated	25/	41 ( 61.0)	1/	15 ( 6.7)	21.88 ( 2.62,182.90)	<.001	9.15 ( 1.36, 61.72)	54.31 (34.76, 73.86)	
Non-mutated	4/	5 ( 80.0)	0/	4 ( 0.0)	27.00 ( 0.85,856.53)	0.024	7.50 ( 0.52,108.28)	80.00 (44.94,100.00)	
Longest duration of prior ESA treatment									0.589
< 6 months	8/	14 ( 57.1)	0/	7 ( 0.0)	19.62 ( 0.94,409.82)	0.013	9.07 ( 0.60,137.64)	57.14 (31.22, 83.07)	
6-12 months	7/	11 ( 63.6)	0/	2 ( 0.0)	8.33 ( 0.32,215.68)	0.111	3.75 ( 0.29, 48.90)	63.64 (35.21, 92.06)	
>12-24 months	8/	11 ( 72.7)	0/	6 ( 0.0)	31.57 ( 1.37,725.23)	0.005	9.92 ( 0.67,146.90)	72.73 (46.41, 99.05)	
> 24 months	6/	10 ( 60.0)	1/	3 ( 33.3)	3.00 ( 0.20, 45.24)	0.435	1.80 ( 0.34, 9.64)	26.67 (-34.7, 88.05)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with least one hospitalization during week 1-24 - Subgroup analysis  
 ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									
<=64 years	1/ 29 ( 3.4)	3/ 16 ( 18.8)	0.15 ( 0.01, 1.63)	0.088	0.18 ( 0.02, 1.63)	-15.3 (-35.5, 4.94)	0.245		
65-74 years	17/ 72 ( 23.6)	5/ 29 ( 17.2)	1.48 ( 0.49, 4.49)	0.485	1.37 ( 0.56, 3.37)	6.37 (-10.5, 23.26)			
>=75 years	16/ 52 ( 30.8)	9/ 31 ( 29.0)	1.09 ( 0.41, 2.88)	0.868	1.06 ( 0.53, 2.10)	1.74 (-18.6, 22.05)			
Gender									
female	13/ 59 ( 22.0)	6/ 26 ( 23.1)	0.94 ( 0.31, 2.83)	0.916	0.95 ( 0.41, 2.23)	-1.04 (-20.4, 18.30)	0.910		
male	21/ 94 ( 22.3)	11/ 50 ( 22.0)	1.02 ( 0.45, 2.33)	0.963	1.02 ( 0.53, 1.93)	0.34 (-13.9, 14.58)			
Race									
Black or African American	0/ 1 ( 0.0)	0	NE	NE	NE	NE	0.828		
White	22/107 ( 20.6)	10/ 51 ( 19.6)	1.06 ( 0.46, 2.45)	0.890	1.05 ( 0.54, 2.05)	0.95 (-12.4, 14.27)			
Not Collected or Reported	12/ 44 ( 27.3)	7/ 24 ( 29.2)	0.91 ( 0.30, 2.74)	0.869	0.94 ( 0.43, 2.06)	-1.89 (-24.3, 20.55)			
Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE			
Region									
North America	3/ 31 ( 9.7)	4/ 19 ( 21.1)	0.40 ( 0.08, 2.04)	0.265	0.46 ( 0.12, 1.83)	-11.4 (-32.5, 9.70)	0.246		
EMEA	31/122 ( 25.4)	13/ 57 ( 22.8)	1.15 ( 0.55, 2.42)	0.707	1.11 ( 0.63, 1.96)	2.60 (-10.8, 15.96)			
MDS WHO classification									
MDS RARS	2/ 7 ( 28.6)	0/ 2 ( 0.0)	2.27 ( 0.08, 67.05)	0.419	1.88 ( 0.12, 28.78)	28.57 (-4.89, 62.04)	0.620		
MDS RCMD	31/145 ( 21.4)	17/ 74 ( 23.0)	0.91 ( 0.47, 1.78)	0.788	0.93 ( 0.55, 1.57)	-1.59 (-13.3, 10.08)			
Other	1/ 1 (100.0)	0	NE	NE	NE	NE			
Time since initial diagnosis									
<= 2 years	11/ 40 ( 27.5)	2/ 19 ( 10.5)	3.22 ( 0.64, 16.31)	0.145	2.61 ( 0.64, 10.64)	16.97 (-2.57, 36.52)	0.195		
>2 to 5 years	12/ 62 ( 19.4)	7/ 34 ( 20.6)	0.93 ( 0.33, 2.63)	0.885	0.94 ( 0.41, 2.16)	-1.23 (-18.0, 15.54)			
> 5 years	11/ 51 ( 21.6)	8/ 23 ( 34.8)	0.52 ( 0.17, 1.53)	0.232	0.62 ( 0.29, 1.33)	-13.2 (-35.7, 9.29)			
ECOG Performance Status									
0	10/ 54 ( 18.5)	4/ 33 ( 12.1)	1.65 ( 0.47, 5.76)	0.433	1.53 ( 0.52, 4.48)	6.40 (-8.81, 21.61)	0.509		
1	22/ 91 ( 24.2)	8/ 32 ( 25.0)	0.96 ( 0.38, 2.43)	0.926	0.97 ( 0.48, 1.95)	-0.82 (-18.2, 16.57)			
2	2/ 8 ( 25.0)	5/ 11 ( 45.5)	0.40 ( 0.05, 2.93)	0.374	0.55 ( 0.14, 2.15)	-20.5 (-62.5, 21.57)			
RBC Transfusions/8 weeks over 16 weeks									
< 6 units	14/ 87 ( 16.1)	6/ 43 ( 14.0)	1.18 ( 0.42, 3.33)	0.751	1.15 ( 0.48, 2.79)	2.14 (-10.8, 15.06)	0.661		
>= 6 units	20/ 66 ( 30.3)	11/ 33 ( 33.3)	0.87 ( 0.36, 2.13)	0.760	0.91 ( 0.50, 1.67)	-3.03 (-22.6, 16.50)			
RBC Transfusions/8 weeks over 16 weeks II									
< 4 units	4/ 46 ( 8.7)	4/ 20 ( 20.0)	0.38 ( 0.08, 1.71)	0.199	0.43 ( 0.12, 1.57)	-11.3 (-30.6, 8.02)	0.155		
>= 4 and < 6 units	10/ 41 ( 24.4)	2/ 23 ( 8.7)	3.39 ( 0.67, 17.05)	0.126	2.80 ( 0.67, 11.72)	15.69 (-1.78, 33.17)			
>= 6 units	20/ 66 ( 30.3)	11/ 33 ( 33.3)	0.87 ( 0.36, 2.13)	0.760	0.91 ( 0.50, 1.67)	-3.03 (-22.6, 16.50)			
IPSS-R risk									
Very Low	7/ 18 ( 38.9)	2/ 6 ( 33.3)	1.27 ( 0.18, 8.89)	0.812	1.17 ( 0.33, 4.16)	5.56 (-38.4, 49.49)	0.655		
Low	24/109 ( 22.0)	12/ 57 ( 21.1)	1.06 ( 0.48, 2.31)	0.886	1.05 ( 0.57, 1.93)	0.97 (-12.2, 14.10)			
Intermediate	3/ 25 ( 12.0)	3/ 13 ( 23.1)	0.45 ( 0.08, 2.66)	0.381	0.52 ( 0.12, 2.22)	-11.1 (-37.3, 15.13)			
High	0/ 1 ( 0.0)	0	NE	NE	NE	NE			
Platelet count									
< 100 x 10 <sup>9</sup> /L	2/ 8 ( 25.0)	1/ 6 ( 16.7)	1.67 ( 0.11, 24.26)	0.717	1.50 ( 0.17, 12.94)	8.33 (-34.0, 50.64)	0.818		
100-400 x 10 <sup>9</sup> /L	29/128 ( 22.7)	15/ 61 ( 24.6)	0.90 ( 0.44, 1.84)	0.769	0.92 ( 0.53, 1.59)	-1.93 (-14.9, 11.08)			
> 400 x 10 <sup>9</sup> /L	3/ 17 ( 17.6)	1/ 9 ( 11.1)	1.71 ( 0.15, 19.36)	0.667	1.59 ( 0.19, 13.15)	6.54 (-20.8, 33.92)			
Renal status: Creatinine Clearance									
40 - <60 mL/min	10/ 33 ( 30.3)	3/ 12 ( 25.0)	1.30 ( 0.29, 5.86)	0.731	1.21 ( 0.40, 3.67)	5.30 (-23.8, 34.39)	0.126		
>= 60 mL/min	22/113 ( 19.5)	13/ 62 ( 21.0)	0.91 ( 0.42, 1.97)	0.813	0.93 ( 0.50, 1.71)	-1.50 (-14.0, 10.99)			
< 40 mL/min	2/ 7 ( 28.6)	1/ 1 (100.0)	0.15 ( 0.00, 5.18)	0.197	0.29 ( 0.09, 0.92)	-71.4 (-100,-37.96)			
Serum EPO								0.179	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with least one hospitalization during week 1-24 - Subgroup analysis

ITT Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
< 100 U/L	10/ 51 ( 19.6)	6/ 31 ( 19.4)	1.02 ( 0.33, 3.14)	0.978	1.01 ( 0.41, 2.51)	0.25 (-17.4, 17.92)			
100- < 200 U/L	11/ 37 ( 29.7)	2/ 19 ( 10.5)	3.60 ( 0.71, 18.28)	0.110	2.82 ( 0.70, 11.47)	19.20 (-0.98, 39.39)			
200-500 U/L	9/ 43 ( 20.9)	6/ 15 ( 40.0)	0.40 ( 0.11, 1.41)	0.150	0.52 ( 0.22, 1.22)	-19.1 (-46.7, 8.54)			
> 500 U/L	3/ 21 ( 14.3)	3/ 11 ( 27.3)	0.44 ( 0.07, 2.70)	0.379	0.52 ( 0.13, 2.18)	-13.0 (-43.3, 17.29)			
SF3B1 mutation status								0.350	
Mutated	31/141 ( 22.0)	16/ 65 ( 24.6)	0.86 ( 0.43, 1.72)	0.677	0.89 ( 0.53, 1.51)	-2.63 (-15.1, 9.88)			
Non-mutated	3/ 12 ( 25.0)	1/ 10 ( 10.0)	3.00 ( 0.26, 34.57)	0.375	2.50 ( 0.31, 20.45)	15.00 (-15.8, 45.76)			
Longest duration of prior ESA treatment								0.161	
< 6 months	14/ 48 ( 29.2)	3/ 20 ( 15.0)	2.33 ( 0.59, 9.24)	0.222	1.94 ( 0.63, 6.03)	14.17 (-6.09, 34.42)			
6-12 months	6/ 34 ( 17.6)	3/ 9 ( 33.3)	0.43 ( 0.08, 2.22)	0.309	0.53 ( 0.16, 1.71)	-15.7 (-49.0, 17.67)			
>12-24 months	10/ 35 ( 28.6)	6/ 21 ( 28.6)	1.00 ( 0.30, 3.31)	1.000	1.00 ( 0.43, 2.35)	0.00 (-24.4, 24.44)			
> 24 months	1/ 31 ( 3.2)	4/ 20 ( 20.0)	0.13 ( 0.01, 1.30)	0.051	0.16 ( 0.02, 1.34)	-16.8 (-35.4, 1.83)			

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one Adverse Event during week 1-24 - Subgroup analysis  
 Safety Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									0.639
<=64 years	26/ 29 ( 89.7)	13/ 16 ( 81.3)			2.00 ( 0.35, 11.32)	0.432	1.10 ( 0.85, 1.44)	8.41 (-13.7, 30.51)	
65-74 years	69/ 72 ( 95.8)	28/ 29 ( 96.6)			0.82 ( 0.08, 8.24)	0.868	0.99 ( 0.91, 1.08)	-0.72 (-8.81, 7.37)	
>=75 years	50/ 52 ( 96.2)	29/ 31 ( 93.5)			1.72 ( 0.23, 12.90)	0.594	1.03 ( 0.92, 1.14)	2.61 (-7.50, 12.71)	
Gender									0.722
female	57/ 59 ( 96.6)	24/ 26 ( 92.3)			2.38 ( 0.32, 17.85)	0.391	1.05 ( 0.93, 1.18)	4.30 (-6.93, 15.54)	
male	88/ 94 ( 93.6)	46/ 50 ( 92.0)			1.28 ( 0.34, 4.75)	0.717	1.02 ( 0.92, 1.12)	1.62 (-7.38, 10.62)	
Race									0.486
Black or African American	1/ 1 (100.0)	0			NE	NE	NE	NE	
White	101/107 ( 94.4)	46/ 51 ( 90.2)			1.83 ( 0.53, 6.30)	0.334	1.05 ( 0.95, 1.16)	4.20 (-5.06, 13.45)	
Not Collected or Reported	42/ 44 ( 95.5)	23/ 24 ( 95.8)			0.91 ( 0.08, 10.62)	0.942	1.00 ( 0.90, 1.11)	-0.38 (-10.5, 9.71)	
Other	1/ 1 (100.0)	1/ 1 (100.0)			NE	NE	NE	NE	
Region									0.654
North America	31/ 31 (100.0)	18/ 19 ( 94.7)			5.11 ( 0.20, 131.96)	0.201	1.06 ( 0.95, 1.17)	5.26 (-4.78, 15.30)	
EMEA	114/122 ( 93.4)	52/ 57 ( 91.2)			1.37 ( 0.43, 4.39)	0.596	1.02 ( 0.93, 1.12)	2.21 (-6.34, 10.77)	
MDS WHO classification									0.301
MDS RARS	7/ 7 (100.0)	1/ 2 ( 50.0)			15.00 ( 0.39, 576.69)	0.061	2.00 ( 0.50, 8.00)	50.00 (-19.3, 100.00)	
MDS RCMD	137/145 ( 94.5)	69/ 74 ( 93.2)			1.24 ( 0.39, 3.94)	0.714	1.01 ( 0.94, 1.09)	1.24 (-5.58, 8.06)	
Other	1/ 1 (100.0)	0			NE	NE	NE	NE	
Time since initial diagnosis									0.924
<= 2 years	36/ 40 ( 90.0)	17/ 19 ( 89.5)			1.06 ( 0.18, 6.36)	0.951	1.01 ( 0.84, 1.21)	0.53 (-16.1, 17.17)	
>2 to 5 years	61/ 62 ( 98.4)	32/ 34 ( 94.1)			3.81 ( 0.33, 43.67)	0.253	1.05 ( 0.96, 1.14)	4.27 (-4.24, 12.78)	
> 5 years	48/ 51 ( 94.1)	21/ 23 ( 91.3)			1.52 ( 0.24, 9.80)	0.658	1.03 ( 0.89, 1.19)	2.81 (-10.4, 16.02)	
ECOG Performance Status									0.758
0	50/ 54 ( 92.6)	30/ 33 ( 90.9)			1.25 ( 0.26, 5.97)	0.781	1.02 ( 0.89, 1.16)	1.68 (-10.4, 13.72)	
1	87/ 91 ( 95.6)	30/ 32 ( 93.8)			1.45 ( 0.25, 8.32)	0.677	1.02 ( 0.92, 1.13)	1.85 (-7.53, 11.24)	
2	8/ 8 (100.0)	10/ 11 ( 90.9)			2.43 ( 0.09, 67.57)	0.394	1.10 ( 0.91, 1.33)	9.09 (-7.90, 26.08)	
RBC Transfusions/8 weeks over 16 weeks									0.225
< 6 units	82/ 87 ( 94.3)	41/ 43 ( 95.3)			0.80 ( 0.15, 4.30)	0.795	0.99 ( 0.91, 1.08)	-1.10 (-9.07, 6.88)	
>= 6 units	63/ 66 ( 95.5)	29/ 33 ( 87.9)			2.90 ( 0.61, 13.79)	0.168	1.09 ( 0.95, 1.25)	7.58 (-4.64, 19.79)	
RBC Transfusions/8 weeks over 16 weeks II									0.057
< 4 units	42/ 46 ( 91.3)	18/ 20 ( 90.0)			1.17 ( 0.20, 6.95)	0.866	1.01 ( 0.85, 1.20)	1.30 (-14.2, 16.77)	
>= 4 and < 6 units	40/ 41 ( 97.6)	23/ 23 (100.0)			0.57 ( 0.02, 14.68)	0.454	0.98 ( 0.93, 1.02)	-2.44 (-7.16, 2.28)	
>= 6 units	63/ 66 ( 95.5)	29/ 33 ( 87.9)			2.90 ( 0.61, 13.79)	0.168	1.09 ( 0.95, 1.25)	7.58 (-4.64, 19.79)	
IPSS-R risk									0.984
Very Low	15/ 18 ( 83.3)	5/ 6 ( 83.3)			1.00 ( 0.08, 11.93)	1.000	1.00 ( 0.66, 1.51)	0.00 (-34.4, 34.43)	
Low	105/109 ( 96.3)	53/ 57 ( 93.0)			1.98 ( 0.48, 8.23)	0.340	1.04 ( 0.96, 1.12)	3.35 (-4.16, 10.86)	
Intermediate	24/ 25 ( 96.0)	12/ 13 ( 92.3)			2.00 ( 0.11, 34.82)	0.633	1.04 ( 0.87, 1.24)	3.69 (-12.7, 20.09)	
High	1/ 1 (100.0)	0			NE	NE	NE	NE	
Platelet count									0.200
< 100 x 10 <sup>9</sup> /L	7/ 8 ( 87.5)	6/ 6 (100.0)			0.38 ( 0.01, 11.17)	0.386	0.88 ( 0.67, 1.14)	-12.5 (-35.4, 10.42)	
100-400 x 10 <sup>9</sup> /L	121/128 ( 94.5)	55/ 61 ( 90.2)			1.89 ( 0.61, 5.87)	0.269	1.05 ( 0.96, 1.15)	4.37 (-4.08, 12.82)	
> 400 x 10 <sup>9</sup> /L	17/ 17 (100.0)	9/ 9 (100.0)			NE	NE	NE	NE	
Renal status: Creatinine Clearance									0.962
40 - <60 mL/min	31/ 33 ( 93.9)	11/ 12 ( 91.7)			1.41 ( 0.12, 17.12)	0.789	1.02 ( 0.85, 1.24)	2.27 (-15.4, 19.90)	
>= 60 mL/min	107/113 ( 94.7)	57/ 62 ( 91.9)			1.56 ( 0.46, 5.35)	0.474	1.03 ( 0.95, 1.12)	2.75 (-5.18, 10.69)	
< 40 mL/min	7/ 7 (100.0)	1/ 1 (100.0)			NE	NE	NE	NE	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one Adverse Event during week 1-24 - Subgroup analysis

Safety Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Serum EPO									0.015
< 100 U/L	46/ 51 ( 90.2)		31/ 31 (100.0)		0.13 ( 0.01, 2.51)	0.074	0.90 ( 0.82, 0.99)	-9.80 (-18.0, -1.64)	
100- < 200 U/L	37/ 37 (100.0)		18/ 19 ( 94.7)		6.08 ( 0.24,156.63)	0.163	1.06 ( 0.95, 1.17)	5.26 (-4.78, 15.30)	
200-500 U/L	41/ 43 ( 95.3)		12/ 15 ( 80.0)		5.13 ( 0.77, 34.31)	0.071	1.19 ( 0.92, 1.55)	15.35 (-5.85, 36.55)	
> 500 U/L	20/ 21 ( 95.2)		9/ 11 ( 81.8)		4.44 ( 0.36, 55.58)	0.223	1.16 ( 0.87, 1.56)	13.42 (-11.1, 37.97)	
SF3B1 mutation status									0.459
Mutated	133/141 ( 94.3)		60/ 65 ( 92.3)		1.39 ( 0.44, 4.41)	0.581	1.02 ( 0.94, 1.11)	2.02 (-5.50, 9.54)	
Non-mutated	12/ 12 (100.0)		9/ 10 ( 90.0)		3.95 ( 0.14,108.09)	0.273	1.11 ( 0.90, 1.37)	10.00 (-8.59, 28.59)	
Longest duration of prior ESA treatment									0.047
< 6 months	48/ 48 (100.0)		18/ 20 ( 90.0)		13.11 ( 0.60,286.14)	0.027	1.11 ( 0.96, 1.29)	10.00 (-3.15, 23.15)	
6-12 months	30/ 34 ( 88.2)		9/ 9 (100.0)		0.36 ( 0.02, 7.25)	0.286	0.88 ( 0.78, 1.00)	-11.8 (-22.6, -0.93)	
>12-24 months	34/ 35 ( 97.1)		19/ 21 ( 90.5)		3.58 ( 0.30, 42.11)	0.288	1.07 ( 0.92, 1.25)	6.67 (-7.05, 20.38)	
> 24 months	28/ 31 ( 90.3)		19/ 20 ( 95.0)		0.49 ( 0.05, 5.08)	0.548	0.95 ( 0.82, 1.11)	-4.68 (-18.8, 9.45)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one Serious Adverse Event during week 1-24 - Subgroup analysis  
 Safety Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/	N (%)	n/	N (%)					
Age									0.934
<=64 years	3/	29 (10.3)	1/	16 (6.3)	1.73 (0.16, 18.16)	0.648	1.66 (0.19, 14.63)	4.09 (-12.1, 20.33)	
65-74 years	17/	72 (23.6)	6/	29 (20.7)	1.18 (0.41, 3.39)	0.753	1.14 (0.50, 2.60)	2.92 (-14.8, 20.63)	
>=75 years	20/	52 (38.5)	9/	31 (29.0)	1.53 (0.59, 3.97)	0.386	1.32 (0.69, 2.53)	9.43 (-11.3, 30.17)	
Gender									0.248
female	16/	59 (27.1)	8/	26 (30.8)	0.84 (0.30, 2.30)	0.732	0.88 (0.43, 1.80)	-3.65 (-24.7, 17.41)	
male	24/	94 (25.5)	8/	50 (16.0)	1.80 (0.74, 4.37)	0.192	1.60 (0.77, 3.29)	9.53 (-3.92, 22.98)	
Race									0.273
Black or African American	0/	1 (0.0)	0		NE	NE	NE	NE	
White	29/	107 (27.1)	9/	51 (17.6)	1.74 (0.75, 4.01)	0.195	1.54 (0.79, 3.00)	9.46 (-3.98, 22.89)	
Not Collected or Reported	11/	44 (25.0)	7/	24 (29.2)	0.81 (0.27, 2.47)	0.712	0.86 (0.38, 1.92)	-4.17 (-26.4, 18.07)	
Other	0/	1 (0.0)	0/	1 (0.0)	NE	NE	NE	NE	
Region									0.573
North America	6/	31 (19.4)	4/	19 (21.1)	0.90 (0.22, 3.72)	0.885	0.92 (0.30, 2.84)	-1.70 (-24.7, 21.31)	
EMEA	34/	122 (27.9)	12/	57 (21.1)	1.45 (0.68, 3.07)	0.332	1.32 (0.74, 2.36)	6.82 (-6.42, 20.06)	
MDS WHO classification									0.743
MDS RARS	2/	7 (28.6)	0/	2 (0.0)	2.27 (0.08, 67.05)	0.419	1.88 (0.12, 28.78)	28.57 (-4.89, 62.04)	
MDS RCMD	37/	145 (25.5)	16/	74 (21.6)	1.24 (0.64, 2.42)	0.525	1.18 (0.70, 1.98)	3.90 (-7.87, 15.66)	
Other	1/	1 (100.0)	0		NE	NE	NE	NE	
Time since initial diagnosis									0.213
<= 2 years	13/	40 (32.5)	2/	19 (10.5)	4.09 (0.82, 20.43)	0.073	3.09 (0.77, 12.33)	21.97 (1.95, 42.00)	
>2 to 5 years	11/	62 (17.7)	8/	34 (23.5)	0.70 (0.25, 1.96)	0.498	0.75 (0.34, 1.69)	-5.79 (-22.9, 11.35)	
> 5 years	16/	51 (31.4)	6/	23 (26.1)	1.30 (0.43, 3.90)	0.647	1.20 (0.54, 2.67)	5.29 (-16.7, 27.29)	
ECOG Performance Status									0.299
0	13/	54 (24.1)	3/	33 (9.1)	3.17 (0.83, 12.12)	0.082	2.65 (0.82, 8.60)	14.98 (-0.06, 30.02)	
1	24/	91 (26.4)	8/	32 (25.0)	1.07 (0.43, 2.71)	0.879	1.05 (0.53, 2.11)	1.37 (-16.1, 18.90)	
2	3/	8 (37.5)	5/	11 (45.5)	0.72 (0.11, 4.62)	0.736	0.83 (0.27, 2.49)	-7.95 (-52.6, 36.67)	
RBC Transfusions/8 weeks over 16 weeks									0.452
< 6 units	19/	87 (21.8)	6/	43 (14.0)	1.72 (0.63, 4.69)	0.285	1.57 (0.67, 3.63)	7.89 (-5.63, 21.40)	
>= 6 units	21/	66 (31.8)	10/	33 (30.3)	1.07 (0.43, 2.65)	0.879	1.05 (0.56, 1.96)	1.52 (-17.8, 20.81)	
RBC Transfusions/8 weeks over 16 weeks II									0.226
< 4 units	7/	46 (15.2)	4/	20 (20.0)	0.72 (0.18, 2.80)	0.634	0.76 (0.25, 2.31)	-4.78 (-25.2, 15.59)	
>= 4 and < 6 units	12/	41 (29.3)	2/	23 (8.7)	4.34 (0.88, 21.49)	0.058	3.37 (0.82, 13.75)	20.57 (2.50, 38.64)	
>= 6 units	21/	66 (31.8)	10/	33 (30.3)	1.07 (0.43, 2.65)	0.879	1.05 (0.56, 1.96)	1.52 (-17.8, 20.81)	
IPSS-R risk									0.329
Very Low	8/	18 (44.4)	2/	6 (33.3)	1.60 (0.23, 11.08)	0.640	1.33 (0.38, 4.63)	11.11 (-33.0, 55.27)	
Low	28/	109 (25.7)	10/	57 (17.5)	1.62 (0.73, 3.64)	0.237	1.46 (0.77, 2.80)	8.14 (-4.69, 20.98)	
Intermediate	4/	25 (16.0)	4/	13 (30.8)	0.43 (0.09, 2.10)	0.296	0.52 (0.15, 1.75)	-14.8 (-43.7, 14.14)	
High	0/	1 (0.0)	0		NE	NE	NE	NE	
Platelet count									0.822
< 100 x 10 <sup>9</sup> /L	3/	8 (37.5)	1/	6 (16.7)	3.00 (0.23, 39.61)	0.411	2.25 (0.30, 16.63)	20.83 (-24.1, 65.72)	
100-400 x 10 <sup>9</sup> /L	33/	128 (25.8)	13/	61 (21.3)	1.28 (0.62, 2.66)	0.504	1.21 (0.69, 2.13)	4.47 (-8.30, 17.24)	
> 400 x 10 <sup>9</sup> /L	4/	17 (23.5)	2/	9 (22.2)	1.08 (0.16, 7.42)	0.941	1.06 (0.24, 4.71)	1.31 (-32.5, 35.13)	
Renal status: Creatinine Clearance									0.039
40 - <60 mL/min	12/	33 (36.4)	2/	12 (16.7)	2.86 (0.53, 15.26)	0.212	2.18 (0.57, 8.36)	19.70 (-7.02, 46.42)	
>= 60 mL/min	26/	113 (23.0)	13/	62 (21.0)	1.13 (0.53, 2.39)	0.757	1.10 (0.61, 1.98)	2.04 (-10.7, 14.80)	
< 40 mL/min	2/	7 (28.6)	1/	1 (100.0)	0.15 (0.00, 5.18)	0.197	0.29 (0.09, 0.92)	-71.4 (-100, -37.96)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

## Dossier zur Nutzenbewertung - Modul 4 B

### Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one Serious Adverse Event during week 1-24 - Subgroup analysis

Safety Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/	N (%)	n/	N (%)					
<b>Serum EPO</b>									
< 100 U/L	12/	51 ( 23.5)	7/	31 ( 22.6)	1.05 ( 0.36, 3.05)	0.922	1.04 ( 0.46, 2.36)	0.95 (-17.8, 19.71)	0.347
100- < 200 U/L	13/	37 ( 35.1)	2/	19 ( 10.5)	4.60 ( 0.92, 23.11)	0.051	3.34 ( 0.84, 13.30)	24.61 ( 3.94, 45.27)	
200-500 U/L	10/	43 ( 23.3)	4/	15 ( 26.7)	0.83 ( 0.22, 3.20)	0.792	0.87 ( 0.32, 2.37)	-3.41 (-29.1, 22.28)	
> 500 U/L	4/	21 ( 19.0)	3/	11 ( 27.3)	0.63 ( 0.11, 3.49)	0.599	0.70 ( 0.19, 2.58)	-8.23 (-39.4, 23.00)	
<b>SF3B1 mutation status</b>									
Mutated	36/	141 ( 25.5)	15/	65 ( 23.1)	1.14 ( 0.57, 2.28)	0.705	1.11 ( 0.65, 1.87)	2.45 (-10.1, 14.97)	0.299
Non-mutated	4/	12 ( 33.3)	1/	10 ( 10.0)	4.50 ( 0.41, 49.08)	0.204	3.33 ( 0.44, 25.23)	23.33 (-9.18, 55.85)	
<b>Longest duration of prior ESA treatment</b>									
< 6 months	15/	48 ( 31.3)	4/	20 ( 20.0)	1.82 ( 0.52, 6.37)	0.350	1.56 ( 0.59, 4.13)	11.25 (-10.6, 33.14)	0.284
6-12 months	7/	34 ( 20.6)	3/	9 ( 33.3)	0.52 ( 0.10, 2.61)	0.426	0.62 ( 0.20, 1.92)	-12.7 (-46.4, 20.92)	
>12-24 months	12/	35 ( 34.3)	4/	21 ( 19.0)	2.22 ( 0.61, 8.08)	0.226	1.80 ( 0.67, 4.86)	15.24 (-7.77, 38.25)	
> 24 months	3/	31 ( 9.7)	4/	20 ( 20.0)	0.43 ( 0.08, 2.16)	0.300	0.48 ( 0.12, 1.94)	-10.3 (-30.7, 10.06)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one Severe Adverse Event (CTCAE >=3) during week 1-24 - Subgroup analysis  
 Safety Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									0.887
<=64 years	5/ 29 ( 17.2)	3/ 16 ( 18.8)			0.90 ( 0.19, 4.39)	0.900	0.92 ( 0.25, 3.36)	-1.51 (-25.1, 22.05)	
65-74 years	23/ 72 ( 31.9)	10/ 29 ( 34.5)			0.89 ( 0.36, 2.22)	0.807	0.93 ( 0.51, 1.70)	-2.54 (-22.9, 17.84)	
>=75 years	26/ 52 ( 50.0)	14/ 31 ( 45.2)			1.21 ( 0.50, 2.96)	0.671	1.11 ( 0.69, 1.78)	4.84 (-17.3, 27.01)	
Gender									0.934
female	22/ 59 ( 37.3)	10/ 26 ( 38.5)			0.95 ( 0.37, 2.46)	0.919	0.97 ( 0.54, 1.75)	-1.17 (-23.6, 21.23)	
male	32/ 94 ( 34.0)	17/ 50 ( 34.0)			1.00 ( 0.49, 2.07)	0.996	1.00 ( 0.62, 1.61)	0.04 (-16.2, 16.30)	
Race									0.303
Black or African American	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
White	39/107 ( 36.4)	15/ 51 ( 29.4)			1.38 ( 0.67, 2.83)	0.385	1.24 ( 0.76, 2.03)	7.04 (-8.44, 22.51)	
Not Collected or Reported	15/ 44 ( 34.1)	11/ 24 ( 45.8)			0.61 ( 0.22, 1.69)	0.345	0.74 ( 0.41, 1.35)	-11.7 (-36.1, 12.62)	
Other	0/ 1 ( 0.0)	1/ 1 (100.0)			0.11 ( 0.00, 10.27)	0.317	0.33 ( 0.03, 4.19)	-100 ( -100,-100.0)	
Region									0.723
North America	10/ 31 ( 32.3)	7/ 19 ( 36.8)			0.82 ( 0.25, 2.71)	0.742	0.88 ( 0.40, 1.91)	-4.58 (-31.8, 22.64)	
EMEA	44/122 ( 36.1)	20/ 57 ( 35.1)			1.04 ( 0.54, 2.01)	0.899	1.03 ( 0.67, 1.57)	0.98 (-14.1, 16.01)	
MDS WHO classification									0.450
MDS RARS	3/ 7 ( 42.9)	0/ 2 ( 0.0)			3.89 ( 0.14,109.99)	0.285	2.63 ( 0.19, 37.14)	42.86 ( 6.20, 79.52)	
MDS RCMD	50/145 ( 34.5)	27/ 74 ( 36.5)			0.92 ( 0.51, 1.64)	0.769	0.95 ( 0.65, 1.38)	-2.00 (-15.4, 11.42)	
Other	1/ 1 (100.0)	0			NE	NE	NE	NE	
Time since initial diagnosis									0.739
<= 2 years	16/ 40 ( 40.0)	6/ 19 ( 31.6)			1.44 ( 0.45, 4.59)	0.535	1.27 ( 0.59, 2.72)	8.42 (-17.4, 34.25)	
>2 to 5 years	19/ 62 ( 30.6)	12/ 34 ( 35.3)			0.81 ( 0.33, 1.97)	0.643	0.87 ( 0.48, 1.57)	-4.65 (-24.4, 15.09)	
> 5 years	19/ 51 ( 37.3)	9/ 23 ( 39.1)			0.92 ( 0.34, 2.54)	0.878	0.95 ( 0.51, 1.77)	-1.88 (-25.8, 22.08)	
ECOG Performance Status									0.348
0	16/ 54 ( 29.6)	6/ 33 ( 18.2)			1.89 ( 0.66, 5.47)	0.236	1.63 ( 0.71, 3.75)	11.45 (-6.48, 29.38)	
1	34/ 91 ( 37.4)	14/ 32 ( 43.8)			0.77 ( 0.34, 1.74)	0.526	0.85 ( 0.53, 1.37)	-6.39 (-26.2, 13.47)	
2	4/ 8 ( 50.0)	7/ 11 ( 63.6)			0.57 ( 0.09, 3.64)	0.563	0.79 ( 0.34, 1.79)	-13.6 (-58.5, 31.18)	
RBC Transfusions/8 weeks over 16 weeks									0.868
< 6 units	29/ 87 ( 33.3)	14/ 43 ( 32.6)			1.04 ( 0.48, 2.26)	0.930	1.02 ( 0.61, 1.73)	0.78 (-16.4, 17.93)	
>= 6 units	25/ 66 ( 37.9)	13/ 33 ( 39.4)			0.94 ( 0.40, 2.21)	0.884	0.96 ( 0.57, 1.62)	-1.52 (-21.9, 18.85)	
RBC Transfusions/8 weeks over 16 weeks II									0.056
< 4 units	9/ 46 ( 19.6)	8/ 20 ( 40.0)			0.36 ( 0.12, 1.16)	0.083	0.49 ( 0.22, 1.08)	-20.4 (-44.8, 3.90)	
>= 4 and < 6 units	20/ 41 ( 48.8)	6/ 23 ( 26.1)			2.70 ( 0.89, 8.22)	0.078	1.87 ( 0.88, 3.98)	22.69 (-0.89, 46.28)	
>= 6 units	25/ 66 ( 37.9)	13/ 33 ( 39.4)			0.94 ( 0.40, 2.21)	0.884	0.96 ( 0.57, 1.62)	-1.52 (-21.9, 18.85)	
IPSS-R risk									0.756
Very Low	9/ 18 ( 50.0)	2/ 6 ( 33.3)			2.00 ( 0.29, 13.81)	0.487	1.50 ( 0.44, 5.09)	16.67 (-27.6, 60.90)	
Low	37/109 ( 33.9)	21/ 57 ( 36.8)			0.88 ( 0.45, 1.72)	0.711	0.92 ( 0.60, 1.42)	-2.90 (-18.3, 12.46)	
Intermediate	8/ 25 ( 32.0)	4/ 13 ( 30.8)			1.06 ( 0.25, 4.50)	0.939	1.04 ( 0.38, 2.81)	1.23 (-29.8, 32.28)	
High	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Platelet count									0.494
< 100 x 10 <sup>9</sup> /L	4/ 8 ( 50.0)	1/ 6 ( 16.7)			5.00 ( 0.39, 64.39)	0.215	3.00 ( 0.44, 20.44)	33.33 (-12.4, 79.05)	
100-400 x 10 <sup>9</sup> /L	45/128 ( 35.2)	23/ 61 ( 37.7)			0.90 ( 0.48, 1.69)	0.734	0.93 ( 0.63, 1.39)	-2.55 (-17.3, 12.16)	
> 400 x 10 <sup>9</sup> /L	5/ 17 ( 29.4)	3/ 9 ( 33.3)			0.83 ( 0.15, 4.72)	0.840	0.88 ( 0.27, 2.88)	-3.92 (-41.6, 33.73)	
Renal status: Creatinine Clearance									0.092
40 - <60 mL/min	15/ 33 ( 45.5)	4/ 12 ( 33.3)			1.67 ( 0.42, 6.64)	0.472	1.36 ( 0.56, 3.30)	12.12 (-19.5, 43.74)	
>= 60 mL/min	37/113 ( 32.7)	22/ 62 ( 35.5)			0.89 ( 0.46, 1.70)	0.715	0.92 ( 0.60, 1.41)	-2.74 (-17.5, 11.98)	
< 40 mL/min	2/ 7 ( 28.6)	1/ 1 (100.0)			0.15 ( 0.00, 5.18)	0.197	0.29 ( 0.09, 0.92)	-71.4 ( -100,-37.96)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one Severe Adverse Event (CTCAE >=3) during week 1-24 - Subgroup analysis

Safety Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/	N (%)	n/	N (%)					
<b>Serum EPO</b>									
< 100 U/L	18/	51 ( 35.3)	11/	31 ( 35.5)	0.99 ( 0.39, 2.52)	0.986	0.99 ( 0.54, 1.82)	-0.19 (-21.5, 21.16)	0.997
100- < 200 U/L	13/	37 ( 35.1)	7/	19 ( 36.8)	0.93 ( 0.29, 2.94)	0.900	0.95 ( 0.46, 1.99)	-1.71 (-28.3, 24.88)	
200-500 U/L	15/	43 ( 34.9)	5/	15 ( 33.3)	1.07 ( 0.31, 3.71)	0.914	1.05 ( 0.46, 2.39)	1.55 (-26.2, 29.34)	
> 500 U/L	7/	21 ( 33.3)	4/	11 ( 36.4)	0.88 ( 0.19, 4.03)	0.866	0.92 ( 0.34, 2.46)	-3.03 (-37.9, 31.82)	
<b>SF3B1 mutation status</b>									
Mutated	50/	141 ( 35.5)	26/	65 ( 40.0)	0.82 ( 0.45, 1.51)	0.531	0.89 ( 0.61, 1.29)	-4.54 (-18.8, 9.75)	0.202
Non-mutated	4/	12 ( 33.3)	1/	10 ( 10.0)	4.50 ( 0.41, 49.08)	0.204	3.33 ( 0.44, 25.23)	23.33 (-9.18, 55.85)	
<b>Longest duration of prior ESA treatment</b>									
< 6 months	15/	48 ( 31.3)	6/	20 ( 30.0)	1.06 ( 0.34, 3.30)	0.920	1.04 ( 0.47, 2.30)	1.25 (-22.7, 25.24)	0.547
6-12 months	12/	34 ( 35.3)	4/	9 ( 44.4)	0.68 ( 0.15, 3.03)	0.618	0.79 ( 0.34, 1.88)	-9.15 (-45.4, 27.07)	
>12-24 months	16/	35 ( 45.7)	8/	21 ( 38.1)	1.37 ( 0.45, 4.13)	0.580	1.20 ( 0.62, 2.31)	7.62 (-18.9, 34.15)	
> 24 months	7/	31 ( 22.6)	8/	20 ( 40.0)	0.44 ( 0.13, 1.49)	0.187	0.56 ( 0.24, 1.31)	-17.4 (-43.5, 8.61)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one Adverse Event leading to permanent discontinuation of study drug during week 1-24 - Subgroup analysis  
 Safety Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/ N (%)	n/ N (%)	n/ N (%)	n/ N (%)					
Age									0.245
<=64 years	0/ 29 ( 0.0)	1/ 16 ( 6.3)			0.18 ( 0.01, 4.56)	0.178	0.19 ( 0.01, 4.39)	-6.25 (-18.1, 5.61)	
65-74 years	5/ 72 ( 6.9)	2/ 29 ( 6.9)			1.01 ( 0.18, 5.51)	0.993	1.01 ( 0.21, 4.90)	0.05 (-10.9, 10.98)	
>=75 years	7/ 52 ( 13.5)	1/ 31 ( 3.2)			4.67 ( 0.55, 39.89)	0.129	4.17 ( 0.54, 32.34)	10.24 (-0.93, 21.40)	
Gender									0.184
female	5/ 59 ( 8.5)	3/ 26 ( 11.5)			0.71 ( 0.16, 3.22)	0.658	0.73 ( 0.19, 2.85)	-3.06 (-17.3, 11.12)	
male	7/ 94 ( 7.4)	1/ 50 ( 2.0)			3.94 ( 0.47, 32.99)	0.176	3.72 ( 0.47, 29.42)	5.45 (-1.13, 12.02)	
Race									0.177
Black or African American	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
White	8/107 ( 7.5)	1/ 51 ( 2.0)			4.04 ( 0.49, 33.21)	0.163	3.81 ( 0.49, 29.68)	5.52 (-0.75, 11.79)	
Not Collected or Reported	4/ 44 ( 9.1)	3/ 24 ( 12.5)			0.70 ( 0.14, 3.42)	0.661	0.73 ( 0.18, 2.98)	-3.41 (-19.1, 12.31)	
Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)			NE	NE	NE	NE	
Region									0.824
North America	1/ 31 ( 3.2)	0/ 19 ( 0.0)			1.92 ( 0.07, 49.50)	0.434	1.88 ( 0.08, 43.82)	3.23 (-2.99, 9.45)	
EMEA	11/122 ( 9.0)	4/ 57 ( 7.0)			1.31 ( 0.40, 4.32)	0.654	1.28 ( 0.43, 3.86)	2.00 (-6.36, 10.35)	
MDS WHO classification									0.889
MDS RARS	1/ 7 ( 14.3)	0/ 2 ( 0.0)			1.15 ( 0.03, 38.88)	0.593	1.13 ( 0.06, 20.71)	14.29 (-11.6, 40.21)	
MDS RCMD	11/145 ( 7.6)	4/ 74 ( 5.4)			1.44 ( 0.44, 4.68)	0.547	1.40 ( 0.46, 4.26)	2.18 (-4.54, 8.90)	
Other	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Time since initial diagnosis									0.203
<= 2 years	0/ 40 ( 0.0)	1/ 19 ( 5.3)			0.15 ( 0.01, 3.92)	0.147	0.16 ( 0.01, 3.82)	-5.26 (-15.3, 4.78)	
>2 to 5 years	3/ 62 ( 4.8)	2/ 34 ( 5.9)			0.81 ( 0.13, 5.12)	0.827	0.82 ( 0.14, 4.68)	-1.04 (-10.6, 8.50)	
> 5 years	9/ 51 ( 17.6)	1/ 23 ( 4.3)			4.71 ( 0.56, 39.65)	0.124	4.06 ( 0.55, 30.19)	13.30 (-0.08, 26.68)	
ECOG Performance Status									0.302
0	6/ 54 ( 11.1)	0/ 33 ( 0.0)			8.98 ( 0.49, 164.82)	0.048	8.04 ( 0.47, 138.17)	11.11 ( 2.73, 19.49)	
1	4/ 91 ( 4.4)	2/ 32 ( 6.3)			0.69 ( 0.12, 3.96)	0.677	0.70 ( 0.14, 3.66)	-1.85 (-11.2, 7.53)	
2	2/ 8 ( 25.0)	2/ 11 ( 18.2)			1.50 ( 0.16, 13.75)	0.726	1.38 ( 0.24, 7.80)	6.82 (-30.9, 44.50)	
RBC Transfusions/8 weeks over 16 weeks									0.079
< 6 units	3/ 87 ( 3.4)	3/ 43 ( 7.0)			0.48 ( 0.09, 2.46)	0.369	0.49 ( 0.10, 2.35)	-3.53 (-12.1, 5.00)	
>= 6 units	9/ 66 ( 13.6)	1/ 33 ( 3.0)			5.05 ( 0.61, 41.71)	0.100	4.50 ( 0.59, 34.03)	10.61 ( 0.47, 20.74)	
RBC Transfusions/8 weeks over 16 weeks II									0.102
< 4 units	0/ 46 ( 0.0)	2/ 20 ( 10.0)			0.08 ( 0.00, 1.74)	0.031	0.09 ( 0.00, 1.78)	-10.0 (-23.1, 3.15)	
>= 4 and < 6 units	3/ 41 ( 7.3)	1/ 23 ( 4.3)			1.74 ( 0.17, 17.73)	0.640	1.68 ( 0.19, 15.26)	2.97 (-8.56, 14.50)	
>= 6 units	9/ 66 ( 13.6)	1/ 33 ( 3.0)			5.05 ( 0.61, 41.71)	0.100	4.50 ( 0.59, 34.03)	10.61 ( 0.47, 20.74)	
IPSS-R risk									0.747
Very Low	2/ 18 ( 11.1)	0/ 6 ( 0.0)			1.97 ( 0.08, 46.85)	0.404	1.84 ( 0.10, 33.82)	11.11 (-3.41, 25.63)	
Low	9/109 ( 8.3)	3/ 57 ( 5.3)			1.62 ( 0.42, 6.24)	0.481	1.57 ( 0.44, 5.57)	2.99 (-4.77, 10.76)	
Intermediate	1/ 25 ( 4.0)	1/ 13 ( 7.7)			0.50 ( 0.03, 8.71)	0.633	0.52 ( 0.04, 7.66)	-3.69 (-20.1, 12.70)	
High	0/ 1 ( 0.0)	0			NE	NE	NE	NE	
Platelet count									0.017
< 100 x 10 <sup>9</sup> /L	0/ 8 ( 0.0)	0/ 6 ( 0.0)			NE	NE	NE	NE	
100-400 x 10 <sup>9</sup> /L	12/128 ( 9.4)	1/ 61 ( 1.6)			6.21 ( 0.79, 48.88)	0.050	5.72 ( 0.76, 42.98)	7.74 ( 1.76, 13.71)	
> 400 x 10 <sup>9</sup> /L	0/ 17 ( 0.0)	3/ 9 ( 33.3)			0.05 ( 0.00, 1.17)	0.013	0.08 ( 0.00, 1.39)	-33.3 (-64.1, -2.54)	
Renal status: Creatinine Clearance									0.459
40 - <60 mL/min	4/ 33 ( 12.1)	0/ 12 ( 0.0)			3.81 ( 0.19, 76.27)	0.212	3.44 ( 0.20, 59.55)	12.12 ( 0.99, 23.26)	
>= 60 mL/min	8/113 ( 7.1)	4/ 62 ( 6.5)			1.10 ( 0.32, 3.83)	0.875	1.10 ( 0.34, 3.50)	0.63 (-7.10, 8.36)	
< 40 mL/min	0/ 7 ( 0.0)	0/ 1 ( 0.0)			NE	NE	NE	NE	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one Adverse Event leading to permanent discontinuation of study drug during week 1-24 - Subgroup analysis

Safety Population

Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
	n/	N (%)	n/	N (%)					
Serum EPO									0.413
< 100 U/L	3/	51 ( 5.9)	3/	31 ( 9.7)	0.58 ( 0.11, 3.09)	0.525	0.61 ( 0.13, 2.83)	-3.80 (-16.0, 8.45)	
100- < 200 U/L	6/	37 ( 16.2)	0/	19 ( 0.0)	8.05 ( 0.43,150.92)	0.066	6.84 ( 0.41,115.37)	16.22 ( 4.34, 28.09)	
200-500 U/L	2/	43 ( 4.7)	0/	15 ( 0.0)	1.87 ( 0.08, 41.12)	0.399	1.82 ( 0.09, 35.87)	4.65 (-1.64, 10.95)	
> 500 U/L	1/	21 ( 4.8)	1/	11 ( 9.1)	0.50 ( 0.03, 8.85)	0.636	0.52 ( 0.04, 7.59)	-4.33 (-23.6, 14.95)	
SF3B1 mutation status									0.412
Mutated	10/	141 ( 7.1)	4/	65 ( 6.2)	1.16 ( 0.35, 3.86)	0.804	1.15 ( 0.38, 3.54)	0.94 (-6.28, 8.16)	
Non-mutated	2/	12 ( 16.7)	0/	10 ( 0.0)	5.00 ( 0.21,117.21)	0.186	4.23 ( 0.23, 79.10)	16.67 (-4.42, 37.75)	
Longest duration of prior ESA treatment									0.477
< 6 months	3/	48 ( 6.3)	2/	20 ( 10.0)	0.60 ( 0.09, 3.90)	0.592	0.63 ( 0.11, 3.46)	-3.75 (-18.6, 11.07)	
6-12 months	0/	34 ( 0.0)	0/	9 ( 0.0)	NE	NE	NE	NE	
>12-24 months	5/	35 ( 14.3)	1/	21 ( 4.8)	3.33 ( 0.36, 30.70)	0.269	3.00 ( 0.38, 23.96)	9.52 (-5.22, 24.27)	
> 24 months	3/	31 ( 9.7)	1/	20 ( 5.0)	2.04 ( 0.20, 21.07)	0.548	1.94 ( 0.22, 17.33)	4.68 (-9.45, 18.80)	

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis  
 Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: Eye disorders	Age							0.967
	<=64 years	2/ 29 ( 6.9)	0/ 16 ( 0.0)	3.00 ( 0.14, 66.40)	0.288	2.83 ( 0.14, 55.64)	6.90 (-2.33, 16.12)	
	65-74 years	5/ 72 ( 6.9)	0/ 29 ( 0.0)	4.81 ( 0.26, 89.78)	0.148	4.52 ( 0.26, 79.23)	6.94 ( 1.07, 12.82)	
	>=75 years	5/ 52 ( 9.6)	1/ 31 ( 3.2)	3.19 ( 0.36, 28.67)	0.280	2.98 ( 0.36, 24.35)	6.39 (-3.75, 16.53)	
	Gender							0.212
	female	10/ 59 ( 16.9)	0/ 26 ( 0.0)	11.24 ( 0.63,199.47)	0.026	9.45 ( 0.57,155.47)	16.95 ( 7.38, 26.52)	
	male	2/ 94 ( 2.1)	1/ 50 ( 2.0)	1.07 ( 0.09, 12.04)	0.959	1.06 ( 0.10, 11.45)	0.13 (-4.73, 4.98)	
	Race							0.601
	Black or African American	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	White	6/107 ( 5.6)	1/ 51 ( 2.0)	2.97 ( 0.35, 25.35)	0.299	2.86 ( 0.35, 23.13)	3.65 (-2.14, 9.43)	
	Not Collected or Reported	6/ 44 ( 13.6)	0/ 24 ( 0.0)	8.27 ( 0.45,153.49)	0.060	7.22 ( 0.42,122.95)	13.64 ( 3.50, 23.78)	
	Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE	
	Region							0.357
	North America	3/ 31 ( 9.7)	1/ 19 ( 5.3)	1.93 ( 0.19, 20.01)	0.580	1.84 ( 0.21, 16.43)	4.41 (-10.0, 18.88)	
	EMEA	9/122 ( 7.4)	0/ 57 ( 0.0)	9.63 ( 0.55,168.32)	0.036	8.96 ( 0.53,151.31)	7.38 ( 2.74, 12.02)	
	MDS WHO classification							0.395
	MDS RARS	1/ 7 ( 14.3)	0/ 2 ( 0.0)	1.15 ( 0.03, 38.88)	0.593	1.13 ( 0.06, 20.71)	14.29 (-11.6, 40.21)	
	MDS RCMD	10/145 ( 6.9)	1/ 74 ( 1.4)	5.41 ( 0.68, 43.08)	0.076	5.10 ( 0.67, 39.11)	5.55 ( 0.65, 10.44)	
	Other	1/ 1 (100.0)	0	NE	NE	NE	NE	
	Time since initial diagnosis							0.812
	<= 2 years	4/ 40 ( 10.0)	0/ 19 ( 0.0)	4.81 ( 0.25, 94.00)	0.157	4.39 ( 0.25, 77.62)	10.00 ( 0.70, 19.30)	
	>2 to 5 years	4/ 62 ( 6.5)	0/ 34 ( 0.0)	5.31 ( 0.28,101.60)	0.132	5.00 ( 0.28, 90.18)	6.45 ( 0.34, 12.57)	
	> 5 years	4/ 51 ( 7.8)	1/ 23 ( 4.3)	1.87 ( 0.20, 17.75)	0.582	1.80 ( 0.21, 15.26)	3.50 (-7.64, 14.63)	
	ECOG Performance Status							0.392
	ECOG Performance Status							
	1	10/ 91 ( 11.0)	0/ 32 ( 0.0)	8.37 ( 0.48,147.11)	0.051	7.53 ( 0.45,124.99)	10.99 ( 4.56, 17.41)	
	2	0/ 8 ( 0.0)	1/ 11 ( 9.1)	0.41 ( 0.01, 11.46)	0.394	0.44 ( 0.02, 9.69)	-9.09 (-26.1, 7.90)	
	RBC Transfusions/last 8 weeks							0.680
	< 6 units	7/ 75 ( 9.3)	0/ 30 ( 0.0)	6.68 ( 0.37,120.70)	0.085	6.12 ( 0.36,103.89)	9.33 ( 2.75, 15.92)	
	>= 6 units	5/ 78 ( 6.4)	1/ 46 ( 2.2)	3.08 ( 0.35, 27.24)	0.290	2.95 ( 0.36, 24.47)	4.24 (-2.64, 11.11)	
	RBC Transfusions/last 8 weeks II							0.670
	< 4 units	0/ 28 ( 0.0)	0/ 11 ( 0.0)	NE	NE	NE	NE	
	>= 4 and < 6 units	7/ 47 ( 14.9)	0/ 19 ( 0.0)	7.22 ( 0.39,133.01)	0.077	6.25 ( 0.37,104.31)	14.89 ( 4.72, 25.07)	
	>= 6 units	5/ 78 ( 6.4)	1/ 46 ( 2.2)	3.08 ( 0.35, 27.24)	0.290	2.95 ( 0.36, 24.47)	4.24 (-2.64, 11.11)	
	RBC Transfusions/8 weeks over 16 weeks							0.177
	< 6 units	10/ 87 ( 11.5)	0/ 43 ( 0.0)	11.79 ( 0.67,206.07)	0.021	10.50 ( 0.63,175.08)	11.49 ( 4.79, 18.20)	
	>= 6 units	2/ 66 ( 3.0)	1/ 33 ( 3.0)	1.00 ( 0.09, 11.45)	1.000	1.00 ( 0.09, 10.63)	0.00 (-7.16, 7.16)	
	RBC Transfusions/8 weeks over 16 weeks II							0.486
	< 4 units	3/ 46 ( 6.5)	0/ 20 ( 0.0)	3.30 ( 0.16, 66.88)	0.246	3.13 ( 0.17, 57.89)	6.52 (-0.61, 13.66)	
	>= 4 and < 6 units	7/ 41 ( 17.1)	0/ 23 ( 0.0)	10.22 ( 0.56,187.62)	0.037	8.57 ( 0.51,143.58)	17.07 ( 5.56, 28.59)	
	>= 6 units	2/ 66 ( 3.0)	1/ 33 ( 3.0)	1.00 ( 0.09, 11.45)	1.000	1.00 ( 0.09, 10.63)	0.00 (-7.16, 7.16)	
	IPSS-R risk							0.320
	Very Low	2/ 18 ( 11.1)	1/ 6 ( 16.7)	0.63 ( 0.05, 8.43)	0.727	0.67 ( 0.07, 6.11)	-5.56 (-38.7, 27.61)	
	Low	7/109 ( 6.4)	0/ 57 ( 0.0)	8.41 ( 0.47,150.03)	0.051	7.91 ( 0.46,136.05)	6.42 ( 1.82, 11.02)	
	Intermediate	3/ 25 ( 12.0)	0/ 13 ( 0.0)	4.20 ( 0.20, 87.72)	0.199	3.77 ( 0.21, 67.89)	12.00 (-0.74, 24.74)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).  
 Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis

Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: Eye disorders	High	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	Platelet count							0.760
	< 100 x 10 <sup>9</sup> /L	1/ 8 ( 12.5)	0/ 6 ( 0.0)	2.60 ( 0.09, 75.49)	0.386	2.33 ( 0.11, 48.99)	12.50 (-10.4, 35.42)	
	100-400 x 10 <sup>9</sup> /L	7/128 ( 5.5)	0/ 61 ( 0.0)	7.59 ( 0.43,135.13)	0.063	7.21 ( 0.42,124.22)	5.47 ( 1.53, 9.41)	
	> 400 x 10 <sup>9</sup> /L	4/ 17 ( 23.5)	1/ 9 ( 11.1)	2.46 ( 0.23, 26.11)	0.454	2.12 ( 0.28, 16.24)	12.42 (-16.4, 41.20)	
	Renal status: Creatinine Clearance							0.798
	40 - <60 mL/min	3/ 33 ( 9.1)	0/ 12 ( 0.0)	2.87 ( 0.14, 59.70)	0.285	2.68 ( 0.15, 48.32)	9.09 (-0.72, 18.90)	
	>= 60 mL/min	7/113 ( 6.2)	1/ 62 ( 1.6)	4.03 ( 0.48, 33.52)	0.166	3.84 ( 0.48, 30.51)	4.58 (-0.86, 10.02)	
	< 40 mL/min	2/ 7 ( 28.6)	0/ 1 ( 0.0)	1.36 ( 0.04, 46.65)	0.564	1.25 ( 0.09, 17.02)	28.57 (-4.89, 62.04)	
	Serum EPO							0.462
	< 100 U/L	8/ 51 ( 15.7)	0/ 31 ( 0.0)	12.31 ( 0.68,221.24)	0.021	10.46 ( 0.62,175.16)	15.69 ( 5.71, 25.67)	
	100- < 200 U/L	1/ 37 ( 2.7)	1/ 19 ( 5.3)	0.50 ( 0.03, 8.46)	0.628	0.51 ( 0.03, 7.76)	-2.56 (-13.9, 8.76)	
	200-500 U/L	2/ 43 ( 4.7)	0/ 15 ( 0.0)	1.87 ( 0.08, 41.12)	0.399	1.82 ( 0.09, 35.87)	4.65 (-1.64, 10.95)	
	> 500 U/L	1/ 21 ( 4.8)	0/ 11 ( 0.0)	1.68 ( 0.06, 44.77)	0.469	1.64 ( 0.07, 37.15)	4.76 (-4.35, 13.87)	
	SF3B1 mutation status							NE
	Mutated	12/141 ( 8.5)	1/ 65 ( 1.5)	5.95 ( 0.76, 46.80)	0.056	5.53 ( 0.73, 41.65)	6.97 ( 1.48, 12.46)	
	Non-mutated	0/ 12 ( 0.0)	0/ 10 ( 0.0)	NE	NE	NE	NE	
	Longest duration of prior ESA treatment							0.870
	< 6 months	3/ 48 ( 6.3)	1/ 20 ( 5.0)	1.27 ( 0.12, 12.96)	0.843	1.25 ( 0.14, 11.31)	1.25 (-10.5, 13.00)	
	6-12 months	3/ 34 ( 8.8)	0/ 9 ( 0.0)	2.11 ( 0.10, 44.61)	0.361	2.00 ( 0.11, 35.58)	8.82 (-0.71, 18.36)	
	>12-24 months	3/ 35 ( 8.6)	0/ 21 ( 0.0)	4.63 ( 0.23, 94.22)	0.172	4.28 ( 0.23, 78.95)	8.57 (-0.70, 17.85)	
	> 24 months	3/ 31 ( 9.7)	0/ 20 ( 0.0)	5.04 ( 0.25,102.87)	0.156	4.59 ( 0.25, 84.46)	9.68 (-0.73, 20.08)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis  
 Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: Gastrointestinal disorders								
	Age							0.936
	<=64 years	14/ 29 ( 48.3)	5/ 16 ( 31.3)	2.05 ( 0.57, 7.41)	0.274	1.54 ( 0.68, 3.50)	17.03 (-12.1, 46.12)	
	65-74 years	39/ 72 ( 54.2)	11/ 29 ( 37.9)	1.93 ( 0.80, 4.67)	0.142	1.43 ( 0.86, 2.38)	16.24 (-4.84, 37.31)	
	>=75 years	25/ 52 ( 48.1)	9/ 31 ( 29.0)	2.26 ( 0.88, 5.84)	0.090	1.66 ( 0.89, 3.07)	19.04 (-1.93, 40.01)	
	Gender							0.385
	female	32/ 59 ( 54.2)	11/ 26 ( 42.3)	1.62 ( 0.64, 4.10)	0.314	1.28 ( 0.77, 2.13)	11.93 (-10.9, 34.78)	
	male	46/ 94 ( 48.9)	14/ 50 ( 28.0)	2.46 ( 1.18, 5.15)	0.016	1.75 ( 1.07, 2.85)	20.94 ( 4.90, 36.97)	
	Race							0.852
	Black or African American	1/ 1 (100.0)	0	NE	NE	NE	NE	
	White	48/107 ( 44.9)	14/ 51 ( 27.5)	2.15 ( 1.04, 4.43)	0.037	1.63 ( 1.00, 2.68)	17.41 ( 1.96, 32.86)	
	Not Collected or Reported	28/ 44 ( 63.6)	10/ 24 ( 41.7)	2.45 ( 0.89, 6.78)	0.083	1.53 ( 0.90, 2.58)	21.97 (-2.34, 46.28)	
	Other	1/ 1 (100.0)	1/ 1 (100.0)	NE	NE	NE	NE	
	Region							0.733
	North America	16/ 31 ( 51.6)	7/ 19 ( 36.8)	1.83 ( 0.57, 5.88)	0.314	1.40 ( 0.71, 2.77)	14.77 (-13.2, 42.70)	
	EMEA	62/122 ( 50.8)	18/ 57 ( 31.6)	2.24 ( 1.16, 4.34)	0.016	1.61 ( 1.06, 2.45)	19.24 ( 4.26, 34.22)	
	MDS WHO classification							0.475
	MDS RARS	3/ 7 ( 42.9)	1/ 2 ( 50.0)	0.75 ( 0.03, 17.51)	0.866	0.86 ( 0.17, 4.37)	-7.14 (-85.5, 71.25)	
	MDS RCMD	74/145 ( 51.0)	24/ 74 ( 32.4)	2.17 ( 1.21, 3.90)	0.009	1.57 ( 1.09, 2.27)	18.60 ( 5.19, 32.02)	
	Other	1/ 1 (100.0)	0	NE	NE	NE	NE	
	Time since initial diagnosis							0.422
	<= 2 years	16/ 40 ( 40.0)	6/ 19 ( 31.6)	1.44 ( 0.45, 4.59)	0.535	1.27 ( 0.59, 2.72)	8.42 (-17.4, 34.25)	
	>2 to 5 years	34/ 62 ( 54.8)	9/ 34 ( 26.5)	3.37 ( 1.36, 8.39)	0.008	2.07 ( 1.13, 3.79)	28.37 ( 9.05, 47.69)	
	> 5 years	28/ 51 ( 54.9)	10/ 23 ( 43.5)	1.58 ( 0.59, 4.27)	0.366	1.26 ( 0.74, 2.14)	11.42 (-13.0, 35.86)	
	ECOG Performance Status	28/ 54 ( 51.9)	12/ 33 ( 36.4)	1.88 ( 0.78, 4.58)	0.162	1.43 ( 0.85, 2.40)	15.49 (-5.65, 36.63)	
	ECOG Performance Status							0.646
	1	47/ 91 ( 51.6)	9/ 32 ( 28.1)	2.73 ( 1.14, 6.54)	0.022	1.84 ( 1.02, 3.31)	23.52 ( 4.87, 42.18)	
	2	3/ 8 ( 37.5)	4/ 11 ( 36.4)	1.05 ( 0.16, 6.92)	0.961	1.03 ( 0.31, 3.38)	1.14 (-42.8, 45.11)	
	RBC Transfusions/last 8 weeks							0.386
	< 6 units	35/ 75 ( 46.7)	7/ 30 ( 23.3)	2.88 ( 1.10, 7.51)	0.028	2.00 ( 1.00, 4.00)	23.33 ( 4.45, 42.22)	
	>= 6 units	43/ 78 ( 55.1)	18/ 46 ( 39.1)	1.91 ( 0.91, 4.01)	0.086	1.41 ( 0.93, 2.13)	16.00 (-1.91, 33.91)	
	RBC Transfusions/last 8 weeks II							0.485
	< 4 units	10/ 28 ( 35.7)	3/ 11 ( 27.3)	1.48 ( 0.32, 6.88)	0.619	1.31 ( 0.44, 3.88)	8.44 (-23.3, 40.19)	
	>= 4 and < 6 units	25/ 47 ( 53.2)	4/ 19 ( 21.1)	4.26 ( 1.23, 14.77)	0.018	2.53 ( 1.02, 6.28)	32.14 ( 8.91, 55.37)	
	>= 6 units	43/ 78 ( 55.1)	18/ 46 ( 39.1)	1.91 ( 0.91, 4.01)	0.086	1.41 ( 0.93, 2.13)	16.00 (-1.91, 33.91)	
	RBC Transfusions/8 weeks over 16 weeks							0.227
	< 6 units	43/ 87 ( 49.4)	11/ 43 ( 25.6)	2.84 ( 1.27, 6.35)	0.010	1.93 ( 1.11, 3.36)	23.84 ( 7.10, 40.59)	
	>= 6 units	35/ 66 ( 53.0)	14/ 33 ( 42.4)	1.53 ( 0.66, 3.56)	0.322	1.25 ( 0.79, 1.98)	10.61 (-10.1, 31.33)	
	RBC Transfusions/8 weeks over 16 weeks II							0.394
	< 4 units	19/ 46 ( 41.3)	5/ 20 ( 25.0)	2.11 ( 0.66, 6.80)	0.209	1.65 ( 0.72, 3.80)	16.30 (-7.41, 40.02)	
	>= 4 and < 6 units	24/ 41 ( 58.5)	6/ 23 ( 26.1)	4.00 ( 1.31, 12.25)	0.013	2.24 ( 1.08, 4.68)	32.45 ( 9.01, 55.89)	
	>= 6 units	35/ 66 ( 53.0)	14/ 33 ( 42.4)	1.53 ( 0.66, 3.56)	0.322	1.25 ( 0.79, 1.98)	10.61 (-10.1, 31.33)	
	IPSS-R risk							0.189
	Very Low	10/ 18 ( 55.6)	4/ 6 ( 66.7)	0.63 ( 0.09, 4.33)	0.640	0.83 ( 0.41, 1.68)	-11.1 (-55.3, 33.04)	
	Low	57/109 ( 52.3)	17/ 57 ( 29.8)	2.58 ( 1.31, 5.09)	0.006	1.75 ( 1.13, 2.71)	22.47 ( 7.34, 37.60)	
	Intermediate	10/ 25 ( 40.0)	4/ 13 ( 30.8)	1.50 ( 0.36, 6.23)	0.581	1.30 ( 0.50, 3.35)	9.23 (-22.4, 40.83)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).  
 Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis

Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
		n/	N (%)	n/	N (%)					
SOC: Gastrointestinal disorders	High	1/	1 (100.0)	0		NE	NE	NE	NE	
	Platelet count									0.742
	< 100 x 10 <sup>9</sup> /L	4/	8 ( 50.0)	1/	6 ( 16.7)	5.00 ( 0.39, 64.39)	0.215	3.00 ( 0.44, 20.44)	33.33 (-12.4, 79.05)	
	100-400 x 10 <sup>9</sup> /L	64/	128 ( 50.0)	20/	61 ( 32.8)	2.05 ( 1.08, 3.88)	0.026	1.53 ( 1.02, 2.27)	17.21 ( 2.59, 31.84)	
	> 400 x 10 <sup>9</sup> /L	10/	17 ( 58.8)	4/	9 ( 44.4)	1.79 ( 0.35, 9.13)	0.493	1.32 ( 0.58, 3.04)	14.38 (-25.6, 54.39)	
	Renal status: Creatinine Clearance									0.010
	40 - <60 mL/min	16/	33 ( 48.5)	3/	12 ( 25.0)	2.82 ( 0.65, 12.33)	0.163	1.94 ( 0.68, 5.49)	23.48 (-6.36, 53.33)	
	>= 60 mL/min	58/	113 ( 51.3)	21/	62 ( 33.9)	2.06 ( 1.08, 3.91)	0.027	1.52 ( 1.02, 2.24)	17.46 ( 2.50, 32.41)	
	< 40 mL/min	4/	7 ( 57.1)	1/	1 (100.0)	0.43 ( 0.01, 14.08)	0.439	0.57 ( 0.30, 1.09)	-42.9 (-79.5, -6.20)	
	Serum EPO									0.308
	< 100 U/L	23/	51 ( 45.1)	12/	31 ( 38.7)	1.30 ( 0.52, 3.23)	0.573	1.17 ( 0.68, 1.99)	6.39 (-15.5, 28.31)	
	100- < 200 U/L	22/	37 ( 59.5)	4/	19 ( 21.1)	5.50 ( 1.52, 19.85)	0.007	2.82 ( 1.14, 7.02)	38.41 (14.19, 62.62)	
	200-500 U/L	22/	43 ( 51.2)	4/	15 ( 26.7)	2.88 ( 0.79, 10.48)	0.103	1.92 ( 0.79, 4.67)	24.50 (-2.41, 51.40)	
	> 500 U/L	11/	21 ( 52.4)	5/	11 ( 45.5)	1.32 ( 0.31, 5.70)	0.714	1.15 ( 0.54, 2.48)	6.93 (-29.4, 43.29)	
	SF3B1 mutation status									0.457
	Mutated	70/	141 ( 49.6)	22/	65 ( 33.8)	1.93 ( 1.05, 3.55)	0.034	1.47 ( 1.00, 2.14)	15.80 ( 1.64, 29.96)	
	Non-mutated	8/	12 ( 66.7)	3/	10 ( 30.0)	4.67 ( 0.77, 28.47)	0.094	2.22 ( 0.80, 6.21)	36.67 (-2.30, 75.63)	
	Longest duration of prior ESA treatment									0.313
	< 6 months	29/	48 ( 60.4)	7/	20 ( 35.0)	2.83 ( 0.96, 8.40)	0.058	1.73 ( 0.91, 3.27)	25.42 ( 0.35, 50.48)	
	6-12 months	20/	34 ( 58.8)	1/	9 ( 11.1)	11.43 ( 1.28, 101.93)	0.012	5.29 ( 0.82, 34.32)	47.71 (21.35, 74.08)	
	>12-24 months	16/	35 ( 45.7)	9/	21 ( 42.9)	1.12 ( 0.38, 3.34)	0.837	1.07 ( 0.58, 1.97)	2.86 (-24.0, 29.70)	
	> 24 months	11/	31 ( 35.5)	6/	20 ( 30.0)	1.28 ( 0.38, 4.29)	0.688	1.18 ( 0.52, 2.69)	5.48 (-20.7, 31.70)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis  
 Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: General disorders and administration site conditions, PT: Asthenia	Age							0.903
	<=64 years	1/ 29 ( 3.4)	0/ 16 ( 0.0)	1.74 ( 0.07, 45.13)	0.458	1.70 ( 0.07, 39.47)	3.45 (-3.19, 10.09)	
	65-74 years	14/ 72 ( 19.4)	2/ 29 ( 6.9)	3.26 ( 0.69, 15.36)	0.120	2.82 ( 0.68, 11.64)	12.55 (-0.44, 25.53)	
	>=75 years	13/ 52 ( 25.0)	4/ 31 ( 12.9)	2.25 ( 0.66, 7.65)	0.189	1.94 ( 0.69, 5.42)	12.10 (-4.57, 28.76)	
	Gender							0.686
	female	13/ 59 ( 22.0)	2/ 26 ( 7.7)	3.39 ( 0.71, 16.27)	0.112	2.86 ( 0.70, 11.80)	14.34 (-0.38, 29.06)	
	male	15/ 94 ( 16.0)	4/ 50 ( 8.0)	2.18 ( 0.68, 6.97)	0.181	1.99 ( 0.70, 5.69)	7.96 (-2.59, 18.51)	
	Race							0.200
	Black or African American	1/ 1 (100.0)	0	NE	NE	NE	NE	
	White	21/107 ( 19.6)	3/ 51 ( 5.9)	3.91 ( 1.11, 13.78)	0.025	3.34 ( 1.04, 10.67)	13.74 ( 3.83, 23.66)	
	Not Collected or Reported	6/ 44 ( 13.6)	3/ 24 ( 12.5)	1.11 ( 0.25, 4.88)	0.896	1.09 ( 0.30, 3.98)	1.14 (-15.5, 17.81)	
	Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE	
	Region							0.598
	North America	3/ 31 ( 9.7)	0/ 19 ( 0.0)	4.79 ( 0.23, 98.01)	0.166	4.37 ( 0.24, 80.31)	9.68 (-0.73, 20.08)	
	EMEA	25/122 ( 20.5)	6/ 57 ( 10.5)	2.19 ( 0.84, 5.68)	0.102	1.95 ( 0.85, 4.48)	9.97 (-0.75, 20.68)	
	MDS WHO classification							0.910
	MDS RARS	2/ 7 ( 28.6)	0/ 2 ( 0.0)	2.27 ( 0.08, 67.05)	0.419	1.88 ( 0.12, 28.78)	28.57 (-4.89, 62.04)	
	MDS RCMD	26/145 ( 17.9)	6/ 74 ( 8.1)	2.48 ( 0.97, 6.32)	0.052	2.21 ( 0.95, 5.13)	9.82 ( 1.01, 18.64)	
	Other	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	Time since initial diagnosis							0.763
	<= 2 years	3/ 40 ( 7.5)	0/ 19 ( 0.0)	3.64 ( 0.18, 74.09)	0.224	3.41 ( 0.19, 62.97)	7.50 (-0.66, 15.66)	
	>2 to 5 years	12/ 62 ( 19.4)	4/ 34 ( 11.8)	1.80 ( 0.53, 6.09)	0.342	1.65 ( 0.57, 4.71)	7.59 (-7.04, 22.22)	
	> 5 years	13/ 51 ( 25.5)	2/ 23 ( 8.7)	3.59 ( 0.74, 17.46)	0.099	2.93 ( 0.72, 11.95)	16.79 ( 0.19, 33.40)	
	ECOG Performance Status							0.782
	ECOG Performance Status	9/ 54 ( 16.7)	3/ 33 ( 9.1)	2.00 ( 0.50, 8.00)	0.323	1.83 ( 0.53, 6.29)	7.58 (-6.39, 21.54)	
	1	16/ 91 ( 17.6)	2/ 32 ( 6.3)	3.20 ( 0.69, 14.77)	0.120	2.81 ( 0.68, 11.57)	11.33 (-0.14, 22.80)	
	2	3/ 8 ( 37.5)	1/ 11 ( 9.1)	6.00 ( 0.49, 73.45)	0.144	4.13 ( 0.52, 32.75)	28.41 (-9.19, 66.01)	
	RBC Transfusions/last 8 weeks							0.885
	< 6 units	11/ 75 ( 14.7)	2/ 30 ( 6.7)	2.41 ( 0.50, 11.57)	0.263	2.20 ( 0.52, 9.34)	8.00 (-3.99, 19.99)	
	>= 6 units	17/ 78 ( 21.8)	4/ 46 ( 8.7)	2.93 ( 0.92, 9.32)	0.061	2.51 ( 0.90, 7.00)	13.10 ( 0.84, 25.36)	
	RBC Transfusions/last 8 weeks II							0.909
	< 4 units	4/ 28 ( 14.3)	1/ 11 ( 9.1)	1.67 ( 0.17, 16.83)	0.666	1.57 ( 0.20, 12.55)	5.19 (-16.2, 26.56)	
	>= 4 and < 6 units	7/ 47 ( 14.9)	1/ 19 ( 5.3)	3.15 ( 0.36, 27.53)	0.281	2.83 ( 0.37, 21.47)	9.63 (-4.67, 23.93)	
	>= 6 units	17/ 78 ( 21.8)	4/ 46 ( 8.7)	2.93 ( 0.92, 9.32)	0.061	2.51 ( 0.90, 7.00)	13.10 ( 0.84, 25.36)	
	RBC Transfusions/8 weeks over 16 weeks							0.206
	< 6 units	17/ 87 ( 19.5)	2/ 43 ( 4.7)	4.98 ( 1.09, 22.65)	0.024	4.20 ( 1.02, 17.36)	14.89 ( 4.45, 25.33)	
	>= 6 units	11/ 66 ( 16.7)	4/ 33 ( 12.1)	1.45 ( 0.42, 4.96)	0.554	1.38 ( 0.47, 3.99)	4.55 (-9.77, 18.86)	
	RBC Transfusions/8 weeks over 16 weeks II							0.387
	< 4 units	6/ 46 ( 13.0)	1/ 20 ( 5.0)	2.85 ( 0.32, 25.37)	0.333	2.61 ( 0.34, 20.28)	8.04 (-5.59, 21.68)	
	>= 4 and < 6 units	11/ 41 ( 26.8)	1/ 23 ( 4.3)	8.07 ( 0.97, 67.19)	0.028	6.17 ( 0.85, 44.80)	22.48 ( 6.56, 38.40)	
	>= 6 units	11/ 66 ( 16.7)	4/ 33 ( 12.1)	1.45 ( 0.42, 4.96)	0.554	1.38 ( 0.47, 3.99)	4.55 (-9.77, 18.86)	
	IPSS-R risk							0.915
	Very Low	2/ 18 ( 11.1)	0/ 6 ( 0.0)	1.97 ( 0.08, 46.85)	0.404	1.84 ( 0.10, 33.82)	11.11 (-3.41, 25.63)	
	Low	23/109 ( 21.1)	6/ 57 ( 10.5)	2.27 ( 0.87, 5.95)	0.089	2.00 ( 0.87, 4.64)	10.57 (-0.48, 21.63)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).  
 Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis

Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: General disorders and administration site conditions, PT: Asthenia	Intermediate	3/ 25 ( 12.0)	0/ 13 ( 0.0)	4.20 ( 0.20, 87.72)	0.199	3.77 ( 0.21, 67.89)	12.00 (-0.74, 24.74)	
	High	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	Platelet count							0.657
	< 100 x 10 <sup>9</sup> /L	1/ 8 ( 12.5)	1/ 6 ( 16.7)	0.71 ( 0.04, 14.35)	0.832	0.75 ( 0.06, 9.72)	-4.17 (-41.8, 33.44)	
	100-400 x 10 <sup>9</sup> /L	24/128 ( 18.8)	5/ 61 ( 8.2)	2.58 ( 0.93, 7.14)	0.061	2.29 ( 0.92, 5.71)	10.55 ( 0.90, 20.20)	
	> 400 x 10 <sup>9</sup> /L	3/ 17 ( 17.6)	0/ 9 ( 0.0)	4.59 ( 0.21, 99.20)	0.189	3.89 ( 0.22, 67.95)	17.65 (-0.47, 35.77)	
	Renal status: Creatinine Clearance							0.002
	40 - <60 mL/min	5/ 33 ( 15.2)	0/ 12 ( 0.0)	4.82 ( 0.25, 94.08)	0.157	4.21 ( 0.25, 70.82)	15.15 ( 2.92, 27.38)	
	>= 60 mL/min	20/113 ( 17.7)	5/ 62 ( 8.1)	2.45 ( 0.87, 6.89)	0.082	2.19 ( 0.87, 5.56)	9.63 (-0.14, 19.40)	
	< 40 mL/min	3/ 7 ( 42.9)	1/ 1 (100.0)	0.26 ( 0.01, 8.52)	0.317	0.43 ( 0.18, 1.01)	-57.1 (-93.8,-20.48)	
	Serum EPO							0.264
	< 100 U/L	8/ 51 ( 15.7)	3/ 31 ( 9.7)	1.74 ( 0.42, 7.11)	0.442	1.62 ( 0.46, 5.66)	6.01 (-8.41, 20.43)	
	100- < 200 U/L	9/ 37 ( 24.3)	0/ 19 ( 0.0)	13.00 ( 0.71,236.63)	0.020	10.00 ( 0.61,163.10)	24.32 (10.50, 38.15)	
	200-500 U/L	7/ 43 ( 16.3)	3/ 15 ( 20.0)	0.78 ( 0.17, 3.49)	0.745	0.81 ( 0.24, 2.75)	-3.72 (-26.8, 19.33)	
	> 500 U/L	4/ 21 ( 19.0)	0/ 11 ( 0.0)	5.91 ( 0.29,120.58)	0.128	4.91 ( 0.29, 83.67)	19.05 ( 2.25, 35.84)	
	SF3B1 mutation status							0.454
	Mutated	25/141 ( 17.7)	6/ 65 ( 9.2)	2.12 ( 0.82, 5.45)	0.114	1.92 ( 0.83, 4.45)	8.50 (-0.95, 17.95)	
Non-mutated	3/ 12 ( 25.0)	0/ 10 ( 0.0)	7.74 ( 0.35,170.10)	0.096	5.92 ( 0.34,102.64)	25.00 ( 0.50, 49.50)		
Longest duration of prior ESA treatment							0.032	
< 6 months	6/ 48 ( 12.5)	0/ 20 ( 0.0)	6.27 ( 0.34,116.77)	0.100	5.57 ( 0.33, 94.47)	12.50 ( 3.14, 21.86)		
6-12 months	8/ 34 ( 23.5)	4/ 9 ( 44.4)	0.38 ( 0.08, 1.79)	0.219	0.53 ( 0.20, 1.37)	-20.9 (-56.4, 14.54)		
>12-24 months	9/ 35 ( 25.7)	1/ 21 ( 4.8)	6.92 ( 0.81, 59.23)	0.050	5.40 ( 0.74, 39.66)	20.95 ( 3.85, 38.06)		
> 24 months	5/ 31 ( 16.1)	1/ 20 ( 5.0)	3.65 ( 0.39, 33.88)	0.233	3.23 ( 0.41, 25.62)	11.13 (-4.96, 27.22)		

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis  
 Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: Nervous system disorders								
	Age							0.034
	<=64 years	6/ 29 ( 20.7)	6/ 16 ( 37.5)	0.43 ( 0.11, 1.68)	0.227	0.55 ( 0.21, 1.43)	-16.8 (-44.7, 11.12)	
	65-74 years	33/ 72 ( 45.8)	5/ 29 ( 17.2)	4.06 ( 1.39, 11.83)	0.008	2.66 ( 1.15, 6.13)	28.59 (10.66, 46.52)	
	>=75 years	21/ 52 ( 40.4)	6/ 31 ( 19.4)	2.82 ( 0.99, 8.06)	0.049	2.09 ( 0.95, 4.60)	21.03 ( 1.76, 40.30)	
	Gender							0.428
	female	22/ 59 ( 37.3)	7/ 26 ( 26.9)	1.61 ( 0.59, 4.45)	0.356	1.38 ( 0.68, 2.83)	10.37 (-10.7, 31.41)	
	male	38/ 94 ( 40.4)	10/ 50 ( 20.0)	2.71 ( 1.21, 6.08)	0.014	2.02 ( 1.10, 3.71)	20.43 ( 5.55, 35.30)	
	Race							0.022
	Black or African American	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	White	46/107 ( 43.0)	8/ 51 ( 15.7)	4.05 ( 1.74, 9.45)	<.001	2.74 ( 1.40, 5.37)	27.30 (13.61, 41.00)	
	Not Collected or Reported	13/ 44 ( 29.5)	8/ 24 ( 33.3)	0.84 ( 0.29, 2.44)	0.748	0.89 ( 0.43, 1.83)	-3.79 (-27.0, 19.39)	
	Other	1/ 1 (100.0)	1/ 1 (100.0)	NE	NE	NE	NE	
	Region							0.646
	North America	15/ 31 ( 48.4)	6/ 19 ( 31.6)	2.03 ( 0.61, 6.72)	0.247	1.53 ( 0.72, 3.26)	16.81 (-10.5, 44.13)	
	EMEA	45/122 ( 36.9)	11/ 57 ( 19.3)	2.44 ( 1.15, 5.19)	0.018	1.91 ( 1.07, 3.41)	17.59 ( 4.24, 30.94)	
	MDS WHO classification							0.218
	MDS RARS	2/ 7 ( 28.6)	1/ 2 ( 50.0)	0.40 ( 0.02, 10.02)	0.593	0.57 ( 0.09, 3.51)	-21.4 (-98.4, 55.52)	
	MDS RCMD	58/145 ( 40.0)	16/ 74 ( 21.6)	2.42 ( 1.27, 4.61)	0.007	1.85 ( 1.15, 2.98)	18.38 ( 6.07, 30.69)	
	Other	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	Time since initial diagnosis							0.005
	<= 2 years	17/ 40 ( 42.5)	3/ 19 ( 15.8)	3.94 ( 0.99, 15.72)	0.045	2.69 ( 0.90, 8.08)	26.71 ( 4.27, 49.15)	
	>2 to 5 years	20/ 62 ( 32.3)	2/ 34 ( 5.9)	7.62 ( 1.66, 35.00)	0.003	5.48 ( 1.36, 22.06)	26.38 (12.31, 40.45)	
	> 5 years	23/ 51 ( 45.1)	12/ 23 ( 52.2)	0.75 ( 0.28, 2.02)	0.575	0.86 ( 0.53, 1.42)	-7.08 (-31.6, 17.49)	
	ECOG Performance Status	17/ 54 ( 31.5)	8/ 33 ( 24.2)	1.44 ( 0.54, 3.83)	0.472	1.30 ( 0.63, 2.67)	7.24 (-11.9, 26.40)	
	ECOG Performance Status							0.092
	1	40/ 91 ( 44.0)	4/ 32 ( 12.5)	5.49 ( 1.78, 16.94)	0.001	3.52 ( 1.37, 9.05)	31.46 (16.12, 46.80)	
	2	3/ 8 ( 37.5)	5/ 11 ( 45.5)	0.72 ( 0.11, 4.62)	0.736	0.83 ( 0.27, 2.49)	-7.95 (-52.6, 36.67)	
	RBC Transfusions/last 8 weeks							0.342
	< 6 units	30/ 75 ( 40.0)	5/ 30 ( 16.7)	3.33 ( 1.15, 9.67)	0.023	2.40 ( 1.03, 5.60)	23.33 ( 5.99, 40.68)	
	>= 6 units	30/ 78 ( 38.5)	12/ 46 ( 26.1)	1.77 ( 0.80, 3.94)	0.161	1.47 ( 0.84, 2.59)	12.37 (-4.29, 29.04)	
	RBC Transfusions/last 8 weeks II							0.614
	< 4 units	8/ 28 ( 28.6)	1/ 11 ( 9.1)	4.00 ( 0.44, 36.58)	0.200	3.14 ( 0.44, 22.28)	19.48 (-4.36, 43.33)	
	>= 4 and < 6 units	22/ 47 ( 46.8)	4/ 19 ( 21.1)	3.30 ( 0.95, 11.44)	0.054	2.22 ( 0.88, 5.59)	25.76 ( 2.53, 48.98)	
	>= 6 units	30/ 78 ( 38.5)	12/ 46 ( 26.1)	1.77 ( 0.80, 3.94)	0.161	1.47 ( 0.84, 2.59)	12.37 (-4.29, 29.04)	
	RBC Transfusions/8 weeks over 16 weeks							0.278
	< 6 units	36/ 87 ( 41.4)	8/ 43 ( 18.6)	3.09 ( 1.28, 7.44)	0.010	2.22 ( 1.13, 4.36)	22.77 ( 7.21, 38.34)	
	>= 6 units	24/ 66 ( 36.4)	9/ 33 ( 27.3)	1.52 ( 0.61, 3.81)	0.368	1.33 ( 0.70, 2.53)	9.09 (-10.0, 28.21)	
	RBC Transfusions/8 weeks over 16 weeks II							0.526
	< 4 units	16/ 46 ( 34.8)	3/ 20 ( 15.0)	3.02 ( 0.77, 11.88)	0.105	2.32 ( 0.76, 7.08)	19.78 (-1.06, 40.62)	
	>= 4 and < 6 units	20/ 41 ( 48.8)	5/ 23 ( 21.7)	3.43 ( 1.07, 10.99)	0.035	2.24 ( 0.97, 5.18)	27.04 ( 4.28, 49.81)	
	>= 6 units	24/ 66 ( 36.4)	9/ 33 ( 27.3)	1.52 ( 0.61, 3.81)	0.368	1.33 ( 0.70, 2.53)	9.09 (-10.0, 28.21)	
	IPSS-R risk							0.571
	Very Low	8/ 18 ( 44.4)	2/ 6 ( 33.3)	1.60 ( 0.23, 11.08)	0.640	1.33 ( 0.38, 4.63)	11.11 (-33.0, 55.27)	
	Low	43/109 ( 39.4)	11/ 57 ( 19.3)	2.72 ( 1.27, 5.84)	0.009	2.04 ( 1.15, 3.65)	20.15 ( 6.40, 33.90)	
	Intermediate	9/ 25 ( 36.0)	4/ 13 ( 30.8)	1.27 ( 0.30, 5.31)	0.750	1.17 ( 0.44, 3.08)	5.23 (-26.1, 36.59)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).  
 Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis

Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153)		Placebo (N=76)		Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
		n/	N (%)	n/	N (%)					
SOC: Nervous system disorders	High	0/	1 ( 0.0)	0		NE	NE	NE	NE	
	Platelet count									0.066
	< 100 x 10 <sup>9</sup> /L	0/	8 ( 0.0)	1/	6 ( 16.7)	0.22 ( 0.01, 6.31)	0.248	0.26 ( 0.01, 5.44)	-16.7 (-46.5, 13.15)	
	100-400 x 10 <sup>9</sup> /L	55/	128 ( 43.0)	12/	61 ( 19.7)	3.08 ( 1.49, 6.33)	0.002	2.18 ( 1.27, 3.77)	23.30 (10.14, 36.45)	
	> 400 x 10 <sup>9</sup> /L	5/	17 ( 29.4)	4/	9 ( 44.4)	0.52 ( 0.10, 2.79)	0.452	0.66 ( 0.23, 1.87)	-15.0 (-54.1, 23.99)	
	Renal status: Creatinine Clearance									0.795
	40 - <60 mL/min	12/	33 ( 36.4)	2/	12 ( 16.7)	2.86 ( 0.53, 15.26)	0.212	2.18 ( 0.57, 8.36)	19.70 (-7.02, 46.42)	
	>= 60 mL/min	47/	113 ( 41.6)	15/	62 ( 24.2)	2.23 ( 1.12, 4.45)	0.022	1.72 ( 1.05, 2.81)	17.40 ( 3.39, 31.41)	
	< 40 mL/min	1/	7 ( 14.3)	0/	1 ( 0.0)	0.69 ( 0.02, 26.90)	0.705	0.75 ( 0.05, 12.34)	14.29 (-11.6, 40.21)	
	Serum EPO									0.504
	< 100 U/L	20/	51 ( 39.2)	10/	31 ( 32.3)	1.35 ( 0.53, 3.47)	0.528	1.22 ( 0.66, 2.25)	6.96 (-14.3, 28.18)	
	100- < 200 U/L	14/	37 ( 37.8)	3/	19 ( 15.8)	3.25 ( 0.80, 13.17)	0.092	2.40 ( 0.78, 7.33)	22.05 (-0.60, 44.70)	
	200-500 U/L	16/	43 ( 37.2)	2/	15 ( 13.3)	3.85 ( 0.77, 19.31)	0.088	2.79 ( 0.73, 10.74)	23.88 ( 1.41, 46.34)	
	> 500 U/L	9/	21 ( 42.9)	2/	11 ( 18.2)	3.38 ( 0.58, 19.60)	0.169	2.36 ( 0.61, 9.07)	24.68 (-6.43, 55.78)	
	SF3B1 mutation status									0.493
	Mutated	56/	141 ( 39.7)	16/	65 ( 24.6)	2.02 ( 1.05, 3.89)	0.035	1.61 ( 1.01, 2.59)	15.10 ( 1.88, 28.33)	
	Non-mutated	4/	12 ( 33.3)	1/	10 ( 10.0)	4.50 ( 0.41, 49.08)	0.204	3.33 ( 0.44, 25.23)	23.33 (-9.18, 55.85)	
	Longest duration of prior ESA treatment									0.794
	< 6 months	18/	48 ( 37.5)	3/	20 ( 15.0)	3.40 ( 0.87, 13.24)	0.069	2.50 ( 0.83, 7.55)	22.50 ( 1.70, 43.30)	
	6-12 months	12/	34 ( 35.3)	2/	9 ( 22.2)	1.91 ( 0.34, 10.68)	0.462	1.59 ( 0.43, 5.85)	13.07 (-18.5, 44.63)	
	>12-24 months	16/	35 ( 45.7)	5/	21 ( 23.8)	2.69 ( 0.81, 8.99)	0.104	1.92 ( 0.82, 4.47)	21.90 (-2.68, 46.49)	
	> 24 months	12/	31 ( 38.7)	6/	20 ( 30.0)	1.47 ( 0.44, 4.89)	0.529	1.29 ( 0.58, 2.88)	8.71 (-17.7, 35.12)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation  
 Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018  
 Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis  
 Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: Nervous system disorders, PT: Dizziness	Age							0.643
	<=64 years	2/ 29 ( 6.9)	1/ 16 ( 6.3)	1.11 ( 0.09, 13.30)	0.934	1.10 ( 0.11, 11.25)	0.65 (-14.4, 15.67)	
	65-74 years	13/ 72 ( 18.1)	2/ 29 ( 6.9)	2.97 ( 0.63, 14.11)	0.156	2.62 ( 0.63, 10.88)	11.16 (-1.65, 23.97)	
	>=75 years	8/ 52 ( 15.4)	1/ 31 ( 3.2)	5.45 ( 0.65, 45.90)	0.087	4.77 ( 0.63, 36.34)	12.16 ( 0.55, 23.77)	
	Gender							0.241
	female	6/ 59 ( 10.2)	2/ 26 ( 7.7)	1.36 ( 0.26, 7.23)	0.720	1.32 ( 0.29, 6.12)	2.48 (-10.3, 15.30)	
	male	17/ 94 ( 18.1)	2/ 50 ( 4.0)	5.30 ( 1.17, 23.96)	0.018	4.52 ( 1.09, 18.79)	14.09 ( 4.60, 23.57)	
	Race							0.794
	Black or African American	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	White	19/107 ( 17.8)	3/ 51 ( 5.9)	3.45 ( 0.97, 12.27)	0.044	3.02 ( 0.94, 9.74)	11.87 ( 2.17, 21.58)	
	Not Collected or Reported	4/ 44 ( 9.1)	1/ 24 ( 4.2)	2.30 ( 0.24, 21.83)	0.461	2.18 ( 0.26, 18.44)	4.92 (-6.74, 16.59)	
	Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE	
	Region							0.531
	North America	8/ 31 ( 25.8)	1/ 19 ( 5.3)	6.26 ( 0.72, 54.75)	0.069	4.90 ( 0.66, 36.19)	20.54 ( 2.16, 38.93)	
	EMEA	15/122 ( 12.3)	3/ 57 ( 5.3)	2.52 ( 0.70, 9.09)	0.146	2.34 ( 0.70, 7.75)	7.03 (-1.19, 15.25)	
	MDS WHO classification							NE
	MDS RARS	0/ 7 ( 0.0)	0/ 2 ( 0.0)	NE	NE	NE	NE	
	MDS RCMD	23/145 ( 15.9)	4/ 74 ( 5.4)	3.30 ( 1.10, 9.93)	0.026	2.93 ( 1.05, 8.17)	10.46 ( 2.59, 18.32)	
	Other	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	Time since initial diagnosis							0.043
	<= 2 years	9/ 40 ( 22.5)	0/ 19 ( 0.0)	11.76 ( 0.65,213.62)	0.026	9.27 ( 0.57,151.37)	22.50 ( 9.56, 35.44)	
	>2 to 5 years	8/ 62 ( 12.9)	0/ 34 ( 0.0)	10.76 ( 0.60,192.45)	0.030	9.44 ( 0.56,158.78)	12.90 ( 4.56, 21.25)	
	> 5 years	6/ 51 ( 11.8)	4/ 23 ( 17.4)	0.63 ( 0.16, 2.50)	0.515	0.68 ( 0.21, 2.17)	-5.63 (-23.5, 12.21)	
	ECOG Performance Status							0.230
	ECOG Performance Status	7/ 54 ( 13.0)	3/ 33 ( 9.1)	1.49 ( 0.36, 6.21)	0.585	1.43 ( 0.40, 5.14)	3.87 (-9.41, 17.16)	
	1	16/ 91 ( 17.6)	1/ 32 ( 3.1)	6.61 ( 0.84, 52.05)	0.042	5.63 ( 0.78, 40.74)	14.46 ( 4.58, 24.33)	
	2	0/ 8 ( 0.0)	0/ 11 ( 0.0)	NE	NE	NE	NE	
	RBC Transfusions/last 8 weeks							0.245
< 6 units	15/ 75 ( 20.0)	1/ 30 ( 3.3)	7.25 ( 0.91, 57.58)	0.033	6.00 ( 0.83, 43.43)	16.67 ( 5.57, 27.77)		
>= 6 units	8/ 78 ( 10.3)	3/ 46 ( 6.5)	1.64 ( 0.41, 6.51)	0.482	1.57 ( 0.44, 5.63)	3.73 (-6.08, 13.55)		
RBC Transfusions/last 8 weeks II							0.443	
< 4 units	4/ 28 ( 14.3)	1/ 11 ( 9.1)	1.67 ( 0.17, 16.83)	0.666	1.57 ( 0.20, 12.55)	5.19 (-16.2, 26.56)		
>= 4 and < 6 units	11/ 47 ( 23.4)	0/ 19 ( 0.0)	12.29 ( 0.69,219.84)	0.022	9.58 ( 0.59,154.93)	23.40 (11.30, 35.51)		
>= 6 units	8/ 78 ( 10.3)	3/ 46 ( 6.5)	1.64 ( 0.41, 6.51)	0.482	1.57 ( 0.44, 5.63)	3.73 (-6.08, 13.55)		
RBC Transfusions/8 weeks over 16 weeks							0.130	
< 6 units	15/ 87 ( 17.2)	1/ 43 ( 2.3)	8.75 ( 1.12, 68.63)	0.015	7.41 ( 1.01, 54.29)	14.92 ( 5.79, 24.04)		
>= 6 units	8/ 66 ( 12.1)	3/ 33 ( 9.1)	1.38 ( 0.34, 5.58)	0.653	1.33 ( 0.38, 4.70)	3.03 (-9.55, 15.61)		
RBC Transfusions/8 weeks over 16 weeks II							0.346	
< 4 units	6/ 46 ( 13.0)	1/ 20 ( 5.0)	2.85 ( 0.32, 25.37)	0.333	2.61 ( 0.34, 20.28)	8.04 (-5.59, 21.68)		
>= 4 and < 6 units	9/ 41 ( 22.0)	0/ 23 ( 0.0)	13.74 ( 0.76,247.92)	0.016	10.86 ( 0.66,178.41)	21.95 ( 9.28, 34.62)		
>= 6 units	8/ 66 ( 12.1)	3/ 33 ( 9.1)	1.38 ( 0.34, 5.58)	0.653	1.33 ( 0.38, 4.70)	3.03 (-9.55, 15.61)		
IPSS-R risk							0.264	
Very Low	1/ 18 ( 5.6)	1/ 6 ( 16.7)	0.29 ( 0.02, 5.60)	0.404	0.33 ( 0.02, 4.55)	-11.1 (-42.8, 20.53)		
Low	18/109 ( 16.5)	3/ 57 ( 5.3)	3.56 ( 1.00, 12.65)	0.039	3.14 ( 0.96, 10.21)	11.25 ( 2.18, 20.32)		

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).  
 Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.  
 Relative risk and risk difference are estimated using normal approximation  
 p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one frequent Adverse Event by SOC, PT during week 1-24 (incidence in either arm >= 10% or >=10 patients) - Subgroup analysis

Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: Nervous system disorders, PT: Dizziness	Intermediate	4/ 25 ( 16.0)	0/ 13 ( 0.0)	5.65 ( 0.28,113.50)	0.132	4.85 ( 0.28, 83.66)	16.00 ( 1.63, 30.37)	
	High	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	Platelet count							0.602
	< 100 x 10 <sup>9</sup> /L	0/ 8 ( 0.0)	0/ 6 ( 0.0)	NE	NE	NE	NE	
	100-400 x 10 <sup>9</sup> /L	19/128 ( 14.8)	4/ 61 ( 6.6)	2.48 ( 0.81, 7.65)	0.104	2.26 ( 0.80, 6.37)	8.29 (-0.46, 17.03)	
	> 400 x 10 <sup>9</sup> /L	4/ 17 ( 23.5)	0/ 9 ( 0.0)	6.33 ( 0.30,132.05)	0.121	5.00 ( 0.30, 83.69)	23.53 ( 3.37, 43.69)	
	Renal status: Creatinine Clearance							0.987
	40 - <60 mL/min	3/ 33 ( 9.1)	0/ 12 ( 0.0)	2.87 ( 0.14, 59.70)	0.285	2.68 ( 0.15, 48.32)	9.09 (-0.72, 18.90)	
	>= 60 mL/min	20/113 ( 17.7)	4/ 62 ( 6.5)	3.12 ( 1.01, 9.58)	0.039	2.74 ( 0.98, 7.67)	11.25 ( 1.92, 20.57)	
	< 40 mL/min	0/ 7 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE	
	Serum EPO							0.501
	< 100 U/L	9/ 51 ( 17.6)	2/ 31 ( 6.5)	3.11 ( 0.63, 15.45)	0.152	2.74 ( 0.63, 11.85)	11.20 (-2.38, 24.77)	
	100- < 200 U/L	6/ 37 ( 16.2)	0/ 19 ( 0.0)	8.05 ( 0.43,150.92)	0.066	6.84 ( 0.41,115.37)	16.22 ( 4.34, 28.09)	
	200-500 U/L	5/ 43 ( 11.6)	2/ 15 ( 13.3)	0.86 ( 0.15, 4.95)	0.863	0.87 ( 0.19, 4.03)	-1.71 (-21.4, 17.99)	
	> 500 U/L	3/ 21 ( 14.3)	0/ 11 ( 0.0)	4.35 ( 0.21, 92.18)	0.195	3.82 ( 0.21, 67.92)	14.29 (-0.68, 29.25)	
	SF3B1 mutation status							NE
	Mutated	23/141 ( 16.3)	4/ 65 ( 6.2)	2.97 ( 0.98, 8.98)	0.045	2.65 ( 0.96, 7.35)	10.16 ( 1.71, 18.60)	
Non-mutated	0/ 12 ( 0.0)	0/ 10 ( 0.0)	NE	NE	NE	NE		
Longest duration of prior ESA treatment							0.784	
< 6 months	6/ 48 ( 12.5)	0/ 20 ( 0.0)	6.27 ( 0.34,116.77)	0.100	5.57 ( 0.33, 94.47)	12.50 ( 3.14, 21.86)		
6-12 months	4/ 34 ( 11.8)	1/ 9 ( 11.1)	1.07 ( 0.10, 10.92)	0.957	1.06 ( 0.13, 8.34)	0.65 (-22.6, 23.87)		
>12-24 months	7/ 35 ( 20.0)	2/ 21 ( 9.5)	2.38 ( 0.44, 12.69)	0.306	2.10 ( 0.48, 9.18)	10.48 (-7.78, 28.73)		
> 24 months	5/ 31 ( 16.1)	1/ 20 ( 5.0)	3.65 ( 0.39, 33.88)	0.233	3.23 ( 0.41, 25.62)	11.13 (-4.96, 27.22)		

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one frequent Adverse Event of CTCAE >=3 by SOC, PT during week 1-24 (incidence in either arm >= 5% or >=10 patients) - Subgroup analysis

Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: Nervous system disorders								
	Age							0.854
	<=64 years	0/ 29 ( 0.0)	0/ 16 ( 0.0)	NE	NE	NE	NE	
	65-74 years	4/ 72 ( 5.6)	0/ 29 ( 0.0)	3.88 ( 0.20, 74.31)	0.197	3.70 ( 0.21, 66.59)	5.56 ( 0.26, 10.85)	
	>=75 years	4/ 52 ( 7.7)	0/ 31 ( 0.0)	5.85 ( 0.30,112.35)	0.116	5.43 ( 0.30, 97.64)	7.69 ( 0.45, 14.93)	
	Gender							0.933
	female	4/ 59 ( 6.8)	0/ 26 ( 0.0)	4.30 ( 0.22, 82.78)	0.176	4.05 ( 0.23, 72.60)	6.78 ( 0.36, 13.19)	
	male	4/ 94 ( 4.3)	0/ 50 ( 0.0)	5.02 ( 0.26, 95.18)	0.140	4.83 ( 0.27, 87.97)	4.26 ( 0.17, 8.34)	
	Race							0.486
	Black or African American	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	White	7/107 ( 6.5)	0/ 51 ( 0.0)	7.69 ( 0.43,137.25)	0.063	7.22 ( 0.42,124.05)	6.54 ( 1.86, 11.23)	
	Not Collected or Reported	1/ 44 ( 2.3)	0/ 24 ( 0.0)	1.69 ( 0.07, 43.09)	0.460	1.67 ( 0.07, 39.40)	2.27 (-2.13, 6.68)	
	Other	0/ 1 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE	
	Region							0.528
	North America	1/ 31 ( 3.2)	0/ 19 ( 0.0)	1.92 ( 0.07, 49.50)	0.434	1.88 ( 0.08, 43.82)	3.23 (-2.99, 9.45)	
	EMEA	7/122 ( 5.7)	0/ 57 ( 0.0)	7.47 ( 0.42,133.05)	0.066	7.07 ( 0.41,121.75)	5.74 ( 1.61, 9.86)	
	MDS WHO classification							NE
	MDS RARS	0/ 7 ( 0.0)	0/ 2 ( 0.0)	NE	NE	NE	NE	
	MDS RCMD	8/145 ( 5.5)	0/ 74 ( 0.0)	9.21 ( 0.52,161.82)	0.040	8.73 ( 0.51,149.25)	5.52 ( 1.80, 9.23)	
	Other	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	Time since initial diagnosis							0.995
	<= 2 years	3/ 40 ( 7.5)	0/ 19 ( 0.0)	3.64 ( 0.18, 74.09)	0.224	3.41 ( 0.19, 62.97)	7.50 (-0.66, 15.66)	
	>2 to 5 years	2/ 62 ( 3.2)	0/ 34 ( 0.0)	2.85 ( 0.13, 61.12)	0.292	2.78 ( 0.14, 56.25)	3.23 (-1.17, 7.62)	
	> 5 years	3/ 51 ( 5.9)	0/ 23 ( 0.0)	3.39 ( 0.17, 68.39)	0.238	3.23 ( 0.17, 60.11)	5.88 (-0.58, 12.34)	
	ECOG Performance Status	3/ 54 ( 5.6)	0/ 33 ( 0.0)	4.55 ( 0.23, 91.00)	0.171	4.33 ( 0.23, 81.22)	5.56 (-0.55, 11.67)	
	ECOG Performance Status							0.990
	1	4/ 91 ( 4.4)	0/ 32 ( 0.0)	3.34 ( 0.18, 63.83)	0.230	3.23 ( 0.18, 58.35)	4.40 ( 0.18, 8.61)	
	2	1/ 8 ( 12.5)	0/ 11 ( 0.0)	4.60 ( 0.16,128.54)	0.241	4.00 ( 0.18, 87.18)	12.50 (-10.4, 35.42)	
	RBC Transfusions/last 8 weeks							0.857
	< 6 units	4/ 75 ( 5.3)	0/ 30 ( 0.0)	3.84 ( 0.20, 73.52)	0.199	3.67 ( 0.20, 66.17)	5.33 ( 0.25, 10.42)	
	>= 6 units	4/ 78 ( 5.1)	0/ 46 ( 0.0)	5.62 ( 0.30,106.75)	0.120	5.35 ( 0.29, 97.25)	5.13 ( 0.23, 10.02)	
	RBC Transfusions/last 8 weeks II							0.872
	< 4 units	2/ 28 ( 7.1)	0/ 11 ( 0.0)	2.17 ( 0.10, 48.86)	0.369	2.07 ( 0.11, 39.96)	7.14 (-2.40, 16.68)	
	>= 4 and < 6 units	2/ 47 ( 4.3)	0/ 19 ( 0.0)	2.14 ( 0.10, 46.74)	0.365	2.08 ( 0.10, 41.48)	4.26 (-1.52, 10.03)	
	>= 6 units	4/ 78 ( 5.1)	0/ 46 ( 0.0)	5.62 ( 0.30,106.75)	0.120	5.35 ( 0.29, 97.25)	5.13 ( 0.23, 10.02)	
	RBC Transfusions/8 weeks over 16 weeks							0.834
	< 6 units	5/ 87 ( 5.7)	0/ 43 ( 0.0)	5.80 ( 0.31,107.35)	0.110	5.50 ( 0.31, 97.23)	5.75 ( 0.86, 10.64)	
	>= 6 units	3/ 66 ( 4.5)	0/ 33 ( 0.0)	3.69 ( 0.19, 73.63)	0.216	3.55 ( 0.19, 66.81)	4.55 (-0.48, 9.57)	
	RBC Transfusions/8 weeks over 16 weeks II							0.959
	< 4 units	2/ 46 ( 4.3)	0/ 20 ( 0.0)	2.30 ( 0.11, 50.17)	0.347	2.23 ( 0.11, 44.54)	4.35 (-1.55, 10.24)	
	>= 4 and < 6 units	3/ 41 ( 7.3)	0/ 23 ( 0.0)	4.27 ( 0.21, 86.45)	0.187	4.00 ( 0.22, 74.20)	7.32 (-0.65, 15.29)	
	>= 6 units	3/ 66 ( 4.5)	0/ 33 ( 0.0)	3.69 ( 0.19, 73.63)	0.216	3.55 ( 0.19, 66.81)	4.55 (-0.48, 9.57)	
	IPSS-R risk							0.642
	Very Low	1/ 18 ( 5.6)	0/ 6 ( 0.0)	1.11 ( 0.04, 30.97)	0.564	1.11 ( 0.05, 24.07)	5.56 (-5.03, 16.14)	
	Low	6/109 ( 5.5)	0/ 57 ( 0.0)	7.22 ( 0.40,130.53)	0.072	6.85 ( 0.39,119.55)	5.50 ( 1.22, 9.79)	
	Intermediate	1/ 25 ( 4.0)	0/ 13 ( 0.0)	1.65 ( 0.06, 43.44)	0.471	1.62 ( 0.07, 37.10)	4.00 (-3.68, 11.68)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Dossier zur Nutzenbewertung - Modul 4 B  
 Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Stand: 27.07.2020

Celgene Corporation

Luspatercept (ACE-536), Protocol ACE-536-MDS-001 - Cutoff Date: 08MAY2018

Proportion of patients with at least one frequent Adverse Event of CTCAE >=3 by SOC, PT during week 1-24 (incidence in either arm >= 5% or >=10 patients) - Subgroup analysis

Safety Population

SOC / PT	Subgroup Level	Luspatercept (N=153) n/ N (%)	Placebo (N=76) n/ N (%)	Odds Ratio (95% CI)	p-Value	Relative Risk (95% CI)	Risk Difference (95% CI)	p-value for interaction
SOC: Nervous system disorders	High	0/ 1 ( 0.0)	0	NE	NE	NE	NE	
	Platelet count							0.482
	< 100 x 10 <sup>9</sup> /L	0/ 8 ( 0.0)	0/ 6 ( 0.0)	NE	NE	NE	NE	
	100-400 x 10 <sup>9</sup> /L	7/128 ( 5.5)	0/ 61 ( 0.0)	7.59 ( 0.43,135.13)	0.063	7.21 ( 0.42,124.22)	5.47 ( 1.53, 9.41)	
	> 400 x 10 <sup>9</sup> /L	1/ 17 ( 5.9)	0/ 9 ( 0.0)	1.73 ( 0.06, 46.77)	0.467	1.67 ( 0.07, 37.21)	5.88 (-5.30, 17.07)	
	Renal status: Creatinine Clearance							0.691
	40 - <60 mL/min	3/ 33 ( 9.1)	0/ 12 ( 0.0)	2.87 ( 0.14, 59.70)	0.285	2.68 ( 0.15, 48.32)	9.09 (-0.72, 18.90)	
	>= 60 mL/min	5/113 ( 4.4)	0/ 62 ( 0.0)	6.34 ( 0.34,116.52)	0.094	6.08 ( 0.34,108.15)	4.42 ( 0.63, 8.22)	
	< 40 mL/min	0/ 7 ( 0.0)	0/ 1 ( 0.0)	NE	NE	NE	NE	
	Serum EPO							0.984
	< 100 U/L	1/ 51 ( 2.0)	0/ 31 ( 0.0)	1.87 ( 0.07, 47.37)	0.436	1.85 ( 0.08, 43.96)	1.96 (-1.84, 5.77)	
	100- < 200 U/L	1/ 37 ( 2.7)	0/ 19 ( 0.0)	1.60 ( 0.06, 41.23)	0.474	1.58 ( 0.07, 37.01)	2.70 (-2.52, 7.93)	
	200-500 U/L	4/ 43 ( 9.3)	0/ 15 ( 0.0)	3.53 ( 0.18, 69.55)	0.225	3.27 ( 0.19, 57.44)	9.30 ( 0.62, 17.98)	
	> 500 U/L	1/ 21 ( 4.8)	0/ 11 ( 0.0)	1.68 ( 0.06, 44.77)	0.469	1.64 ( 0.07, 37.15)	4.76 (-4.35, 13.87)	
	SF3B1 mutation status							0.629
	Mutated	7/141 ( 5.0)	0/ 65 ( 0.0)	7.30 ( 0.41,129.86)	0.068	6.97 ( 0.40,120.26)	4.96 ( 1.38, 8.55)	
	Non-mutated	1/ 12 ( 8.3)	0/ 10 ( 0.0)	2.74 ( 0.10, 74.87)	0.361	2.54 ( 0.11, 56.25)	8.33 (-7.30, 23.97)	
	Longest duration of prior ESA treatment							0.988
	< 6 months	2/ 48 ( 4.2)	0/ 20 ( 0.0)	2.20 ( 0.10, 47.99)	0.358	2.14 ( 0.11, 42.74)	4.17 (-1.49, 9.82)	
	6-12 months	2/ 34 ( 5.9)	0/ 9 ( 0.0)	1.46 ( 0.06, 33.14)	0.461	1.43 ( 0.07, 27.40)	5.88 (-2.03, 13.79)	
	>12-24 months	2/ 35 ( 5.7)	0/ 21 ( 0.0)	3.21 ( 0.15, 70.12)	0.269	3.06 ( 0.15, 60.74)	5.71 (-1.98, 13.40)	
	> 24 months	1/ 31 ( 3.2)	0/ 20 ( 0.0)	2.02 ( 0.08, 51.96)	0.422	1.97 ( 0.08, 46.08)	3.23 (-2.99, 9.45)	

Subgroup analysis is done for SOCs / PTs with significant overall treatment effect (p-value from CMH-Test <=0.05).

Odds ratio (luspatercept over placebo) with 95% CI and p-value are estimated from Cochran Mantel-Haenszel test.

Relative risk and risk difference are estimated using normal approximation

p-Value for interaction from test for heterogeneity of the relative risks using Cochran's Q statistic.

Table 14.2.18.7.1.1.1 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	9	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	32.1%	18.8%
		OR (95% CI) [b]	2.05 (0.46, 9.06)	
		p-value [b]	0.3425	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6178	
		RR (95% CI) [d]	1.71 (0.54, 5.43)	
		Risk Diff (95% CI) [e]	13.4% (-12.4, 39.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.6%	27.3%
	OR (95% CI) [b]	1.41 (0.30, 6.68)		
	p-value [b]	0.6635		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7534		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 513 von 1722

Table 14.2.18.7.1.1.1 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.27 (0.42, 3.81)	
		Risk Diff (95% CI) [e]	7.3% (-24.7, 39.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.1 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	21
	Number of subjects with clinically meaningful improvement at Week 24	n	10	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	10.3%
		OR (95% CI) [b]	1.44 (0.37, 5.68)	
		p-value [b]	0.5988	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8942	
		RR (95% CI) [d]	1.38 (0.41, 4.66)	
		Risk Diff (95% CI) [e]	3.9% (-9.8, 17.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.7%	14.3%
		OR (95% CI) [b]	1.76 (0.43, 7.24)	
		p-value [b]	0.4302	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5668	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.1 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.59 (0.49, 5.18)	
		Risk Diff (95% CI) [e]	8.4% (-11.0, 27.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.1 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	12	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.5%	19.4%
		OR (95% CI) [b]	1.28 (0.43, 3.86)	
		p-value [b]	0.6583	
		RR (95% CI) [d]	1.22 (0.51, 2.91)	
		Risk Diff (95% CI) [e]	4.2% (-14.0, 22.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.3%	28.6%
		OR (95% CI) [b]	1.03 (0.32, 3.30)	
		p-value [b]	0.9544	
		RR (95% CI) [d]	1.02 (0.45, 2.34)	
		Risk Diff (95% CI) [e]	0.7% (-23.1, 24.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 517 von 1722

Table 14.2.18.7.1.2.1 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	8	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.6%	6.3%
		OR (95% CI) [b]	6.00 (0.68, 53.29)	
		p-value [b]	0.1078	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1078	
		RR (95% CI) [d]	4.57 (0.63, 33.31)	
		Risk Diff (95% CI) [e]	22.3% (1.8, 42.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%
		OR (95% CI) [b]	4.44 (0.48, 40.84)	
		p-value [b]	0.1874	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1177	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.1 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	3.38 (0.48, 23.92)	
		Risk Diff (95% CI) [e]	21.7% (-2.9, 46.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.1 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	22
	Number of subjects with clinically meaningful improvement at Week 24	n	4	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.7%	27.6%
		OR (95% CI) [b]	0.16 (0.04, 0.58)	
		p-value [b]	0.0055	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0469	
		RR (95% CI) [d]	0.21 (0.07, 0.63)	
		Risk Diff (95% CI) [e]	-21.9% (-39.0, -4.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	9.1%	36.4%
		OR (95% CI) [b]	0.18 (0.05, 0.67)	
		p-value [b]	0.0111	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1522		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 520 von 1722

Table 14.2.18.7.1.2.1 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.25 (0.08, 0.74)	
		Risk Diff (95% CI) [e]	-27.3% (-49.1, -5.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 521 von 1722

Table 14.2.18.7.1.2.1 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	13	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.5%	29.0%
		OR (95% CI) [b]	0.84 (0.31, 2.27)	
		p-value [b]	0.7257	
		RR (95% CI) [d]	0.88 (0.43, 1.81)	
		Risk Diff (95% CI) [e]	-3.5% (-23.5, 16.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.7%	42.9%
		OR (95% CI) [b]	0.62 (0.21, 1.83)	
		p-value [b]	0.3868	
		RR (95% CI) [d]	0.74 (0.38, 1.44)	
		Risk Diff (95% CI) [e]	-11.1% (-36.7, 14.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 522 von 1722

Table 14.2.18.7.1.3.1 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	12.5%
		OR (95% CI) [b]	1.52 (0.26, 8.93)	
		p-value [b]	0.6419	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7279	
		RR (95% CI) [d]	1.43 (0.31, 6.53)	
		Risk Diff (95% CI) [e]	5.4% (-16.2, 26.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.2%	18.2%
	OR (95% CI) [b]	1.07 (0.17, 6.59)		
	p-value [b]	0.9407		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8129		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.1 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.06 (0.24, 4.65)	
		Risk Diff (95% CI) [e]	1.0% (-26.3, 28.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.1 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	22
	Number of subjects with clinically meaningful improvement at Week 24	n	13	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.6%	31.0%
		OR (95% CI) [b]	0.51 (0.19, 1.37)	
		p-value [b]	0.1789	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3225	
		RR (95% CI) [d]	0.60 (0.29, 1.24)	
		Risk Diff (95% CI) [e]	-12.5% (-31.6, 6.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.5%	40.9%
		OR (95% CI) [b]	0.61 (0.21, 1.76)	
		p-value [b]	0.3578	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6940		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 525 von 1722

Table 14.2.18.7.1.3.1 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.37, 1.42)	
		Risk Diff (95% CI) [e]	-11.4% (-35.9, 13.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 526 von 1722

Table 14.2.18.7.1.3.1 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	12	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.5%	22.6%
		OR (95% CI) [b]	1.05 (0.36, 3.05)	
		p-value [b]	0.9214	
		RR (95% CI) [d]	1.04 (0.46, 2.36)	
		Risk Diff (95% CI) [e]	0.9% (-17.8, 19.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.3%	33.3%
		OR (95% CI) [b]	0.83 (0.27, 2.56)	
		p-value [b]	0.7426	
		RR (95% CI) [d]	0.88 (0.41, 1.90)	
		Risk Diff (95% CI) [e]	-4.1% (-28.6, 20.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 527 von 1722

Table 14.2.18.7.1.4.1 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	11	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	39.3%	12.5%
		OR (95% CI) [b]	4.53 (0.86, 23.93)	
		p-value [b]	0.0753	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1022	
		RR (95% CI) [d]	3.14 (0.79, 12.44)	
		Risk Diff (95% CI) [e]	26.8% (2.5, 51.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	42.3%
		OR (95% CI) [b]	3.30 (0.59, 18.40)	
		p-value [b]	0.1733	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1273	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.1 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.33 (0.61, 8.81)	
		Risk Diff (95% CI) [e]	24.1% (-5.5, 53.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.1 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	21
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.6%	13.8%
		OR (95% CI) [b]	0.59 (0.15, 2.25)	
		p-value [b]	0.4368	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6344	
		RR (95% CI) [d]	0.62 (0.19, 2.04)	
		Risk Diff (95% CI) [e]	-5.2% (-19.4, 8.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.6%	19.0%
		OR (95% CI) [b]	0.67 (0.17, 2.69)	
		p-value [b]	0.5734	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9982		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.1 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.23, 2.27)	
		Risk Diff (95% CI) [e]	-5.4% (-25.0, 14.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.1 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	12	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.5%	25.8%
		OR (95% CI) [b]	0.88 (0.32, 2.48)	
		p-value [b]	0.8159	
		RR (95% CI) [d]	0.91 (0.42, 1.98)	
		Risk Diff (95% CI) [e]	-2.3% (-21.6, 17.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.3%	38.1%
		OR (95% CI) [b]	0.67 (0.22, 2.04)	
		p-value [b]	0.4828	
		RR (95% CI) [d]	0.77 (0.37, 1.58)	
		Risk Diff (95% CI) [e]	-8.8% (-33.8, 16.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 532 von 1722

Table 14.2.18.7.1.5.1 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	3	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.7%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9673	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9651	
		RR (95% CI) [d]	4.10 (0.23, 74.73)	
		Risk Diff (95% CI) [e]	9.1% (-5.2, 23.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.5%	0.0%
	OR (95% CI) [b]	NE (NE, NE)		
	p-value [b]	0.9537		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9491		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 533 von 1722

Table 14.2.18.7.1.5.1 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	3.11 (0.17, 55.65)	
		Risk Diff (95% CI) [e]	8.8% (-8.2, 25.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.1 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	21
	Number of subjects with clinically meaningful improvement at Week 24	n	9	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.9%	13.8%
		OR (95% CI) [b]	0.92 (0.26, 3.27)	
		p-value [b]	0.9002	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4301	
		RR (95% CI) [d]	0.93 (0.31, 2.79)	
		Risk Diff (95% CI) [e]	-0.9% (-15.7, 13.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.5%
	OR (95% CI) [b]	1.09 (0.29, 4.06)		
	p-value [b]	0.8945		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2108		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 535 von 1722

Table 14.2.18.7.1.5.1 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.07 (0.37, 3.09)	
		Risk Diff (95% CI) [e]	1.4% (-19.2, 22.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 536 von 1722

Table 14.2.18.7.1.5.1 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluatable subjects at baseline	N	51	31
	Number of HRQoL evaluatable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	6	7
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	11.8%	22.6%
		OR (95% CI) [b]	0.46 (0.14, 1.51)	
		p-value [b]	0.2002	
		RR (95% CI) [d]	0.52 (0.19, 1.41)	
		Risk Diff (95% CI) [e]	-10.8% (-28.0, 6.4)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.6%	33.3%
		OR (95% CI) [b]	0.34 (0.10, 1.20)	
		p-value [b]	0.0944	
		RR (95% CI) [d]	0.44 (0.17, 1.14)	
		Risk Diff (95% CI) [e]	-18.7% (-41.6, 4.2)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 537 von 1722

Table 14.2.18.7.1.6.1 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	12.5%
		OR (95% CI) [b]	2.33 (0.42, 12.91)	
		p-value [b]	0.3317	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2900	
		RR (95% CI) [d]	2.00 (0.47, 8.50)	
		Risk Diff (95% CI) [e]	12.5% (-10.3, 35.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.9%	18.2%
	OR (95% CI) [b]	1.66 (0.29, 9.64)		
	p-value [b]	0.5735		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3363		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 538 von 1722

Table 14.2.18.7.1.6.1 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.48 (0.36, 6.03)	
		Risk Diff (95% CI) [e]	8.7% (-19.7, 37.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.1 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	21
	Number of subjects with clinically meaningful improvement at Week 24	n	8	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.4%	20.7%
		OR (95% CI) [b]	0.49 (0.15, 1.58)	
		p-value [b]	0.2349	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5563	
		RR (95% CI) [d]	0.55 (0.21, 1.45)	
		Risk Diff (95% CI) [e]	-9.3% (-25.8, 7.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.2%
	OR (95% CI) [b]	0.56 (0.16, 1.88)		
	p-value [b]	0.3442		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9339		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 540 von 1722

Table 14.2.18.7.1.6.1 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.64 (0.25, 1.60)	
		Risk Diff (95% CI) [e]	-10.4% (-32.8, 12.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.1 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluatable subjects at baseline	N	51	31
	Number of HRQoL evaluatable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	11	8
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	21.6%	25.8%
		OR (95% CI) [b]	0.79 (0.28, 2.25)	
		p-value [b]	0.6595	
		RR (95% CI) [d]	0.84 (0.38, 1.85)	
		Risk Diff (95% CI) [e]	-4.2% (-23.3, 14.9)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.8%	38.1%
		OR (95% CI) [b]	0.60 (0.19, 1.82)	
	p-value [b]	0.3646		
	RR (95% CI) [d]	0.70 (0.33, 1.48)		
	Risk Diff (95% CI) [e]	-11.3% (-36.1, 13.5)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.1 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	11	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	39.3%	12.5%
		OR (95% CI) [b]	4.53 (0.86, 23.93)	
		p-value [b]	0.0753	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0259	
		RR (95% CI) [d]	3.14 (0.79, 12.44)	
		Risk Diff (95% CI) [e]	26.8% (2.5, 51.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	42.3%	18.2%
	OR (95% CI) [b]	3.30 (0.59, 18.40)		
	p-value [b]	0.1733		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0277		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.1 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.33 (0.61, 8.81)	
		Risk Diff (95% CI) [e]	24.1% (-5.5, 53.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.1 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	43	22
	Number of subjects with clinically meaningful improvement at Week 24	n	10	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	37.9%
		OR (95% CI) [b]	0.27 (0.10, 0.75)	
		p-value [b]	0.0113	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4006	
		RR (95% CI) [d]	0.38 (0.18, 0.79)	
		Risk Diff (95% CI) [e]	-23.6% (-43.1, -4.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.3%
	OR (95% CI) [b]	0.30 (0.10, 0.91)		
	p-value [b]	0.0326		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9042		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 545 von 1722

Table 14.2.18.7.1.7.1 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.47 (0.23, 0.92)	
		Risk Diff (95% CI) [e]	-26.7% (-51.2, -2.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.1 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	11	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.6%	35.5%
		OR (95% CI) [b]	0.50 (0.19, 1.35)	
		p-value [b]	0.1714	
		RR (95% CI) [d]	0.61 (0.30, 1.23)	
		Risk Diff (95% CI) [e]	-13.9% (-34.2, 6.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.8%	52.4%
		OR (95% CI) [b]	0.33 (0.11, 1.00)	
		p-value [b]	0.0504	
		RR (95% CI) [d]	0.51 (0.27, 0.98)	
		Risk Diff (95% CI) [e]	-25.6% (-50.9, -0.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 547 von 1722

Table 14.2.18.7.1.8.1 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	3	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.7%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9555	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9582	
		RR (95% CI) [d]	4.10 (0.23, 74.73)	
		Risk Diff (95% CI) [e]	9.1% (-5.2, 23.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.5%	0.0%
	OR (95% CI) [b]	NE (NE, NE)		
	p-value [b]	0.9582		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9601		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 548 von 1722

Table 14.2.18.7.1.8.1 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	3.11 (0.17, 55.65)	
		Risk Diff (95% CI) [e]	8.8% (-8.2, 25.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.1 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	22
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.1%	6.9%
		OR (95% CI) [b]	1.04 (0.19, 5.69)	
		p-value [b]	0.9653	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5587	
		RR (95% CI) [d]	1.04 (0.21, 5.04)	
		Risk Diff (95% CI) [e]	0.2% (-10.8, 11.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.4%	9.1%
		OR (95% CI) [b]	1.28 (0.23, 7.20)	
		p-value [b]	0.7779	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8100		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.1 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.25 (0.26, 5.94)	
		Risk Diff (95% CI) [e]	2.3% (-13.0, 17.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.1 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	9	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.6%	9.7%
		OR (95% CI) [b]	2.00 (0.50, 8.04)	
		p-value [b]	0.3289	
		RR (95% CI) [d]	1.82 (0.53, 6.23)	
		Risk Diff (95% CI) [e]	8.0% (-6.8, 22.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.0%	14.3%
		OR (95% CI) [b]	1.69 (0.40, 7.04)	
		p-value [b]	0.4728	
		RR (95% CI) [d]	1.54 (0.46, 5.08)	
		Risk Diff (95% CI) [e]	7.7% (-11.9, 27.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 552 von 1722

Table 14.2.18.7.1.9.1 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	6	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.4%	18.8%
		OR (95% CI) [b]	1.18 (0.25, 5.55)	
		p-value [b]	0.8323	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9433	
		RR (95% CI) [d]	1.14 (0.33, 3.96)	
		Risk Diff (95% CI) [e]	2.7% (-21.8, 27.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.1%
	OR (95% CI) [b]	0.80 (0.16, 4.00)		
	p-value [b]	0.7859		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7810		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.1 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.85 (0.26, 2.79)	
		Risk Diff (95% CI) [e]	-4.2% (-35.1, 26.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.1 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	22
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.9%	20.7%
		OR (95% CI) [b]	0.57 (0.18, 1.77)	
		p-value [b]	0.3267	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3346	
		RR (95% CI) [d]	0.62 (0.24, 1.59)	
		Risk Diff (95% CI) [e]	-7.8% (-24.5, 8.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.5%
	OR (95% CI) [b]	0.69 (0.21, 2.25)		
	p-value [b]	0.5345		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6134		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 555 von 1722

Table 14.2.18.7.1.9.1 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.31, 1.84)	
		Risk Diff (95% CI) [e]	-6.8% (-28.9, 15.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.1 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	40	21
	Number of subjects with clinically meaningful improvement at Week 24	n	10	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.6%	16.1%
		OR (95% CI) [b]	1.27 (0.39, 4.13)	
		p-value [b]	0.6932	
		RR (95% CI) [d]	1.22 (0.46, 3.23)	
		Risk Diff (95% CI) [e]	3.5% (-13.4, 20.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	23.8%
		OR (95% CI) [b]	1.07 (0.31, 3.66)	
		p-value [b]	0.9183	
		RR (95% CI) [d]	1.05 (0.41, 2.67)	
		Risk Diff (95% CI) [e]	1.2% (-21.4, 23.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.1 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	12.5%
		OR (95% CI) [b]	1.52 (0.26, 8.93)	
		p-value [b]	0.6419	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5323	
		RR (95% CI) [d]	1.43 (0.31, 6.53)	
		Risk Diff (95% CI) [e]	5.4% (-16.2, 26.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.2%	18.2%
	OR (95% CI) [b]	1.07 (0.17, 6.59)		
	p-value [b]	0.9407		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6346		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.1 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.06 (0.24, 4.65)	
		Risk Diff (95% CI) [e]	1.0% (-26.3, 28.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.1 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	42	22
	Number of subjects with clinically meaningful improvement at Week 24	n	8	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.4%	20.7%
		OR (95% CI) [b]	0.49 (0.15, 1.58)	
		p-value [b]	0.2349	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5563	
		RR (95% CI) [d]	0.55 (0.21, 1.45)	
		Risk Diff (95% CI) [e]	-9.3% (-25.8, 7.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.0%	27.3%
		OR (95% CI) [b]	0.63 (0.19, 2.11)	
		p-value [b]	0.4517	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9837		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 560 von 1722

Table 14.2.18.7.1.10.1 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.70 (0.28, 1.76)	
		Risk Diff (95% CI) [e]	-8.2% (-30.3, 13.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.1 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	39	21
	Number of subjects with clinically meaningful improvement at Week 24	n	11	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.6%	25.8%
		OR (95% CI) [b]	0.79 (0.28, 2.25)	
		p-value [b]	0.6595	
		RR (95% CI) [d]	0.84 (0.38, 1.85)	
		Risk Diff (95% CI) [e]	-4.2% (-23.3, 14.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.2%	38.1%
		OR (95% CI) [b]	0.64 (0.21, 1.96)	
		p-value [b]	0.4336	
		RR (95% CI) [d]	0.74 (0.35, 1.55)	
		Risk Diff (95% CI) [e]	-9.9% (-35.0, 15.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 562 von 1722

Table 14.2.18.7.1.11.1 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	11	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	39.3%	43.8%
		OR (95% CI) [b]	0.83 (0.24, 2.89)	
		p-value [b]	0.7721	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8838	
		RR (95% CI) [d]	0.90 (0.44, 1.85)	
		Risk Diff (95% CI) [e]	-4.5% (-34.8, 25.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	42.3%	63.6%
	OR (95% CI) [b]	0.42 (0.10, 1.79)		
	p-value [b]	0.2411		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7322		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.1 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.66 (0.35, 1.25)	
		Risk Diff (95% CI) [e]	-21.3% (-55.5, 12.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.1 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	43	22
	Number of subjects with clinically meaningful improvement at Week 24	n	8	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.4%	13.8%
		OR (95% CI) [b]	0.81 (0.22, 2.92)	
		p-value [b]	0.7432	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9143	
		RR (95% CI) [d]	0.83 (0.27, 2.54)	
		Risk Diff (95% CI) [e]	-2.4% (-17.0, 12.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.6%
	OR (95% CI) [b]	1.03 (0.27, 3.88)		
	p-value [b]	0.9668		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5269		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.1 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.02 (0.35, 3.03)	
		Risk Diff (95% CI) [e]	0.4% (-19.5, 20.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 566 von 1722

Table 14.2.18.7.1.11.1 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	40	21
	Number of subjects with clinically meaningful improvement at Week 24	n	9	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.6%	22.6%
		OR (95% CI) [b]	0.73 (0.24, 2.22)	
		p-value [b]	0.5854	
		RR (95% CI) [d]	0.78 (0.32, 1.89)	
		Risk Diff (95% CI) [e]	-4.9% (-23.0, 13.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.5%	33.3%
		OR (95% CI) [b]	0.58 (0.18, 1.87)	
		p-value [b]	0.3634	
		RR (95% CI) [d]	0.68 (0.29, 1.56)	
		Risk Diff (95% CI) [e]	-10.8% (-34.8, 13.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 567 von 1722

Table 14.2.18.7.1.12.1 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	25	10
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	12.5%
		OR (95% CI) [b]	1.52 (0.26, 8.93)	
		p-value [b]	0.6419	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5809	
		RR (95% CI) [d]	1.43 (0.31, 6.53)	
		Risk Diff (95% CI) [e]	5.4% (-16.2, 26.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%
		OR (95% CI) [b]	1.00 (0.16, 6.25)	
		p-value [b]	1.0000	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7154	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.1 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.00 (0.23, 4.33)	
		Risk Diff (95% CI) [e]	0.0% (-29.3, 29.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.1 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	22
	Number of subjects with clinically meaningful improvement at Week 24	n	10	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	6.9%
		OR (95% CI) [b]	2.25 (0.46, 10.97)	
		p-value [b]	0.3159	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3307	
		RR (95% CI) [d]	2.07 (0.48, 8.88)	
		Risk Diff (95% CI) [e]	7.4% (-4.9, 19.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.7%
		OR (95% CI) [b]	2.94 (0.58, 14.79)	
		p-value [b]	0.1906	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1564	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.1 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.50 (0.60, 10.44)	
		Risk Diff (95% CI) [e]	13.6% (-3.6, 30.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.1 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.7%	16.1%
		OR (95% CI) [b]	0.83 (0.24, 2.88)	
		p-value [b]	0.7655	
		RR (95% CI) [d]	0.85 (0.30, 2.45)	
		Risk Diff (95% CI) [e]	-2.4% (-18.4, 13.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.1%	23.8%
		OR (95% CI) [b]	0.66 (0.18, 2.40)	
		p-value [b]	0.5268	
		RR (95% CI) [d]	0.72 (0.26, 1.99)	
		Risk Diff (95% CI) [e]	-6.7% (-28.3, 14.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 572 von 1722

Table 14.2.18.7.1.13.1 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	12.5%
		OR (95% CI) [b]	1.52 (0.26, 8.93)	
		p-value [b]	0.6419	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6755	
		RR (95% CI) [d]	1.43 (0.31, 6.53)	
		Risk Diff (95% CI) [e]	5.4% (-16.2, 26.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.2%
	OR (95% CI) [b]	1.07 (0.17, 6.59)		
	p-value [b]	0.9407		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7461		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.1 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.06 (0.24, 4.65)	
		Risk Diff (95% CI) [e]	1.0% (-26.3, 28.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 574 von 1722

Table 14.2.18.7.1.13.1 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	21
	Number of subjects with clinically meaningful improvement at Week 24	n	13	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.6%	10.3%
		OR (95% CI) [b]	1.98 (0.52, 7.54)	
		p-value [b]	0.3184	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4165	
		RR (95% CI) [d]	1.80 (0.55, 5.83)	
		Risk Diff (95% CI) [e]	8.2% (-6.1, 22.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.5%	14.3%
		OR (95% CI) [b]	2.52 (0.63, 10.03)	
		p-value [b]	0.1911	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1815	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 575 von 1722

Table 14.2.18.7.1.13.1 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.07 (0.66, 6.48)	
		Risk Diff (95% CI) [e]	15.3% (-4.9, 35.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 576 von 1722

Table 14.2.18.7.1.13.1 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	13	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.5%	25.8%
		OR (95% CI) [b]	0.98 (0.35, 2.73)	
		p-value [b]	0.9746	
		RR (95% CI) [d]	0.99 (0.46, 2.11)	
		Risk Diff (95% CI) [e]	-0.3% (-19.8, 19.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.7%	38.1%
		OR (95% CI) [b]	0.75 (0.25, 2.26)	
		p-value [b]	0.6154	
		RR (95% CI) [d]	0.83 (0.41, 1.69)	
		Risk Diff (95% CI) [e]	-6.4% (-31.6, 18.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 577 von 1722

Table 14.2.18.7.1.14.1 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	0	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	6.3%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9665	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9674	
		RR (95% CI) [d]	0.20 (0.01, 4.53)	
		Risk Diff (95% CI) [e]	-7.1% (-21.4, 7.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	9.1%
	OR (95% CI) [b]	NE (NE, NE)		
	p-value [b]	0.9622		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9639		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 578 von 1722

Table 14.2.18.7.1.14.1 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.15 (0.01, 3.38)	
		Risk Diff (95% CI) [e]	-10.6% (-30.0, 8.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 579 von 1722

Table 14.2.18.7.1.14.1 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	44	21
	Number of subjects with clinically meaningful improvement at Week 24	n	6	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.6%	3.4%
		OR (95% CI) [b]	2.63 (0.30, 22.83)	
		p-value [b]	0.3819	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3320	
		RR (95% CI) [d]	2.49 (0.31, 19.74)	
		Risk Diff (95% CI) [e]	5.1% (-4.2, 14.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.6%	4.8%
		OR (95% CI) [b]	3.16 (0.36, 28.08)	
		p-value [b]	0.3024	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2087		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 580 von 1722

Table 14.2.18.7.1.14.1 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.86 (0.37, 22.29)	
		Risk Diff (95% CI) [e]	8.9% (-4.8, 22.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 581 von 1722

Table 14.2.18.7.1.14.1 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	41	21
	Number of subjects with clinically meaningful improvement at Week 24	n	5	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.8%	12.9%
		OR (95% CI) [b]	0.73 (0.18, 2.97)	
		p-value [b]	0.6642	
		RR (95% CI) [d]	0.76 (0.22, 2.62)	
		Risk Diff (95% CI) [e]	-3.1% (-17.4, 11.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.2%	19.0%
		OR (95% CI) [b]	0.59 (0.14, 2.48)	
		p-value [b]	0.4717	
		RR (95% CI) [d]	0.64 (0.19, 2.14)	
		Risk Diff (95% CI) [e]	-6.9% (-26.4, 12.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 582 von 1722

Table 14.2.18.7.1.15.1 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects at baseline	N	28	16
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	5	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	6.3%
		OR (95% CI) [b]	3.26 (0.35, 30.74)	
		p-value [b]	0.3017	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1878	
		RR (95% CI) [d]	2.86 (0.37, 22.36)	
		Risk Diff (95% CI) [e]	11.6% (-6.9, 30.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.2%	9.1%
	OR (95% CI) [b]	2.38 (0.24, 23.17)		
	p-value [b]	0.4549		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2236		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.1 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.12 (0.28, 16.07)	
		Risk Diff (95% CI) [e]	10.1% (-12.6, 32.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 584 von 1722

Table 14.2.18.7.1.15.1 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects at baseline	N	70	29
	Number of HRQoL evaluable subjects at Week 24	N	43	21
	Number of subjects with clinically meaningful improvement at Week 24	n	3	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.3%	10.3%
		OR (95% CI) [b]	0.39 (0.07, 2.05)	
		p-value [b]	0.2648	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7295	
		RR (95% CI) [d]	0.41 (0.09, 1.93)	
		Risk Diff (95% CI) [e]	-6.1% (-18.1, 6.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.0%	14.3%
		OR (95% CI) [b]	0.45 (0.08, 2.45)	
		p-value [b]	0.3557	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9888		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 585 von 1722

Table 14.2.18.7.1.15.1 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.49 (0.11, 2.22)	
		Risk Diff (95% CI) [e]	-7.3% (-24.1, 9.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 586 von 1722

Table 14.2.18.7.1.15.1 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects at baseline	N	51	31
	Number of HRQoL evaluable subjects at Week 24	N	40	21
	Number of subjects with clinically meaningful improvement at Week 24	n	5	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.8%	16.1%
		OR (95% CI) [b]	0.57 (0.15, 2.14)	
		p-value [b]	0.4003	
		RR (95% CI) [d]	0.61 (0.19, 1.93)	
		Risk Diff (95% CI) [e]	-6.3% (-21.6, 9.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.5%	23.8%
		OR (95% CI) [b]	0.46 (0.12, 1.81)	
		p-value [b]	0.2640	
		RR (95% CI) [d]	0.53 (0.17, 1.61)	
		Risk Diff (95% CI) [e]	-11.3% (-32.2, 9.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 587 von 1722

Table 14.2.18.7.1.1.2 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	34
	Number of subjects with clinically meaningful improvement at Week 24	n	18	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.4%	22.0%
		OR (95% CI) [b]	0.85 (0.37, 1.98)	
		p-value [b]	0.7077	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0578	
		RR (95% CI) [d]	0.88 (0.45, 1.71)	
		Risk Diff (95% CI) [e]	-2.6% (-16.7, 11.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.9%	32.4%
		OR (95% CI) [b]	0.77 (0.31, 1.89)	
		p-value [b]	0.5651	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0513	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 588 von 1722

Table 14.2.18.7.1.1.2 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.83 (0.44, 1.55)	
		Risk Diff (95% CI) [e]	-5.5% (-24.5, 13.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.2 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	13	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.2%	3.8%
		OR (95% CI) [b]	7.56 (0.93, 61.26)	
		p-value [b]	0.0582	
		RR (95% CI) [d]	6.04 (0.83, 43.72)	
		Risk Diff (95% CI) [e]	19.4% (6.1, 32.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.5%	5.3%
		OR (95% CI) [b]	7.55 (0.91, 62.57)	
		p-value [b]	0.0611	
		RR (95% CI) [d]	5.61 (0.79, 39.91)	
		Risk Diff (95% CI) [e]	24.3% (7.5, 41.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 590 von 1722

Table 14.2.18.7.1.2.2 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	35
	Number of subjects with clinically meaningful improvement at Week 24	n	12	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.9%	20.0%
		OR (95% CI) [b]	0.59 (0.24, 1.49)	
		p-value [b]	0.2653	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8456	
		RR (95% CI) [d]	0.65 (0.30, 1.39)	
		Risk Diff (95% CI) [e]	-7.1% (-20.1, 5.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.9%	28.6%
		OR (95% CI) [b]	0.55 (0.21, 1.43)	
		p-value [b]	0.2174	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9412	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 591 von 1722

Table 14.2.18.7.1.2.2 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.63 (0.30, 1.30)	
		Risk Diff (95% CI) [e]	-10.7% (-28.2, 6.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.2 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	13	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.2%	30.8%
		OR (95% CI) [b]	0.68 (0.24, 1.92)	
		p-value [b]	0.4671	
		RR (95% CI) [d]	0.75 (0.36, 1.59)	
		Risk Diff (95% CI) [e]	-7.6% (-28.5, 13.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.5%	42.1%
		OR (95% CI) [b]	0.58 (0.19, 1.76)	
		p-value [b]	0.3342	
		RR (95% CI) [d]	0.70 (0.35, 1.41)	
		Risk Diff (95% CI) [e]	-12.6% (-38.5, 13.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 593 von 1722

Table 14.2.18.7.1.3.2 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	35
	Number of subjects with clinically meaningful improvement at Week 24	n	20	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.5%	24.0%
		OR (95% CI) [b]	0.87 (0.38, 1.96)	
		p-value [b]	0.7330	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8014	
		RR (95% CI) [d]	0.90 (0.48, 1.68)	
		Risk Diff (95% CI) [e]	-2.5% (-17.0, 12.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.9%	34.3%
		OR (95% CI) [b]	0.82 (0.34, 1.95)	
		p-value [b]	0.6469	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7442		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 594 von 1722

Table 14.2.18.7.1.3.2 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.87 (0.48, 1.57)	
		Risk Diff (95% CI) [e]	-4.4% (-23.6, 14.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 595 von 1722

Table 14.2.18.7.1.3.2 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	10	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	23.1%
		OR (95% CI) [b]	0.72 (0.23, 2.27)	
		p-value [b]	0.5799	
		RR (95% CI) [d]	0.77 (0.31, 1.90)	
		Risk Diff (95% CI) [e]	-5.2% (-24.3, 13.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.7%	31.6%
		OR (95% CI) [b]	0.64 (0.19, 2.11)	
		p-value [b]	0.4608	
		RR (95% CI) [d]	0.72 (0.31, 1.70)	
	Risk Diff (95% CI) [e]	-8.9% (-33.1, 15.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 596 von 1722

Table 14.2.18.7.1.4.2 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	34
	Number of subjects with clinically meaningful improvement at Week 24	n	16	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.2%	18.0%
		OR (95% CI) [b]	0.95 (0.38, 2.33)	
		p-value [b]	0.9049	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6943	
		RR (95% CI) [d]	0.96 (0.46, 2.00)	
		Risk Diff (95% CI) [e]	-0.8% (-13.9, 12.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.9%	26.5%
		OR (95% CI) [b]	0.87 (0.34, 2.25)	
		p-value [b]	0.7757	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7035	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.2 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.90 (0.45, 1.83)	
		Risk Diff (95% CI) [e]	-2.6% (-20.6, 15.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 598 von 1722

Table 14.2.18.7.1.4.2 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	13	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.2%	19.2%
		OR (95% CI) [b]	1.27 (0.40, 4.03)	
		p-value [b]	0.6854	
		RR (95% CI) [d]	1.21 (0.48, 3.03)	
		Risk Diff (95% CI) [e]	4.0% (-14.8, 22.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.5%	26.3%
		OR (95% CI) [b]	1.17 (0.35, 3.93)	
		p-value [b]	0.7946	
		RR (95% CI) [d]	1.12 (0.47, 2.71)	
		Risk Diff (95% CI) [e]	3.2% (-20.7, 27.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 599 von 1722

Table 14.2.18.7.1.5.2 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	34
	Number of subjects with clinically meaningful improvement at Week 24	n	12	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.9%	14.0%
		OR (95% CI) [b]	0.91 (0.33, 2.48)	
		p-value [b]	0.8538	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7095	
		RR (95% CI) [d]	0.92 (0.39, 2.19)	
		Risk Diff (95% CI) [e]	-1.1% (-12.9, 10.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.9%	20.6%
		OR (95% CI) [b]	0.84 (0.30, 2.38)	
		p-value [b]	0.7450	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6929	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 600 von 1722

Table 14.2.18.7.1.5.2 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.87 (0.38, 2.01)	
		Risk Diff (95% CI) [e]	-2.7% (-19.1, 13.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 601 von 1722

Table 14.2.18.7.1.5.2 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.7%	15.4%
		OR (95% CI) [b]	0.66 (0.17, 2.57)	
		p-value [b]	0.5495	
		RR (95% CI) [d]	0.70 (0.21, 2.26)	
		Risk Diff (95% CI) [e]	-4.7% (-20.7, 11.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.6%	21.1%
		OR (95% CI) [b]	0.59 (0.15, 2.40)	
		p-value [b]	0.4631	
		RR (95% CI) [d]	0.65 (0.21, 2.04)	
	Risk Diff (95% CI) [e]	-7.4% (-28.4, 13.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 602 von 1722

Table 14.2.18.7.1.6.2 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	34
	Number of subjects with clinically meaningful improvement at Week 24	n	14	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.1%	20.0%
		OR (95% CI) [b]	0.71 (0.29, 1.74)	
		p-value [b]	0.4519	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7332	
		RR (95% CI) [d]	0.75 (0.36, 1.57)	
		Risk Diff (95% CI) [e]	-4.9% (-18.2, 8.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.9%	29.4%
		OR (95% CI) [b]	0.63 (0.25, 1.63)	
		p-value [b]	0.3441	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7468	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.2 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.71 (0.35, 1.43)	
		Risk Diff (95% CI) [e]	-8.5% (-26.7, 9.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 604 von 1722

Table 14.2.18.7.1.6.2 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluatable subjects at baseline	N	56	26
	Number of HRQoL evaluatable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	12	6
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	21.4%	23.1%
		OR (95% CI) [b]	0.91 (0.30, 2.77)	
		p-value [b]	0.8666	
		RR (95% CI) [d]	0.93 (0.39, 2.20)	
		Risk Diff (95% CI) [e]	-1.6% (-21.1, 17.8)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.3%	31.6%
		OR (95% CI) [b]	0.81 (0.25, 2.63)	
		p-value [b]	0.7285	
		RR (95% CI) [d]	0.86 (0.38, 1.96)	
		Risk Diff (95% CI) [e]	-4.3% (-29.0, 20.4)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 605 von 1722

Table 14.2.18.7.1.7.2 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	66	35
	Number of subjects with clinically meaningful improvement at Week 24	n	17	16
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.3%	32.0%
		OR (95% CI) [b]	0.48 (0.21, 1.05)	
		p-value [b]	0.0662	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4053	
		RR (95% CI) [d]	0.57 (0.32, 1.03)	
		Risk Diff (95% CI) [e]	-13.7% (-28.8, 1.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.8%	45.7%
		OR (95% CI) [b]	0.41 (0.17, 0.98)	
		p-value [b]	0.0443	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4453		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 606 von 1722

Table 14.2.18.7.1.7.2 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.56 (0.33, 0.97)	
		Risk Diff (95% CI) [e]	-20.0% (-39.5, -0.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.2 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	15	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	26.8%	30.8%
		OR (95% CI) [b]	0.82 (0.30, 2.29)	
		p-value [b]	0.7089	
		RR (95% CI) [d]	0.87 (0.42, 1.79)	
		Risk Diff (95% CI) [e]	-4.0% (-25.2, 17.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.1%	42.1%
		OR (95% CI) [b]	0.71 (0.24, 2.14)	
		p-value [b]	0.5450	
		RR (95% CI) [d]	0.81 (0.41, 1.58)	
		Risk Diff (95% CI) [e]	-8.0% (-34.3, 18.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 608 von 1722

Table 14.2.18.7.1.8.2 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	35
	Number of subjects with clinically meaningful improvement at Week 24	n	9	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.7%	4.0%
		OR (95% CI) [b]	2.57 (0.53, 12.39)	
		p-value [b]	0.2392	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5173	
		RR (95% CI) [d]	2.42 (0.54, 10.77)	
		Risk Diff (95% CI) [e]	5.7% (-2.4, 13.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.4%	5.7%
		OR (95% CI) [b]	2.56 (0.52, 12.56)	
		p-value [b]	0.2468	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4835	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.2 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.35 (0.54, 10.29)	
		Risk Diff (95% CI) [e]	7.7% (-3.5, 18.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 610 von 1722

Table 14.2.18.7.1.8.2 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	11.5%
		OR (95% CI) [b]	1.28 (0.31, 5.27)	
		p-value [b]	0.7346	
		RR (95% CI) [d]	1.24 (0.36, 4.29)	
		Risk Diff (95% CI) [e]	2.7% (-12.6, 18.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.2%	15.8%
		OR (95% CI) [b]	1.19 (0.28, 5.06)	
		p-value [b]	0.8186	
		RR (95% CI) [d]	1.15 (0.34, 3.87)	
		Risk Diff (95% CI) [e]	2.4% (-17.6, 22.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 611 von 1722

Table 14.2.18.7.1.9.2 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	66	35
	Number of subjects with clinically meaningful improvement at Week 24	n	13	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.0%	12.0%
		OR (95% CI) [b]	1.19 (0.42, 3.35)	
		p-value [b]	0.7398	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3774	
		RR (95% CI) [d]	1.16 (0.47, 2.88)	
		Risk Diff (95% CI) [e]	2.0% (-9.5, 13.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.7%	17.1%
		OR (95% CI) [b]	1.19 (0.41, 3.45)	
		p-value [b]	0.7548	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2932	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 612 von 1722

Table 14.2.18.7.1.9.2 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.15 (0.48, 2.76)	
		Risk Diff (95% CI) [e]	2.6% (-13.2, 18.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.2 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	12	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.4%	30.8%
		OR (95% CI) [b]	0.61 (0.21, 1.75)	
		p-value [b]	0.3617	
		RR (95% CI) [d]	0.70 (0.32, 1.50)	
		Risk Diff (95% CI) [e]	-9.3% (-30.1, 11.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.3%	42.1%
		OR (95% CI) [b]	0.52 (0.17, 1.59)	
		p-value [b]	0.2492	
		RR (95% CI) [d]	0.65 (0.32, 1.32)	
		Risk Diff (95% CI) [e]	-14.8% (-40.6, 11.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 614 von 1722

Table 14.2.18.7.1.10.2 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	63	35
	Number of subjects with clinically meaningful improvement at Week 24	n	14	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.1%	18.0%
		OR (95% CI) [b]	0.81 (0.32, 2.02)	
		p-value [b]	0.6478	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6688	
		RR (95% CI) [d]	0.84 (0.39, 1.79)	
		Risk Diff (95% CI) [e]	-2.9% (-15.8, 9.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	25.7%
		OR (95% CI) [b]	0.83 (0.32, 2.16)	
		p-value [b]	0.6961	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5235	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 615 von 1722

Table 14.2.18.7.1.10.2 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.86 (0.42, 1.79)	
		Risk Diff (95% CI) [e]	-3.5% (-21.2, 14.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 616 von 1722

Table 14.2.18.7.1.10.2 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	10	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	26.9%
		OR (95% CI) [b]	0.59 (0.20, 1.78)	
		p-value [b]	0.3490	
		RR (95% CI) [d]	0.66 (0.28, 1.55)	
		Risk Diff (95% CI) [e]	-9.1% (-28.8, 10.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.7%	36.8%
		OR (95% CI) [b]	0.50 (0.16, 1.62)	
		p-value [b]	0.2508	
		RR (95% CI) [d]	0.62 (0.28, 1.38)	
		Risk Diff (95% CI) [e]	-14.1% (-39.1, 10.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 617 von 1722

Table 14.2.18.7.1.11.2 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	66	35
	Number of subjects with clinically meaningful improvement at Week 24	n	16	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.2%	22.0%
		OR (95% CI) [b]	0.74 (0.31, 1.74)	
		p-value [b]	0.4857	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9950	
		RR (95% CI) [d]	0.78 (0.39, 1.55)	
		Risk Diff (95% CI) [e]	-4.8% (-18.6, 9.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.2%	31.4%
		OR (95% CI) [b]	0.70 (0.28, 1.73)	
		p-value [b]	0.4385	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9456	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 618 von 1722

Table 14.2.18.7.1.11.2 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.77 (0.40, 1.48)	
		Risk Diff (95% CI) [e]	-7.2% (-25.7, 11.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 619 von 1722

Table 14.2.18.7.1.11.2 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	43	19
	Number of subjects with clinically meaningful improvement at Week 24	n	12	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.4%	26.9%
		OR (95% CI) [b]	0.74 (0.25, 2.17)	
		p-value [b]	0.5835	
		RR (95% CI) [d]	0.80 (0.35, 1.79)	
		Risk Diff (95% CI) [e]	-5.5% (-25.6, 14.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.9%	36.8%
		OR (95% CI) [b]	0.66 (0.21, 2.09)	
		p-value [b]	0.4829	
		RR (95% CI) [d]	0.76 (0.35, 1.62)	
		Risk Diff (95% CI) [e]	-8.9% (-34.4, 16.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 620 von 1722

Table 14.2.18.7.1.14.2 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	34
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.5%	10.0%
		OR (95% CI) [b]	0.73 (0.22, 2.44)	
		p-value [b]	0.6121	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4573	
		RR (95% CI) [d]	0.75 (0.25, 2.25)	
		Risk Diff (95% CI) [e]	-2.5% (-12.4, 7.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.4%	14.7%
		OR (95% CI) [b]	0.68 (0.20, 2.32)	
		p-value [b]	0.5338	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4563	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 621 von 1722

Table 14.2.18.7.1.14.2 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.71 (0.24, 2.07)	
		Risk Diff (95% CI) [e]	-4.3% (-18.2, 9.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 622 von 1722

Table 14.2.18.7.1.14.2 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	4	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.1%	3.8%
		OR (95% CI) [b]	1.92 (0.20, 18.11)	
		p-value [b]	0.5677	
		RR (95% CI) [d]	1.86 (0.22, 15.81)	
		Risk Diff (95% CI) [e]	3.3% (-6.7, 13.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	9.1%	5.3%
		OR (95% CI) [b]	1.80 (0.19, 17.26)	
		p-value [b]	0.6104	
		RR (95% CI) [d]	1.73 (0.21, 14.45)	
		Risk Diff (95% CI) [e]	3.8% (-9.3, 17.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 623 von 1722

Table 14.2.18.7.1.12.2 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	35
	Number of subjects with clinically meaningful improvement at Week 24	n	11	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.8%	12.0%
		OR (95% CI) [b]	0.98 (0.34, 2.84)	
		p-value [b]	0.9758	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4662	
		RR (95% CI) [d]	0.99 (0.39, 2.51)	
		Risk Diff (95% CI) [e]	-0.2% (-11.3, 11.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.4%	17.1%
		OR (95% CI) [b]	0.95 (0.32, 2.83)	
		p-value [b]	0.9257	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5153	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.2 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.96 (0.39, 2.37)	
		Risk Diff (95% CI) [e]	-0.7% (-16.0, 14.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 625 von 1722

Table 14.2.18.7.1.12.2 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	43	18
	Number of subjects with clinically meaningful improvement at Week 24	n	11	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.6%	11.5%
		OR (95% CI) [b]	1.87 (0.48, 7.39)	
		p-value [b]	0.3695	
		RR (95% CI) [d]	1.70 (0.52, 5.59)	
		Risk Diff (95% CI) [e]	8.1% (-8.0, 24.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.6%	16.7%
		OR (95% CI) [b]	1.72 (0.42, 7.08)	
		p-value [b]	0.4536	
		RR (95% CI) [d]	1.53 (0.49, 4.86)	
		Risk Diff (95% CI) [e]	8.9% (-12.7, 30.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 626 von 1722

Table 14.2.18.7.1.13.2 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	34
	Number of subjects with clinically meaningful improvement at Week 24	n	19	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.4%	18.0%
		OR (95% CI) [b]	1.17 (0.49, 2.82)	
		p-value [b]	0.7271	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7490	
		RR (95% CI) [d]	1.14 (0.56, 2.32)	
		Risk Diff (95% CI) [e]	2.4% (-11.0, 15.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.4%	26.5%
		OR (95% CI) [b]	1.10 (0.43, 2.78)	
		p-value [b]	0.8413	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7614	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 627 von 1722

Table 14.2.18.7.1.13.2 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.07 (0.54, 2.11)	
		Risk Diff (95% CI) [e]	1.9% (-16.5, 20.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 628 von 1722

Table 14.2.18.7.1.13.2 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	44	19
	Number of subjects with clinically meaningful improvement at Week 24	n	12	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.4%	15.4%
		OR (95% CI) [b]	1.50 (0.43, 5.19)	
		p-value [b]	0.5224	
		RR (95% CI) [d]	1.39 (0.50, 3.91)	
		Risk Diff (95% CI) [e]	6.0% (-11.5, 23.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.3%	21.1%
		OR (95% CI) [b]	1.41 (0.39, 5.09)	
		p-value [b]	0.6038	
		RR (95% CI) [d]	1.30 (0.48, 3.51)	
		Risk Diff (95% CI) [e]	6.2% (-16.3, 28.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 629 von 1722

Table 14.2.18.7.1.15.2 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects at baseline	N	93	50
	Number of HRQoL evaluable subjects at Week 24	N	67	34
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.7%	12.0%
		OR (95% CI) [b]	0.79 (0.26, 2.35)	
		p-value [b]	0.6661	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7695	
		RR (95% CI) [d]	0.81 (0.30, 2.14)	
		Risk Diff (95% CI) [e]	-2.3% (-13.2, 8.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.4%	17.6%
		OR (95% CI) [b]	0.72 (0.23, 2.24)	
		p-value [b]	0.5746	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7994	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.2 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.76 (0.30, 1.96)	
		Risk Diff (95% CI) [e]	-4.2% (-19.4, 11.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 631 von 1722

Table 14.2.18.7.1.15.2 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects at baseline	N	56	26
	Number of HRQoL evaluable subjects at Week 24	N	42	19
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.1%	11.5%
		OR (95% CI) [b]	0.59 (0.12, 2.85)	
		p-value [b]	0.5112	
		RR (95% CI) [d]	0.62 (0.15, 2.57)	
		Risk Diff (95% CI) [e]	-4.4% (-18.4, 9.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	9.5%	15.8%
		OR (95% CI) [b]	0.56 (0.11, 2.80)	
		p-value [b]	0.4813	
		RR (95% CI) [d]	0.60 (0.15, 2.43)	
		Risk Diff (95% CI) [e]	-6.3% (-24.9, 12.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 632 von 1722

Table 14.2.18.7.1.1.4 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	8	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.8%	26.3%
		OR (95% CI) [b]	0.97 (0.27, 3.57)	
		p-value [b]	0.9682	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4786	
		RR (95% CI) [d]	0.98 (0.38, 2.56)	
		Risk Diff (95% CI) [e]	-0.5% (-25.6, 24.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.6%	33.3%
		OR (95% CI) [b]	0.76 (0.20, 2.93)	
		p-value [b]	0.6924	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3310		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 633 von 1722

Table 14.2.18.7.1.1.4 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.83 (0.33, 2.09)	
		Risk Diff (95% CI) [e]	-5.7% (-34.6, 23.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 634 von 1722

Table 14.2.18.7.1.1.4 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	38
	Number of subjects with clinically meaningful improvement at Week 24	n	23	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.5%	12.3%
		OR (95% CI) [b]	1.73 (0.69, 4.31)	
		p-value [b]	0.2395	
		RR (95% CI) [d]	1.59 (0.72, 3.48)	
		Risk Diff (95% CI) [e]	7.2% (-3.9, 18.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.0%	18.4%
		OR (95% CI) [b]	1.73 (0.67, 4.47)	
		p-value [b]	0.2606	
		RR (95% CI) [d]	1.52 (0.72, 3.23)	
		Risk Diff (95% CI) [e]	9.6% (-6.1, 25.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 635 von 1722

Table 14.2.18.7.1.2.4 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	8	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.8%	31.6%
		OR (95% CI) [b]	0.75 (0.21, 2.65)	
		p-value [b]	0.6595	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8169	
		RR (95% CI) [d]	0.82 (0.34, 1.99)	
		Risk Diff (95% CI) [e]	-5.8% (-31.7, 20.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.6%	40.0%
		OR (95% CI) [b]	0.57 (0.15, 2.13)	
		p-value [b]	0.4044	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9708		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 636 von 1722

Table 14.2.18.7.1.2.4 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.69 (0.29, 1.62)	
		Risk Diff (95% CI) [e]	-12.4% (-42.1, 17.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.4 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	39
	Number of subjects with clinically meaningful improvement at Week 24	n	17	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.4%	21.1%
		OR (95% CI) [b]	0.63 (0.28, 1.43)	
		p-value [b]	0.2704	
		RR (95% CI) [d]	0.68 (0.35, 1.33)	
		Risk Diff (95% CI) [e]	-6.6% (-19.0, 5.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.7%	30.8%
		OR (95% CI) [b]	0.59 (0.25, 1.40)	
		p-value [b]	0.2293	
		RR (95% CI) [d]	0.67 (0.36, 1.27)	
		Risk Diff (95% CI) [e]	-10.0% (-27.0, 6.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 638 von 1722

Table 14.2.18.7.1.3.4 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	9	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	29.0%	21.1%
		OR (95% CI) [b]	1.53 (0.40, 5.91)	
		p-value [b]	0.5339	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2906	
		RR (95% CI) [d]	1.38 (0.49, 3.86)	
		Risk Diff (95% CI) [e]	8.0% (-16.3, 32.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.0%	26.7%
		OR (95% CI) [b]	1.24 (0.31, 4.96)	
		p-value [b]	0.7636	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3953	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 639 von 1722

Table 14.2.18.7.1.3.4 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.16 (0.43, 3.16)	
		Risk Diff (95% CI) [e]	4.4% (-23.6, 32.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.4 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	39
	Number of subjects with clinically meaningful improvement at Week 24	n	21	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.8%	24.6%
		OR (95% CI) [b]	0.66 (0.31, 1.43)	
		p-value [b]	0.2962	
		RR (95% CI) [d]	0.72 (0.40, 1.32)	
		Risk Diff (95% CI) [e]	-6.8% (-19.9, 6.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.6%	35.9%
		OR (95% CI) [b]	0.61 (0.27, 1.40)	
		p-value [b]	0.2454	
		RR (95% CI) [d]	0.71 (0.41, 1.25)	
		Risk Diff (95% CI) [e]	-10.3% (-28.1, 7.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 641 von 1722

Table 14.2.18.7.1.4.4 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.4%	21.1%
		OR (95% CI) [b]	0.90 (0.22, 3.72)	
		p-value [b]	0.8842	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7790	
		RR (95% CI) [d]	0.92 (0.30, 2.84)	
		Risk Diff (95% CI) [e]	-1.7% (-24.7, 21.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.7%	26.7%
		OR (95% CI) [b]	0.72 (0.17, 3.07)	
		p-value [b]	0.6547	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6274		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.4 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.78 (0.26, 2.33)	
		Risk Diff (95% CI) [e]	-6.0% (-32.8, 20.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.4 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	38
	Number of subjects with clinically meaningful improvement at Week 24	n	23	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.5%	17.5%
		OR (95% CI) [b]	1.14 (0.50, 2.59)	
		p-value [b]	0.7577	
		RR (95% CI) [d]	1.11 (0.57, 2.17)	
		Risk Diff (95% CI) [e]	1.9% (-10.2, 14.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.0%	26.3%
		OR (95% CI) [b]	1.09 (0.46, 2.60)	
	p-value [b]	0.8432		
	RR (95% CI) [d]	1.07 (0.56, 2.01)		
	Risk Diff (95% CI) [e]	1.7% (-15.3, 18.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 644 von 1722

Table 14.2.18.7.1.5.4 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	3	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.7%	21.1%
		OR (95% CI) [b]	0.40 (0.08, 2.04)	
		p-value [b]	0.2708	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3226	
		RR (95% CI) [d]	0.46 (0.12, 1.83)	
		Risk Diff (95% CI) [e]	-11.4% (-32.5, 9.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.3%	26.7%
		OR (95% CI) [b]	0.32 (0.06, 1.66)	
		p-value [b]	0.1739	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2472		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 645 von 1722

Table 14.2.18.7.1.5.4 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.39 (0.10, 1.51)	
		Risk Diff (95% CI) [e]	-16.3% (-41.3, 8.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.4 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	38
	Number of subjects with clinically meaningful improvement at Week 24	n	15	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.7%	12.3%
		OR (95% CI) [b]	1.04 (0.40, 2.71)	
		p-value [b]	0.9357	
		RR (95% CI) [d]	1.04 (0.45, 2.40)	
		Risk Diff (95% CI) [e]	0.4% (-10.0, 10.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.3%	18.4%
		OR (95% CI) [b]	0.99 (0.37, 2.68)	
	p-value [b]	0.9865		
	RR (95% CI) [d]	0.99 (0.44, 2.23)		
	Risk Diff (95% CI) [e]	-0.1% (-15.0, 14.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 647 von 1722

Table 14.2.18.7.1.6.4 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.1%	15.8%
		OR (95% CI) [b]	1.03 (0.22, 4.89)	
		p-value [b]	0.9747	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7054	
		RR (95% CI) [d]	1.02 (0.27, 3.80)	
		Risk Diff (95% CI) [e]	0.3% (-20.6, 21.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.2%	20.0%
		OR (95% CI) [b]	0.83 (0.17, 4.09)	
		p-value [b]	0.8222	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8017		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 648 von 1722

Table 14.2.18.7.1.6.4 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.86 (0.24, 3.13)	
		Risk Diff (95% CI) [e]	-2.8% (-27.2, 21.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 649 von 1722

Table 14.2.18.7.1.6.4 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	38
	Number of subjects with clinically meaningful improvement at Week 24	n	21	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.8%	22.8%
		OR (95% CI) [b]	0.73 (0.34, 1.60)	
		p-value [b]	0.4333	
		RR (95% CI) [d]	0.78 (0.42, 1.44)	
		Risk Diff (95% CI) [e]	-5.0% (-17.9, 7.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.6%	34.2%
		OR (95% CI) [b]	0.66 (0.29, 1.52)	
	p-value [b]	0.3323		
	RR (95% CI) [d]	0.75 (0.42, 1.33)		
	Risk Diff (95% CI) [e]	-8.6% (-26.4, 9.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 650 von 1722

Table 14.2.18.7.1.7.4 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.6%	26.3%
		OR (95% CI) [b]	0.82 (0.22, 3.07)	
		p-value [b]	0.7642	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5847	
		RR (95% CI) [d]	0.86 (0.32, 2.32)	
		Risk Diff (95% CI) [e]	-3.7% (-28.4, 20.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.1%	33.3%
		OR (95% CI) [b]	0.64 (0.16, 2.50)	
		p-value [b]	0.5178	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7066	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 651 von 1722

Table 14.2.18.7.1.7.4 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.28, 1.90)	
		Risk Diff (95% CI) [e]	-9.2% (-37.7, 19.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.4 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	81	39
	Number of subjects with clinically meaningful improvement at Week 24	n	25	19
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.2%	33.3%
		OR (95% CI) [b]	0.54 (0.27, 1.09)	
		p-value [b]	0.0847	
		RR (95% CI) [d]	0.64 (0.38, 1.05)	
		Risk Diff (95% CI) [e]	-12.1% (-26.4, 2.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.9%	48.7%
		OR (95% CI) [b]	0.47 (0.21, 1.03)	
	p-value [b]	0.0594		
	RR (95% CI) [d]	0.63 (0.40, 1.00)		
	Risk Diff (95% CI) [e]	-17.9% (-36.5, 0.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 653 von 1722

Table 14.2.18.7.1.8.4 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.1%	10.5%
		OR (95% CI) [b]	1.63 (0.28, 9.41)	
		p-value [b]	0.5821	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8432	
		RR (95% CI) [d]	1.53 (0.33, 7.13)	
		Risk Diff (95% CI) [e]	5.6% (-13.3, 24.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.2%	13.3%
		OR (95% CI) [b]	1.35 (0.23, 7.98)	
		p-value [b]	0.7375	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7114	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.4 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.29 (0.28, 5.89)	
		Risk Diff (95% CI) [e]	3.9% (-18.1, 25.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 655 von 1722

Table 14.2.18.7.1.8.4 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	39
	Number of subjects with clinically meaningful improvement at Week 24	n	12	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.2%	5.3%
		OR (95% CI) [b]	2.04 (0.55, 7.53)	
		p-value [b]	0.2857	
		RR (95% CI) [d]	1.93 (0.57, 6.58)	
		Risk Diff (95% CI) [e]	4.9% (-3.1, 12.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.6%	7.7%
		OR (95% CI) [b]	2.06 (0.55, 7.76)	
	p-value [b]	0.2869		
	RR (95% CI) [d]	1.90 (0.57, 6.36)		
	Risk Diff (95% CI) [e]	6.9% (-4.4, 18.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 656 von 1722

Table 14.2.18.7.1.9.4 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	8	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.8%	10.5%
		OR (95% CI) [b]	2.96 (0.56, 15.73)	
		p-value [b]	0.2037	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1039	
		RR (95% CI) [d]	2.45 (0.58, 10.35)	
		Risk Diff (95% CI) [e]	15.3% (-5.4, 36.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.6%	13.3%
		OR (95% CI) [b]	2.48 (0.45, 13.51)	
		p-value [b]	0.2950	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1435		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.4 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.07 (0.50, 8.55)	
		Risk Diff (95% CI) [e]	14.3% (-9.4, 37.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 658 von 1722

Table 14.2.18.7.1.9.4 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	81	39
	Number of subjects with clinically meaningful improvement at Week 24	n	17	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.4%	21.1%
		OR (95% CI) [b]	0.63 (0.28, 1.43)	
		p-value [b]	0.2704	
		RR (95% CI) [d]	0.68 (0.35, 1.33)	
		Risk Diff (95% CI) [e]	-6.6% (-19.0, 5.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.0%	30.8%
		OR (95% CI) [b]	0.60 (0.25, 1.42)	
		p-value [b]	0.2435	
		RR (95% CI) [d]	0.68 (0.36, 1.28)	
		Risk Diff (95% CI) [e]	-9.8% (-26.8, 7.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 659 von 1722

Table 14.2.18.7.1.10.4 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	6	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.4%	15.8%
		OR (95% CI) [b]	1.28 (0.28, 5.86)	
		p-value [b]	0.7505	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3967	
		RR (95% CI) [d]	1.23 (0.35, 4.33)	
		Risk Diff (95% CI) [e]	3.6% (-17.9, 25.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.7%	20.0%
		OR (95% CI) [b]	1.04 (0.22, 4.92)	
		p-value [b]	0.9572	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5396	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 660 von 1722

Table 14.2.18.7.1.10.4 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.03 (0.30, 3.57)	
		Risk Diff (95% CI) [e]	0.7% (-24.4, 25.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 661 von 1722

Table 14.2.18.7.1.10.4 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	78	39
	Number of subjects with clinically meaningful improvement at Week 24	n	18	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.3%	22.8%
		OR (95% CI) [b]	0.61 (0.27, 1.35)	
		p-value [b]	0.2225	
		RR (95% CI) [d]	0.67 (0.35, 1.27)	
		Risk Diff (95% CI) [e]	-7.6% (-20.2, 5.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.1%	33.3%
		OR (95% CI) [b]	0.60 (0.26, 1.40)	
		p-value [b]	0.2380	
		RR (95% CI) [d]	0.69 (0.38, 1.26)	
	Risk Diff (95% CI) [e]	-10.3% (-27.8, 7.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 662 von 1722

Table 14.2.18.7.1.11.4 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	7	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.6%	31.6%
		OR (95% CI) [b]	0.63 (0.18, 2.28)	
		p-value [b]	0.4830	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7446	
		RR (95% CI) [d]	0.72 (0.28, 1.81)	
		Risk Diff (95% CI) [e]	-9.0% (-34.6, 16.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.1%	40.0%
		OR (95% CI) [b]	0.48 (0.13, 1.82)	
		p-value [b]	0.2785	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5210		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 663 von 1722

Table 14.2.18.7.1.11.4 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.60 (0.25, 1.48)	
		Risk Diff (95% CI) [e]	-15.9% (-45.1, 13.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 664 von 1722

Table 14.2.18.7.1.11.4 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	80	39
	Number of subjects with clinically meaningful improvement at Week 24	n	21	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.8%	21.1%
		OR (95% CI) [b]	0.81 (0.37, 1.79)	
		p-value [b]	0.6062	
		RR (95% CI) [d]	0.85 (0.45, 1.60)	
		Risk Diff (95% CI) [e]	-3.3% (-15.9, 9.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.3%	30.8%
		OR (95% CI) [b]	0.80 (0.34, 1.86)	
		p-value [b]	0.6056	
		RR (95% CI) [d]	0.85 (0.47, 1.55)	
		Risk Diff (95% CI) [e]	-4.5% (-21.9, 12.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 665 von 1722

Table 14.2.18.7.1.12.4 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	14
	Number of subjects with clinically meaningful improvement at Week 24	n	6	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.4%	5.3%
		OR (95% CI) [b]	4.32 (0.48, 39.00)	
		p-value [b]	0.1929	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2167	
		RR (95% CI) [d]	3.68 (0.48, 28.24)	
		Risk Diff (95% CI) [e]	14.1% (-3.1, 31.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.7%	7.1%
	OR (95% CI) [b]	3.39 (0.37, 31.29)		
	p-value [b]	0.2819		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3041		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.4 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.90 (0.38, 21.81)	
		Risk Diff (95% CI) [e]	13.5% (-6.4, 33.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.4 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluatable subjects at baseline	N	118	57
	Number of HRQoL evaluatable subjects at Week 24	N	81	39
	Number of subjects with clinically meaningful improvement at Week 24	n	16	8
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	13.6%	14.0%
		OR (95% CI) [b]	0.96 (0.38, 2.40)	
		p-value [b]	0.9317	
		RR (95% CI) [d]	0.97 (0.44, 2.12)	
		Risk Diff (95% CI) [e]	-0.5% (-11.4, 10.5)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.8%	20.5%
		OR (95% CI) [b]	0.95 (0.37, 2.47)	
		p-value [b]	0.9224	
		RR (95% CI) [d]	0.96 (0.45, 2.05)	
		Risk Diff (95% CI) [e]	-0.8% (-16.1, 14.6)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 668 von 1722

Table 14.2.18.7.1.13.4 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	7	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.6%	15.8%
		OR (95% CI) [b]	1.56 (0.35, 6.92)	
		p-value [b]	0.5620	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7651	
		RR (95% CI) [d]	1.43 (0.42, 4.87)	
		Risk Diff (95% CI) [e]	6.8% (-15.2, 28.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.1%	20.0%
		OR (95% CI) [b]	1.27 (0.28, 5.84)	
		p-value [b]	0.7569	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9167	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 669 von 1722

Table 14.2.18.7.1.13.4 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.21 (0.36, 4.01)	
		Risk Diff (95% CI) [e]	4.1% (-21.4, 29.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 670 von 1722

Table 14.2.18.7.1.13.4 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	38
	Number of subjects with clinically meaningful improvement at Week 24	n	24	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.3%	17.5%
		OR (95% CI) [b]	1.20 (0.53, 2.72)	
		p-value [b]	0.6617	
		RR (95% CI) [d]	1.16 (0.60, 2.26)	
		Risk Diff (95% CI) [e]	2.8% (-9.5, 15.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.3%	26.3%
		OR (95% CI) [b]	1.16 (0.49, 2.75)	
		p-value [b]	0.7386	
		RR (95% CI) [d]	1.11 (0.59, 2.09)	
		Risk Diff (95% CI) [e]	3.0% (-14.2, 20.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 671 von 1722

Table 14.2.18.7.1.14.4 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.1%	10.5%
		OR (95% CI) [b]	1.63 (0.28, 9.41)	
		p-value [b]	0.5821	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4541	
		RR (95% CI) [d]	1.53 (0.33, 7.13)	
		Risk Diff (95% CI) [e]	5.6% (-13.3, 24.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.2%	13.3%
		OR (95% CI) [b]	1.35 (0.23, 7.98)	
		p-value [b]	0.7375	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5345	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 672 von 1722

Table 14.2.18.7.1.14.4 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.29 (0.28, 5.89)	
		Risk Diff (95% CI) [e]	3.9% (-18.1, 25.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 673 von 1722

Table 14.2.18.7.1.14.4 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	82	38
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.1%	7.0%
		OR (95% CI) [b]	0.71 (0.19, 2.62)	
		p-value [b]	0.6072	
		RR (95% CI) [d]	0.72 (0.21, 2.47)	
		Risk Diff (95% CI) [e]	-1.9% (-9.7, 5.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.3%	10.5%
		OR (95% CI) [b]	0.67 (0.18, 2.53)	
	p-value [b]	0.5561		
	RR (95% CI) [d]	0.70 (0.21, 2.32)		
	Risk Diff (95% CI) [e]	-3.2% (-14.5, 8.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 674 von 1722

Table 14.2.18.7.1.15.4 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects at baseline	N	31	19
	Number of HRQoL evaluable subjects at Week 24	N	29	15
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.9%	15.8%
		OR (95% CI) [b]	0.79 (0.16, 3.99)	
		p-value [b]	0.7756	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9052	
		RR (95% CI) [d]	0.82 (0.20, 3.26)	
		Risk Diff (95% CI) [e]	-2.9% (-23.1, 17.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.8%	20.0%
		OR (95% CI) [b]	0.64 (0.12, 3.32)	
		p-value [b]	0.5955	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9569	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 675 von 1722

Table 14.2.18.7.1.15.4 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.69 (0.18, 2.69)	
		Risk Diff (95% CI) [e]	-6.2% (-30.0, 17.6)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 676 von 1722

Table 14.2.18.7.1.15.4 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects at baseline	N	118	57
	Number of HRQoL evaluable subjects at Week 24	N	80	38
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.6%	10.5%
		OR (95% CI) [b]	0.70 (0.24, 2.08)	
		p-value [b]	0.5225	
		RR (95% CI) [d]	0.72 (0.27, 1.94)	
		Risk Diff (95% CI) [e]	-2.9% (-12.2, 6.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.3%	15.8%
		OR (95% CI) [b]	0.68 (0.22, 2.06)	
		p-value [b]	0.4910	
		RR (95% CI) [d]	0.71 (0.27, 1.86)	
		Risk Diff (95% CI) [e]	-4.5% (-18.0, 9.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 677 von 1722

Table 14.2.18.7.1.1.5 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	51
	Number of subjects with clinically meaningful improvement at Week 24	n	29	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.6%	16.2%
		OR (95% CI) [b]	1.34 (0.64, 2.81)	
		p-value [b]	0.4413	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9556	
		RR (95% CI) [d]	1.27 (0.69, 2.34)	
		Risk Diff (95% CI) [e]	4.4% (-6.4, 15.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.6%	23.5%
	OR (95% CI) [b]	1.24 (0.57, 2.69)		
	p-value [b]	0.5866		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9651		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 678 von 1722

Table 14.2.18.7.1.1.5 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.17 (0.65, 2.10)	
		Risk Diff (95% CI) [e]	4.1% (-10.4, 18.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.5 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	1	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9541	
		RR (95% CI) [d]	1.13 (0.06, 20.71)	
		Risk Diff (95% CI) [e]	2.1% (-48.0, 52.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9643		
	RR (95% CI) [d]	1.50 (0.08, 26.86)		
	Risk Diff (95% CI) [e]	8.3% (-46.2, 62.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 680 von 1722

Table 14.2.18.7.1.2.5 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	52
	Number of subjects with clinically meaningful improvement at Week 24	n	22	18
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.6%	24.3%
		OR (95% CI) [b]	0.58 (0.29, 1.16)	
		p-value [b]	0.1210	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9469	
		RR (95% CI) [d]	0.64 (0.37, 1.12)	
		Risk Diff (95% CI) [e]	-8.7% (-20.2, 2.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.0%
	OR (95% CI) [b]	0.50 (0.24, 1.05)		
	p-value [b]	0.0668		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9585		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 681 von 1722

Table 14.2.18.7.1.2.5 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.61 (0.36, 1.03)	
		Risk Diff (95% CI) [e]	-13.7% (-28.8, 1.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 682 von 1722

Table 14.2.18.7.1.2.5 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	2	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.6%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9497	
		RR (95% CI) [d]	1.88 (0.12, 28.78)	
		Risk Diff (95% CI) [e]	14.6% (-38.4, 67.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	40.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9609		
	RR (95% CI) [d]	2.50 (0.17, 37.26)		
	Risk Diff (95% CI) [e]	25.0% (-32.7, 82.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.5 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	52
	Number of subjects with clinically meaningful improvement at Week 24	n	28	18
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.9%	24.3%
		OR (95% CI) [b]	0.77 (0.39, 1.51)	
		p-value [b]	0.4487	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9460	
		RR (95% CI) [d]	0.82 (0.48, 1.37)	
		Risk Diff (95% CI) [e]	-4.5% (-16.3, 7.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.7%	34.6%
	OR (95% CI) [b]	0.69 (0.34, 1.41)		
	p-value [b]	0.3042		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9580		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 684 von 1722

Table 14.2.18.7.1.3.5 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.77 (0.47, 1.26)	
		Risk Diff (95% CI) [e]	-7.9% (-23.4, 7.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 685 von 1722

Table 14.2.18.7.1.3.5 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	2	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.6%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9474	
		RR (95% CI) [d]	1.88 (0.12, 28.78)	
		Risk Diff (95% CI) [e]	14.6% (-38.4, 67.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	40.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9593		
	RR (95% CI) [d]	2.50 (0.17, 37.26)		
	Risk Diff (95% CI) [e]	25.0% (-32.7, 82.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 686 von 1722

Table 14.2.18.7.1.4.5 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	51
	Number of subjects with clinically meaningful improvement at Week 24	n	27	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	17.6%
		OR (95% CI) [b]	1.11 (0.53, 2.31)	
		p-value [b]	0.7772	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2967	
		RR (95% CI) [d]	1.09 (0.60, 1.98)	
		Risk Diff (95% CI) [e]	1.6% (-9.3, 12.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.7%	25.5%
	OR (95% CI) [b]	1.01 (0.47, 2.18)		
	p-value [b]	0.9760		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4485		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 687 von 1722

Table 14.2.18.7.1.4.5 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.01 (0.57, 1.79)	
		Risk Diff (95% CI) [e]	0.2% (-14.4, 14.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.5 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	1	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	50.0%
		OR (95% CI) [b]	0.17 (0.01, 5.45)	
		p-value [b]	0.3140	
		RR (95% CI) [d]	0.29 (0.03, 2.80)	
		Risk Diff (95% CI) [e]	-35.7% (-100.0, 38.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%	50.0%
		OR (95% CI) [b]	0.25 (0.01, 8.56)	
	p-value [b]	0.4419		
	RR (95% CI) [d]	0.40 (0.04, 3.74)		
	Risk Diff (95% CI) [e]	-30.0% (-100.0, 47.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 689 von 1722

Table 14.2.18.7.1.4.5 (AMNOG)

EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

---

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 690 von 1722

Table 14.2.18.7.1.5.5 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	51
	Number of subjects with clinically meaningful improvement at Week 24	n	18	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.8%	14.9%
		OR (95% CI) [b]	0.84 (0.37, 1.88)	
		p-value [b]	0.6689	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9993	
		RR (95% CI) [d]	0.86 (0.43, 1.72)	
		Risk Diff (95% CI) [e]	-2.1% (-11.9, 7.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.1%	21.6%
	OR (95% CI) [b]	0.75 (0.33, 1.74)		
	p-value [b]	0.5059		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9988		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 691 von 1722

Table 14.2.18.7.1.5.5 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.79 (0.41, 1.56)	
		Risk Diff (95% CI) [e]	-4.4% (-17.8, 9.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.5 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24		0	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
		p-value [b]	1.0000	
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
	p-value [b]	1.0000		
	RR (95% CI) [d]	NE (NE, NE)		
	Risk Diff (95% CI) [e]	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 693 von 1722

Table 14.2.18.7.1.6.5 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	51
	Number of subjects with clinically meaningful improvement at Week 24	n	25	16
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.7%	21.6%
		OR (95% CI) [b]	0.78 (0.39, 1.58)	
		p-value [b]	0.4908	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9529	
		RR (95% CI) [d]	0.82 (0.47, 1.44)	
		Risk Diff (95% CI) [e]	-3.9% (-15.2, 7.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.8%	31.4%
	OR (95% CI) [b]	0.68 (0.33, 1.44)		
	p-value [b]	0.3154		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9630		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 694 von 1722

Table 14.2.18.7.1.6.5 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.76 (0.45, 1.29)	
		Risk Diff (95% CI) [e]	-7.6% (-22.7, 7.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 695 von 1722

Table 14.2.18.7.1.6.5 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	1	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9541	
		RR (95% CI) [d]	1.13 (0.06, 20.71)	
		Risk Diff (95% CI) [e]	2.1% (-48.0, 52.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9643		
	RR (95% CI) [d]	1.50 (0.08, 26.86)		
	Risk Diff (95% CI) [e]	8.3% (-46.2, 62.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 696 von 1722

Table 14.2.18.7.1.7.5 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	104	52
	Number of subjects with clinically meaningful improvement at Week 24	n	29	24
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.6%	32.4%
		OR (95% CI) [b]	0.54 (0.29, 1.02)	
		p-value [b]	0.0569	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9611	
		RR (95% CI) [d]	0.63 (0.40, 1.01)	
		Risk Diff (95% CI) [e]	-11.9% (-24.4, 0.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.9%
	OR (95% CI) [b]	0.45 (0.23, 0.90)		
	p-value [b]	0.0245		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9547		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 697 von 1722

Table 14.2.18.7.1.7.5 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.60 (0.39, 0.93)	
		Risk Diff (95% CI) [e]	-18.3% (-34.3, -2.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 698 von 1722

Table 14.2.18.7.1.7.5 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	2	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.6%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9632	
		RR (95% CI) [d]	1.88 (0.12, 28.78)	
		Risk Diff (95% CI) [e]	14.6% (-38.4, 67.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	40.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9577		
	RR (95% CI) [d]	2.50 (0.17, 37.26)		
	Risk Diff (95% CI) [e]	25.0% (-32.7, 82.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 699 von 1722

Table 14.2.18.7.1.8.5 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	52
	Number of subjects with clinically meaningful improvement at Week 24	n	15	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.6%	6.8%
		OR (95% CI) [b]	1.64 (0.57, 4.71)	
		p-value [b]	0.3559	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9660	
		RR (95% CI) [d]	1.57 (0.60, 4.16)	
		Risk Diff (95% CI) [e]	3.9% (-3.8, 11.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.3%	9.6%
	OR (95% CI) [b]	1.57 (0.54, 4.58)		
	p-value [b]	0.4117		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9600		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 700 von 1722

Table 14.2.18.7.1.8.5 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.49 (0.57, 3.86)	
		Risk Diff (95% CI) [e]	4.7% (-5.8, 15.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 701 von 1722

Table 14.2.18.7.1.8.5 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	1	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9641	
		RR (95% CI) [d]	1.13 (0.06, 20.71)	
		Risk Diff (95% CI) [e]	2.1% (-48.0, 52.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9580		
	RR (95% CI) [d]	1.50 (0.08, 26.86)		
	Risk Diff (95% CI) [e]	8.3% (-46.2, 62.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 702 von 1722

Table 14.2.18.7.1.9.5 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	104	52
	Number of subjects with clinically meaningful improvement at Week 24	n	25	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.7%	18.9%
		OR (95% CI) [b]	0.92 (0.45, 1.91)	
		p-value [b]	0.8299	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9997	
		RR (95% CI) [d]	0.94 (0.52, 1.69)	
		Risk Diff (95% CI) [e]	-1.2% (-12.1, 9.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.0%	26.9%
	OR (95% CI) [b]	0.86 (0.40, 1.84)		
	p-value [b]	0.6950		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9993		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.5 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.89 (0.51, 1.57)	
		Risk Diff (95% CI) [e]	-2.9% (-17.5, 11.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 704 von 1722

Table 14.2.18.7.1.9.5 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24		0	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
		p-value [b]	1.0000	
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
	p-value [b]	1.0000		
	RR (95% CI) [d]	NE (NE, NE)		
	Risk Diff (95% CI) [e]	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 705 von 1722

Table 14.2.18.7.1.10.5 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	101	52
	Number of subjects with clinically meaningful improvement at Week 24	n	21	16
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.9%	21.6%
		OR (95% CI) [b]	0.63 (0.31, 1.31)	
		p-value [b]	0.2166	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9487	
		RR (95% CI) [d]	0.69 (0.38, 1.24)	
		Risk Diff (95% CI) [e]	-6.7% (-17.8, 4.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.8%	30.8%
	OR (95% CI) [b]	0.59 (0.28, 1.26)		
	p-value [b]	0.1745		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9597		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 706 von 1722

Table 14.2.18.7.1.10.5 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.68 (0.39, 1.18)	
		Risk Diff (95% CI) [e]	-10.0% (-24.8, 4.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 707 von 1722

Table 14.2.18.7.1.10.5 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	2	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.6%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9509	
		RR (95% CI) [d]	1.88 (0.12, 28.78)	
		Risk Diff (95% CI) [e]	14.6% (-38.4, 67.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	40.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9615		
	RR (95% CI) [d]	2.50 (0.17, 37.26)		
	Risk Diff (95% CI) [e]	25.0% (-32.7, 82.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 708 von 1722

Table 14.2.18.7.1.11.5 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	103	52
	Number of subjects with clinically meaningful improvement at Week 24	n	27	16
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	21.6%
		OR (95% CI) [b]	0.86 (0.43, 1.72)	
		p-value [b]	0.6669	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8946	
		RR (95% CI) [d]	0.89 (0.51, 1.54)	
		Risk Diff (95% CI) [e]	-2.5% (-13.9, 8.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.2%	30.8%
	OR (95% CI) [b]	0.80 (0.38, 1.67)		
	p-value [b]	0.5501		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9043		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 709 von 1722

Table 14.2.18.7.1.11.5 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.85 (0.51, 1.43)	
		Risk Diff (95% CI) [e]	-4.6% (-19.7, 10.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 710 von 1722

Table 14.2.18.7.1.11.5 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24	n	0	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	100.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.8938	
		RR (95% CI) [d]	0.08 (0.00, 1.15)	
		Risk Diff (95% CI) [e]	-77.1% (-100.0, -31.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	100.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9033	
		RR (95% CI) [d]	0.10 (0.01, 1.49)	
		Risk Diff (95% CI) [e]	-75.0% (-100.0, -27.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 711 von 1722

Table 14.2.18.7.1.11.5 (AMNOG)

EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

---

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 712 von 1722

Table 14.2.18.7.1.12.5 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	104	52
	Number of subjects with clinically meaningful improvement at Week 24	n	22	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.6%	12.2%
		OR (95% CI) [b]	1.34 (0.58, 3.07)	
		p-value [b]	0.4961	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9989	
		RR (95% CI) [d]	1.28 (0.62, 2.64)	
		Risk Diff (95% CI) [e]	3.4% (-6.1, 13.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.2%	17.3%
	OR (95% CI) [b]	1.28 (0.54, 3.03)		
	p-value [b]	0.5710		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9992		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 713 von 1722

Table 14.2.18.7.1.12.5 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.22 (0.61, 2.46)	
		Risk Diff (95% CI) [e]	3.8% (-9.1, 16.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.5 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	1
	Number of subjects with clinically meaningful improvement at Week 24		0	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
		p-value [b]	1.0000	
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
	p-value [b]	1.0000		
	RR (95% CI) [d]	NE (NE, NE)		
	Risk Diff (95% CI) [e]	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 715 von 1722

Table 14.2.18.7.1.13.5 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	51
	Number of subjects with clinically meaningful improvement at Week 24	n	31	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.0%	17.6%
		OR (95% CI) [b]	1.32 (0.64, 2.71)	
		p-value [b]	0.4464	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9987	
		RR (95% CI) [d]	1.25 (0.70, 2.24)	
		Risk Diff (95% CI) [e]	4.4% (-6.6, 15.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.5%
	OR (95% CI) [b]	1.22 (0.57, 2.61)		
	p-value [b]	0.5998		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9990		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 716 von 1722

Table 14.2.18.7.1.13.5 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.16 (0.67, 2.02)	
		Risk Diff (95% CI) [e]	4.0% (-10.8, 18.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 717 von 1722

Table 14.2.18.7.1.13.5 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24		0	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
		p-value [b]	1.0000	
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
	p-value [b]	1.0000		
	RR (95% CI) [d]	NE (NE, NE)		
	Risk Diff (95% CI) [e]	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 718 von 1722

Table 14.2.18.7.1.14.5 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	105	51
	Number of subjects with clinically meaningful improvement at Week 24	n	11	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.8%	8.1%
		OR (95% CI) [b]	0.96 (0.34, 2.71)	
		p-value [b]	0.9369	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9998	
		RR (95% CI) [d]	0.96 (0.37, 2.50)	
		Risk Diff (95% CI) [e]	-0.3% (-7.9, 7.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.5%	11.8%
	OR (95% CI) [b]	0.88 (0.31, 2.52)		
	p-value [b]	0.8087		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9996		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 719 von 1722

Table 14.2.18.7.1.14.5 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.89 (0.35, 2.27)	
		Risk Diff (95% CI) [e]	-1.3% (-11.9, 9.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 720 von 1722

Table 14.2.18.7.1.14.5 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24		0	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
		p-value [b]	1.0000	
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
	p-value [b]	1.0000		
	RR (95% CI) [d]	NE (NE, NE)		
	Risk Diff (95% CI) [e]	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 721 von 1722

Table 14.2.18.7.1.15.5 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects at baseline	N	141	74
	Number of HRQoL evaluable subjects at Week 24	N	103	51
	Number of subjects with clinically meaningful improvement at Week 24	n	13	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.2%	12.2%
		OR (95% CI) [b]	0.73 (0.30, 1.81)	
		p-value [b]	0.5001	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9990	
		RR (95% CI) [d]	0.76 (0.34, 1.69)	
		Risk Diff (95% CI) [e]	-2.9% (-11.8, 5.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.6%
	OR (95% CI) [b]	0.67 (0.27, 1.70)		
	p-value [b]	0.4035		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9985		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 722 von 1722

Table 14.2.18.7.1.15.5 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.33, 1.56)	
		Risk Diff (95% CI) [e]	-5.0% (-17.3, 7.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 723 von 1722

Table 14.2.18.7.1.15.5 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects at baseline	N	7	2
	Number of HRQoL evaluable subjects at Week 24	N	5	2
	Number of subjects with clinically meaningful improvement at Week 24		0	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
		p-value [b]	1.0000	
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
	p-value [b]	1.0000		
	RR (95% CI) [d]	NE (NE, NE)		
	Risk Diff (95% CI) [e]	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 724 von 1722

Table 14.2.18.7.1.1.6 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	10.5%
		OR (95% CI) [b]	1.86 (0.35, 9.96)	
		p-value [b]	0.4688	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1626	
		RR (95% CI) [d]	1.71 (0.39, 7.44)	
		Risk Diff (95% CI) [e]	7.4% (-10.9, 25.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%
	OR (95% CI) [b]	1.40 (0.25, 7.93)		
	p-value [b]	0.7037		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3168		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 725 von 1722

Table 14.2.18.7.1.1.6 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.31 (0.32, 5.45)	
		Risk Diff (95% CI) [e]	5.2% (-20.3, 30.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.6 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	23
	Number of subjects with clinically meaningful improvement at Week 24	n	19	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	31.7%	14.7%
		OR (95% CI) [b]	2.69 (0.90, 8.03)	
		p-value [b]	0.0765	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0320	
		RR (95% CI) [d]	2.15 (0.88, 5.25)	
		Risk Diff (95% CI) [e]	17.0% (0.2, 33.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	42.2%	21.7%
		OR (95% CI) [b]	2.63 (0.83, 8.34)	
		p-value [b]	0.1004	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0561	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 727 von 1722

Table 14.2.18.7.1.1.6 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.94 (0.83, 4.53)	
		Risk Diff (95% CI) [e]	20.5% (-1.7, 42.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.6 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	5	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.0%	21.7%
		OR (95% CI) [b]	0.40 (0.10, 1.55)	
		p-value [b]	0.1850	
		RR (95% CI) [d]	0.46 (0.15, 1.43)	
		Risk Diff (95% CI) [e]	-11.7% (-30.5, 7.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.7%	27.8%
		OR (95% CI) [b]	0.45 (0.11, 1.82)	
	p-value [b]	0.2619		
	RR (95% CI) [d]	0.53 (0.18, 1.59)		
	Risk Diff (95% CI) [e]	-13.1% (-36.9, 10.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 729 von 1722

Table 14.2.18.7.1.2.6 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	7	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	21.1%
		OR (95% CI) [b]	0.82 (0.21, 3.24)	
		p-value [b]	0.7774	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5064	
		RR (95% CI) [d]	0.85 (0.28, 2.56)	
		Risk Diff (95% CI) [e]	-3.1% (-25.0, 18.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%
	OR (95% CI) [b]	0.56 (0.13, 2.42)		
	p-value [b]	0.4376		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8658		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 730 von 1722

Table 14.2.18.7.1.2.6 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.66 (0.23, 1.85)	
		Risk Diff (95% CI) [e]	-11.5% (-41.7, 18.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 731 von 1722

Table 14.2.18.7.1.2.6 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	24
	Number of subjects with clinically meaningful improvement at Week 24	n	14	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.3%	29.4%
		OR (95% CI) [b]	0.73 (0.28, 1.89)	
		p-value [b]	0.5169	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5265	
		RR (95% CI) [d]	0.79 (0.40, 1.59)	
		Risk Diff (95% CI) [e]	-6.1% (-24.8, 12.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.1%
	OR (95% CI) [b]	0.63 (0.23, 1.77)		
	p-value [b]	0.3821		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7462		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 732 von 1722

Table 14.2.18.7.1.2.6 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.39, 1.42)	
		Risk Diff (95% CI) [e]	-10.6% (-34.5, 13.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.6 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	4	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	17.4%
		OR (95% CI) [b]	0.41 (0.09, 1.82)	
		p-value [b]	0.2435	
		RR (95% CI) [d]	0.46 (0.13, 1.68)	
		Risk Diff (95% CI) [e]	-9.4% (-26.6, 7.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.8%	22.2%
		OR (95% CI) [b]	0.47 (0.10, 2.14)	
	p-value [b]	0.3271		
	RR (95% CI) [d]	0.53 (0.15, 1.87)		
	Risk Diff (95% CI) [e]	-10.5% (-32.5, 11.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 734 von 1722

Table 14.2.18.7.1.3.6 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	10.5%
		OR (95% CI) [b]	1.86 (0.35, 9.96)	
		p-value [b]	0.4688	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2198	
		RR (95% CI) [d]	1.71 (0.39, 7.44)	
		Risk Diff (95% CI) [e]	7.4% (-10.9, 25.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%	16.7%
		OR (95% CI) [b]	1.40 (0.25, 7.93)	
		p-value [b]	0.7037	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4121	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.6 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.31 (0.32, 5.45)	
		Risk Diff (95% CI) [e]	5.2% (-20.3, 30.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.6 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	24
	Number of subjects with clinically meaningful improvement at Week 24	n	17	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.3%	32.4%
		OR (95% CI) [b]	0.83 (0.33, 2.06)	
		p-value [b]	0.6824	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5217	
		RR (95% CI) [d]	0.88 (0.47, 1.65)	
		Risk Diff (95% CI) [e]	-4.0% (-23.4, 15.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	37.8%
	OR (95% CI) [b]	0.72 (0.26, 1.96)		
	p-value [b]	0.5170		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7689		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.6 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.82 (0.46, 1.46)	
		Risk Diff (95% CI) [e]	-8.1% (-32.5, 16.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 738 von 1722

Table 14.2.18.7.1.3.6 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	6	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	21.7%
		OR (95% CI) [b]	0.49 (0.13, 1.81)	
		p-value [b]	0.2861	
		RR (95% CI) [d]	0.55 (0.19, 1.62)	
		Risk Diff (95% CI) [e]	-9.7% (-28.9, 9.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.6%	27.8%
		OR (95% CI) [b]	0.56 (0.14, 2.16)	
	p-value [b]	0.3982		
	RR (95% CI) [d]	0.64 (0.22, 1.80)		
	Risk Diff (95% CI) [e]	-10.1% (-34.5, 14.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 739 von 1722

Table 14.2.18.7.1.4.6 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	8	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.5%	21.1%
		OR (95% CI) [b]	0.97 (0.25, 3.73)	
		p-value [b]	0.9620	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9513	
		RR (95% CI) [d]	0.97 (0.34, 2.83)	
		Risk Diff (95% CI) [e]	-0.5% (-22.8, 21.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.16, 2.82)	
		p-value [b]	0.5817	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6573	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.6 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.28, 2.04)	
		Risk Diff (95% CI) [e]	-8.3% (-38.9, 22.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.6 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	23
	Number of subjects with clinically meaningful improvement at Week 24	n	15	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	20.6%
		OR (95% CI) [b]	1.29 (0.47, 3.55)	
		p-value [b]	0.6279	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7055	
		RR (95% CI) [d]	1.21 (0.55, 2.68)	
		Risk Diff (95% CI) [e]	4.4% (-13.0, 21.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	33.3%
	OR (95% CI) [b]	1.14 (0.39, 3.38)		
	p-value [b]	0.8091		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9460		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 742 von 1722

Table 14.2.18.7.1.4.6 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.10 (0.52, 2.30)	
		Risk Diff (95% CI) [e]	2.9% (-20.4, 26.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.6 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	6	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	13.0%
		OR (95% CI) [b]	0.91 (0.21, 4.01)	
		p-value [b]	0.8998	
		RR (95% CI) [d]	0.92 (0.25, 3.36)	
		Risk Diff (95% CI) [e]	-1.0% (-17.5, 15.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.6%	16.7%
		OR (95% CI) [b]	1.07 (0.23, 4.90)	
	p-value [b]	0.9292		
	RR (95% CI) [d]	1.06 (0.30, 3.74)		
	Risk Diff (95% CI) [e]	1.0% (-20.5, 22.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 744 von 1722

Table 14.2.18.7.1.5.6 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	3	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.7%	15.8%
		OR (95% CI) [b]	0.44 (0.08, 2.45)	
		p-value [b]	0.3513	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8231	
		RR (95% CI) [d]	0.49 (0.11, 2.19)	
		Risk Diff (95% CI) [e]	-8.1% (-26.5, 10.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	9.4%	25.0%
		OR (95% CI) [b]	0.31 (0.05, 1.82)	
		p-value [b]	0.1942	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5318	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 745 von 1722

Table 14.2.18.7.1.5.6 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.38 (0.09, 1.61)	
		Risk Diff (95% CI) [e]	-15.6% (-42.1, 10.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.6 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	23
	Number of subjects with clinically meaningful improvement at Week 24	n	11	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.3%	14.7%
		OR (95% CI) [b]	1.30 (0.41, 4.12)	
		p-value [b]	0.6536	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4186	
		RR (95% CI) [d]	1.25 (0.47, 3.29)	
		Risk Diff (95% CI) [e]	3.6% (-11.8, 19.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.4%	21.7%
		OR (95% CI) [b]	1.16 (0.35, 3.87)	
		p-value [b]	0.8036	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5877	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 747 von 1722

Table 14.2.18.7.1.5.6 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.12 (0.44, 2.85)	
		Risk Diff (95% CI) [e]	2.7% (-18.3, 23.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.6 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	13.0%
		OR (95% CI) [b]	0.58 (0.12, 2.83)	
		p-value [b]	0.5005	
		RR (95% CI) [d]	0.61 (0.15, 2.52)	
		Risk Diff (95% CI) [e]	-5.0% (-20.7, 10.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.8%	16.7%
		OR (95% CI) [b]	0.67 (0.13, 3.37)	
	p-value [b]	0.6238		
	RR (95% CI) [d]	0.71 (0.18, 2.82)		
	Risk Diff (95% CI) [e]	-4.9% (-25.2, 15.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 749 von 1722

Table 14.2.18.7.1.6.6 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	26.3%
		OR (95% CI) [b]	0.61 (0.17, 2.27)	
		p-value [b]	0.4627	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5590	
		RR (95% CI) [d]	0.68 (0.25, 1.87)	
		Risk Diff (95% CI) [e]	-8.4% (-31.5, 14.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%	41.7%
		OR (95% CI) [b]	0.39 (0.09, 1.62)	
		p-value [b]	0.1965	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2741	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 750 von 1722

Table 14.2.18.7.1.6.6 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.53 (0.21, 1.34)	
		Risk Diff (95% CI) [e]	-19.8% (-51.1, 11.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.6 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	23
	Number of subjects with clinically meaningful improvement at Week 24	n	14	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.3%	26.5%
		OR (95% CI) [b]	0.85 (0.32, 2.23)	
		p-value [b]	0.7340	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7492	
		RR (95% CI) [d]	0.88 (0.43, 1.82)	
		Risk Diff (95% CI) [e]	-3.1% (-21.4, 15.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.1%	39.1%
		OR (95% CI) [b]	0.70 (0.25, 2.00)	
		p-value [b]	0.5093	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5169	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 752 von 1722

Table 14.2.18.7.1.6.6 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.80 (0.41, 1.55)	
		Risk Diff (95% CI) [e]	-8.0% (-32.1, 16.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.6 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.0%	8.7%
		OR (95% CI) [b]	1.17 (0.21, 6.51)	
		p-value [b]	0.8606	
		RR (95% CI) [d]	1.15 (0.24, 5.49)	
		Risk Diff (95% CI) [e]	1.3% (-12.9, 15.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.7%	11.1%
		OR (95% CI) [b]	1.38 (0.24, 7.93)	
	p-value [b]	0.7188		
	RR (95% CI) [d]	1.32 (0.28, 6.16)		
	Risk Diff (95% CI) [e]	3.6% (-15.2, 22.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 754 von 1722

Table 14.2.18.7.1.7.6 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	7	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	21.1%
		OR (95% CI) [b]	0.82 (0.21, 3.24)	
		p-value [b]	0.7774	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5399	
		RR (95% CI) [d]	0.85 (0.28, 2.56)	
		Risk Diff (95% CI) [e]	-3.1% (-25.0, 18.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%	33.3%
		OR (95% CI) [b]	0.56 (0.13, 2.42)	
		p-value [b]	0.4376	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9383	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 755 von 1722

Table 14.2.18.7.1.7.6 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.66 (0.23, 1.85)	
		Risk Diff (95% CI) [e]	-11.5% (-41.7, 18.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 756 von 1722

Table 14.2.18.7.1.7.6 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	44	24
	Number of subjects with clinically meaningful improvement at Week 24	n	18	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	30.0%	41.2%
		OR (95% CI) [b]	0.61 (0.25, 1.47)	
		p-value [b]	0.2736	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7129	
		RR (95% CI) [d]	0.73 (0.42, 1.27)	
		Risk Diff (95% CI) [e]	-11.2% (-31.4, 9.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	40.9%
	OR (95% CI) [b]	0.49 (0.18, 1.36)		
	p-value [b]	0.1717		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9547		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 757 von 1722

Table 14.2.18.7.1.7.6 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.70 (0.43, 1.15)	
		Risk Diff (95% CI) [e]	-17.4% (-41.9, 7.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 758 von 1722

Table 14.2.18.7.1.7.6 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	7	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.0%	26.1%
		OR (95% CI) [b]	0.46 (0.14, 1.57)	
		p-value [b]	0.2162	
		RR (95% CI) [d]	0.54 (0.20, 1.42)	
		Risk Diff (95% CI) [e]	-12.1% (-32.4, 8.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.6%	33.3%
		OR (95% CI) [b]	0.52 (0.14, 1.87)	
	p-value [b]	0.3165		
	RR (95% CI) [d]	0.62 (0.24, 1.56)		
	Risk Diff (95% CI) [e]	-12.7% (-38.4, 12.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 759 von 1722

Table 14.2.18.7.1.8.6 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	8	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.5%	10.5%
		OR (95% CI) [b]	2.19 (0.42, 11.52)	
		p-value [b]	0.3533	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9533	
		RR (95% CI) [d]	1.95 (0.46, 8.30)	
		Risk Diff (95% CI) [e]	10.0% (-8.7, 28.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	16.7%
		OR (95% CI) [b]	1.67 (0.30, 9.27)	
		p-value [b]	0.5596	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9504	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.6 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.50 (0.37, 6.08)	
		Risk Diff (95% CI) [e]	8.3% (-17.5, 34.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 761 von 1722

Table 14.2.18.7.1.8.6 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	24
	Number of subjects with clinically meaningful improvement at Week 24	n	6	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.0%	8.8%
		OR (95% CI) [b]	1.15 (0.27, 4.92)	
		p-value [b]	0.8523	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9503	
		RR (95% CI) [d]	1.13 (0.30, 4.24)	
		Risk Diff (95% CI) [e]	1.2% (-11.0, 13.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.3%
	OR (95% CI) [b]	1.08 (0.24, 4.75)		
	p-value [b]	0.9220		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9483		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 762 von 1722

Table 14.2.18.7.1.8.6 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.07 (0.29, 3.89)	
		Risk Diff (95% CI) [e]	0.8% (-15.7, 17.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 763 von 1722

Table 14.2.18.7.1.8.6 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	3	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9496	
		RR (95% CI) [d]	3.29 (0.18, 61.27)	
		Risk Diff (95% CI) [e]	4.8% (-4.2, 13.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	8.8%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9479		
	RR (95% CI) [d]	3.80 (0.21, 69.76)		
	Risk Diff (95% CI) [e]	7.4% (-4.9, 19.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 764 von 1722

Table 14.2.18.7.1.9.6 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	11	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.2%	15.8%
		OR (95% CI) [b]	2.10 (0.51, 8.64)	
		p-value [b]	0.3062	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3276	
		RR (95% CI) [d]	1.79 (0.56, 5.66)	
		Risk Diff (95% CI) [e]	12.4% (-9.2, 34.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.4%	25.0%
		OR (95% CI) [b]	1.57 (0.35, 7.02)	
		p-value [b]	0.5539	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5863	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.6 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.38 (0.46, 4.09)	
		Risk Diff (95% CI) [e]	9.4% (-20.1, 38.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.6 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	44	24
	Number of subjects with clinically meaningful improvement at Week 24	n	9	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.0%	23.5%
		OR (95% CI) [b]	0.57 (0.20, 1.66)	
		p-value [b]	0.3054	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7874	
		RR (95% CI) [d]	0.64 (0.27, 1.50)	
		Risk Diff (95% CI) [e]	-8.5% (-25.4, 8.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.5%
	OR (95% CI) [b]	0.51 (0.17, 1.58)		
	p-value [b]	0.2450		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5984		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 767 von 1722

Table 14.2.18.7.1.9.6 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.61 (0.27, 1.38)	
		Risk Diff (95% CI) [e]	-12.9% (-35.2, 9.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 768 von 1722

Table 14.2.18.7.1.9.6 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.0%	13.0%
		OR (95% CI) [b]	0.74 (0.16, 3.40)	
		p-value [b]	0.6998	
		RR (95% CI) [d]	0.77 (0.20, 2.94)	
		Risk Diff (95% CI) [e]	-3.0% (-19.1, 13.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.7%	16.7%
		OR (95% CI) [b]	0.86 (0.18, 4.11)	
	p-value [b]	0.8522		
	RR (95% CI) [d]	0.88 (0.24, 3.28)		
	Risk Diff (95% CI) [e]	-2.0% (-22.9, 19.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 769 von 1722

Table 14.2.18.7.1.10.6 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	10.5%
		OR (95% CI) [b]	1.86 (0.35, 9.96)	
		p-value [b]	0.4688	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8293	
		RR (95% CI) [d]	1.71 (0.39, 7.44)	
		Risk Diff (95% CI) [e]	7.4% (-10.9, 25.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%	16.7%
		OR (95% CI) [b]	1.40 (0.25, 7.93)	
		p-value [b]	0.7037	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8479	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 770 von 1722

Table 14.2.18.7.1.10.6 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.31 (0.32, 5.45)	
		Risk Diff (95% CI) [e]	5.2% (-20.3, 30.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.6 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	42	24
	Number of subjects with clinically meaningful improvement at Week 24	n	11	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.3%	35.3%
		OR (95% CI) [b]	0.41 (0.16, 1.08)	
		p-value [b]	0.0700	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2072	
		RR (95% CI) [d]	0.52 (0.26, 1.05)	
		Risk Diff (95% CI) [e]	-17.0% (-35.8, 1.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.2%
	OR (95% CI) [b]	0.35 (0.12, 1.02)		
	p-value [b]	0.0543		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1169		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 772 von 1722

Table 14.2.18.7.1.10.6 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.52 (0.27, 1.00)	
		Risk Diff (95% CI) [e]	-23.8% (-47.8, 0.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 773 von 1722

Table 14.2.18.7.1.10.6 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	33	18
	Number of subjects with clinically meaningful improvement at Week 24	n	6	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	8.7%
		OR (95% CI) [b]	1.43 (0.27, 7.70)	
		p-value [b]	0.6759	
		RR (95% CI) [d]	1.38 (0.30, 6.32)	
		Risk Diff (95% CI) [e]	3.3% (-11.3, 17.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.2%	11.1%
		OR (95% CI) [b]	1.78 (0.32, 9.88)	
	p-value [b]	0.5111		
	RR (95% CI) [d]	1.64 (0.37, 7.29)		
	Risk Diff (95% CI) [e]	7.1% (-12.5, 26.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 774 von 1722

Table 14.2.18.7.1.11.6 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	31	12
	Number of subjects with clinically meaningful improvement at Week 24	n	7	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	15.8%
		OR (95% CI) [b]	1.17 (0.27, 5.12)	
		p-value [b]	0.8382	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4893	
		RR (95% CI) [d]	1.14 (0.33, 3.91)	
		Risk Diff (95% CI) [e]	2.2% (-18.2, 22.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.6%	25.0%
		OR (95% CI) [b]	0.88 (0.18, 4.14)	
		p-value [b]	0.8663	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8023		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 775 von 1722

Table 14.2.18.7.1.11.6 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.90 (0.28, 2.93)	
		Risk Diff (95% CI) [e]	-2.4% (-31.0, 26.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 776 von 1722

Table 14.2.18.7.1.11.6 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	44	24
	Number of subjects with clinically meaningful improvement at Week 24	n	14	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.3%	29.4%
		OR (95% CI) [b]	0.73 (0.28, 1.89)	
		p-value [b]	0.5169	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7858	
		RR (95% CI) [d]	0.79 (0.40, 1.59)	
		Risk Diff (95% CI) [e]	-6.1% (-24.8, 12.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.8%
	OR (95% CI) [b]	0.65 (0.23, 1.83)		
	p-value [b]	0.4180		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9709		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 777 von 1722

Table 14.2.18.7.1.11.6 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.76 (0.40, 1.45)	
		Risk Diff (95% CI) [e]	-9.8% (-33.9, 14.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 778 von 1722

Table 14.2.18.7.1.11.6 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.0%	21.7%
		OR (95% CI) [b]	0.59 (0.16, 2.09)	
		p-value [b]	0.4106	
		RR (95% CI) [d]	0.64 (0.23, 1.81)	
		Risk Diff (95% CI) [e]	-7.7% (-27.1, 11.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.6%	27.8%
		OR (95% CI) [b]	0.67 (0.18, 2.54)	
	p-value [b]	0.5595		
	RR (95% CI) [d]	0.74 (0.27, 2.01)		
	Risk Diff (95% CI) [e]	-7.2% (-31.9, 17.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 779 von 1722

Table 14.2.18.7.1.12.6 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	11
	Number of subjects with clinically meaningful improvement at Week 24	n	10	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.6%	5.3%
		OR (95% CI) [b]	6.21 (0.73, 52.66)	
		p-value [b]	0.0942	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2493	
		RR (95% CI) [d]	4.87 (0.67, 35.33)	
		Risk Diff (95% CI) [e]	20.4% (3.4, 37.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.3%	9.1%
		OR (95% CI) [b]	4.55 (0.51, 40.51)	
		p-value [b]	0.1749	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3986	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.6 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	3.44 (0.49, 23.88)	
		Risk Diff (95% CI) [e]	22.2% (-1.2, 45.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 781 von 1722

Table 14.2.18.7.1.12.6 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	24
	Number of subjects with clinically meaningful improvement at Week 24	n	10	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.7%	20.6%
		OR (95% CI) [b]	0.77 (0.26, 2.26)	
		p-value [b]	0.6356	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8995	
		RR (95% CI) [d]	0.81 (0.34, 1.93)	
		Risk Diff (95% CI) [e]	-3.9% (-20.5, 12.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%
	OR (95% CI) [b]	0.69 (0.22, 2.14)		
	p-value [b]	0.5248		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7412		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 782 von 1722

Table 14.2.18.7.1.12.6 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.76 (0.33, 1.75)	
		Risk Diff (95% CI) [e]	-6.9% (-28.8, 14.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 783 von 1722

Table 14.2.18.7.1.12.6 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	33	18
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.0%	4.3%
		OR (95% CI) [b]	0.92 (0.08, 10.65)	
		p-value [b]	0.9446	
		RR (95% CI) [d]	0.92 (0.09, 9.64)	
		Risk Diff (95% CI) [e]	-0.3% (-10.3, 9.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	6.1%	5.6%
		OR (95% CI) [b]	1.10 (0.09, 13.00)	
	p-value [b]	0.9416		
	RR (95% CI) [d]	1.09 (0.11, 11.22)		
	Risk Diff (95% CI) [e]	0.5% (-12.8, 13.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 784 von 1722

Table 14.2.18.7.1.13.6 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	11	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.2%	10.5%
		OR (95% CI) [b]	3.34 (0.66, 16.92)	
		p-value [b]	0.1453	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2041	
		RR (95% CI) [d]	2.68 (0.66, 10.90)	
		Risk Diff (95% CI) [e]	17.7% (-2.1, 37.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.4%	16.7%
		OR (95% CI) [b]	2.62 (0.49, 14.11)	
		p-value [b]	0.2626	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4118	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 785 von 1722

Table 14.2.18.7.1.13.6 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.06 (0.53, 7.98)	
		Risk Diff (95% CI) [e]	17.7% (-9.0, 44.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 786 von 1722

Table 14.2.18.7.1.13.6 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	23
	Number of subjects with clinically meaningful improvement at Week 24	n	10	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.7%	17.6%
		OR (95% CI) [b]	0.93 (0.31, 2.84)	
		p-value [b]	0.9033	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9654	
		RR (95% CI) [d]	0.94 (0.38, 2.37)	
		Risk Diff (95% CI) [e]	-1.0% (-16.9, 14.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%
	OR (95% CI) [b]	0.81 (0.25, 2.60)		
	p-value [b]	0.7225		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7403		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.6 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.85 (0.35, 2.05)	
		Risk Diff (95% CI) [e]	-3.9% (-25.5, 17.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.6 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	10	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	21.7%
		OR (95% CI) [b]	0.90 (0.27, 3.02)	
		p-value [b]	0.8644	
		RR (95% CI) [d]	0.92 (0.35, 2.39)	
		Risk Diff (95% CI) [e]	-1.7% (-21.9, 18.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.4%	27.8%
		OR (95% CI) [b]	1.08 (0.30, 3.85)	
	p-value [b]	0.9015		
	RR (95% CI) [d]	1.06 (0.43, 2.63)		
	Risk Diff (95% CI) [e]	1.6% (-24.1, 27.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 789 von 1722

Table 14.2.18.7.1.14.6 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	32	12
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.8%	15.8%
		OR (95% CI) [b]	0.78 (0.17, 3.69)	
		p-value [b]	0.7583	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9159	
		RR (95% CI) [d]	0.81 (0.22, 3.05)	
		Risk Diff (95% CI) [e]	-3.0% (-22.4, 16.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.6%	25.0%
		OR (95% CI) [b]	0.56 (0.11, 2.80)	
		p-value [b]	0.4765	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6670	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 790 von 1722

Table 14.2.18.7.1.14.6 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.63 (0.18, 2.22)	
		Risk Diff (95% CI) [e]	-9.4% (-36.9, 18.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 791 von 1722

Table 14.2.18.7.1.14.6 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	23
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.7%	5.9%
		OR (95% CI) [b]	1.14 (0.20, 6.59)	
		p-value [b]	0.8813	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8860	
		RR (95% CI) [d]	1.13 (0.22, 5.87)	
		Risk Diff (95% CI) [e]	0.8% (-9.3, 10.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	8.9%
	OR (95% CI) [b]	1.02 (0.17, 6.06)		
	p-value [b]	0.9788		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9812		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.6 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.02 (0.20, 5.17)	
		Risk Diff (95% CI) [e]	0.2% (-14.0, 14.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.6 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects at baseline	N	50	23
	Number of HRQoL evaluable subjects at Week 24	N	34	18
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.0%	4.3%
		OR (95% CI) [b]	0.92 (0.08, 10.65)	
		p-value [b]	0.9446	
		RR (95% CI) [d]	0.92 (0.09, 9.64)	
		Risk Diff (95% CI) [e]	-0.3% (-10.3, 9.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	5.9%	5.6%
		OR (95% CI) [b]	1.06 (0.09, 12.58)	
	p-value [b]	0.9617		
	RR (95% CI) [d]	1.06 (0.10, 10.90)		
	Risk Diff (95% CI) [e]	0.3% (-12.9, 13.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 794 von 1722

Table 14.2.18.7.1.15.6 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects at baseline	N	39	19
	Number of HRQoL evaluable subjects at Week 24	N	31	12
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.7%	10.5%
		OR (95% CI) [b]	0.71 (0.11, 4.64)	
		p-value [b]	0.7192	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8701	
		RR (95% CI) [d]	0.73 (0.13, 4.01)	
		Risk Diff (95% CI) [e]	-2.8% (-19.0, 13.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	9.7%	16.7%
		OR (95% CI) [b]	0.54 (0.08, 3.69)	
		p-value [b]	0.5261	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6544	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.6 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.58 (0.11, 3.06)	
		Risk Diff (95% CI) [e]	-7.0% (-30.5, 16.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.6 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects at baseline	N	60	34
	Number of HRQoL evaluable subjects at Week 24	N	45	23
	Number of subjects with clinically meaningful improvement at Week 24	n	8	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.3%	17.6%
		OR (95% CI) [b]	0.72 (0.23, 2.28)	
		p-value [b]	0.5736	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8598	
		RR (95% CI) [d]	0.76 (0.29, 2.00)	
		Risk Diff (95% CI) [e]	-4.3% (-19.7, 11.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.8%	26.1%
		OR (95% CI) [b]	0.61 (0.18, 2.04)	
		p-value [b]	0.4251	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6782	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 797 von 1722

Table 14.2.18.7.1.15.6 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.68 (0.27, 1.73)	
		Risk Diff (95% CI) [e]	-8.3% (-29.4, 12.8)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 798 von 1722

Table 14.2.18.7.1.15.6 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluatable subjects at baseline	N	50	23
	Number of HRQoL evaluatable subjects at Week 24	N	33	18
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	4.0%	4.3%
		OR (95% CI) [b]	0.92 (0.08, 10.65)	
		p-value [b]	0.9446	
		RR (95% CI) [d]	0.92 (0.09, 9.64)	
		Risk Diff (95% CI) [e]	-0.3% (-10.3, 9.6)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	6.1%	5.6%
		OR (95% CI) [b]	1.10 (0.09, 12.99)	
	p-value [b]	0.9418		
	RR (95% CI) [d]	1.09 (0.11, 11.22)		
	Risk Diff (95% CI) [e]	0.5% (-12.8, 13.9)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 799 von 1722

Table 14.2.18.7.1.1.7 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	44
	Number of subjects with clinically meaningful improvement at Week 24	n	29	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.6%	13.8%
		OR (95% CI) [b]	1.61 (0.71, 3.64)	
		p-value [b]	0.2507	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6016	
		RR (95% CI) [d]	1.49 (0.75, 2.95)	
		Risk Diff (95% CI) [e]	6.7% (-4.0, 17.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.1%
	OR (95% CI) [b]	1.45 (0.62, 3.37)		
	p-value [b]	0.3938		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8028		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 800 von 1722

Table 14.2.18.7.1.1.7 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.33 (0.68, 2.56)	
		Risk Diff (95% CI) [e]	6.6% (-7.9, 21.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.7 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluatable subjects at baseline	N	8	11
	Number of HRQoL evaluatable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	27.3%
		OR (95% CI) [b]	0.89 (0.11, 7.11)	
		p-value [b]	0.9116	
		RR (95% CI) [d]	0.92 (0.20, 4.28)	
		Risk Diff (95% CI) [e]	-2.3% (-42.2, 37.6)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	50.0%	33.3%
		OR (95% CI) [b]	2.00 (0.18, 22.06)	
	p-value [b]	0.5714		
	RR (95% CI) [d]	1.50 (0.39, 5.77)		
	Risk Diff (95% CI) [e]	16.7% (-41.2, 74.5)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 802 von 1722

Table 14.2.18.7.1.2.7 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	45
	Number of subjects with clinically meaningful improvement at Week 24	n	22	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.6%	20.0%
		OR (95% CI) [b]	0.74 (0.35, 1.58)	
		p-value [b]	0.4359	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9792	
		RR (95% CI) [d]	0.78 (0.42, 1.45)	
		Risk Diff (95% CI) [e]	-4.4% (-15.8, 7.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.6%
	OR (95% CI) [b]	0.64 (0.29, 1.41)		
	p-value [b]	0.2676		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3422		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.7 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.71 (0.39, 1.28)	
		Risk Diff (95% CI) [e]	-8.3% (-23.6, 7.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 804 von 1722

Table 14.2.18.7.1.2.7 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	3	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	37.5%	45.5%
		OR (95% CI) [b]	0.72 (0.11, 4.62)	
		p-value [b]	0.7291	
		RR (95% CI) [d]	0.83 (0.27, 2.49)	
		Risk Diff (95% CI) [e]	-8.0% (-52.6, 36.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	75.0%	55.6%
		OR (95% CI) [b]	2.40 (0.18, 32.85)	
	p-value [b]	0.5123		
	RR (95% CI) [d]	1.35 (0.60, 3.04)		
	Risk Diff (95% CI) [e]	19.4% (-34.0, 72.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 805 von 1722

Table 14.2.18.7.1.3.7 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	45
	Number of subjects with clinically meaningful improvement at Week 24	n	27	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	20.0%
		OR (95% CI) [b]	0.95 (0.45, 1.98)	
		p-value [b]	0.8859	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7880	
		RR (95% CI) [d]	0.96 (0.53, 1.73)	
		Risk Diff (95% CI) [e]	-0.9% (-12.5, 10.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.2%
	OR (95% CI) [b]	0.83 (0.38, 1.81)		
	p-value [b]	0.6406		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4464		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.7 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.87 (0.50, 1.53)	
		Risk Diff (95% CI) [e]	-3.7% (-19.2, 11.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.7 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluatable subjects at baseline	N	8	11
	Number of HRQoL evaluatable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	3	5
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	37.5%	45.5%
		OR (95% CI) [b]	0.72 (0.11, 4.62)	
		p-value [b]	0.7291	
		RR (95% CI) [d]	0.83 (0.27, 2.49)	
		Risk Diff (95% CI) [e]	-8.0% (-52.6, 36.7)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	75.0%	55.6%
		OR (95% CI) [b]	2.40 (0.18, 32.88)	
	p-value [b]	0.5121		
	RR (95% CI) [d]	1.35 (0.60, 3.04)		
	Risk Diff (95% CI) [e]	19.4% (-34.0, 72.9)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 808 von 1722

Table 14.2.18.7.1.4.7 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	44
	Number of subjects with clinically meaningful improvement at Week 24	n	27	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	13.8%
		OR (95% CI) [b]	1.47 (0.65, 3.34)	
		p-value [b]	0.3537	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2355	
		RR (95% CI) [d]	1.38 (0.69, 2.77)	
		Risk Diff (95% CI) [e]	5.3% (-5.3, 15.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.2%
	OR (95% CI) [b]	1.31 (0.56, 3.08)		
	p-value [b]	0.5319		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6990		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.7 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.23 (0.63, 2.41)	
		Risk Diff (95% CI) [e]	4.8% (-9.7, 19.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.7 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluatable subjects at baseline	N	8	11
	Number of HRQoL evaluatable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	2	5
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	45.5%
		OR (95% CI) [b]	0.40 (0.05, 2.93)	
		p-value [b]	0.3674	
		RR (95% CI) [d]	0.55 (0.14, 2.15)	
		Risk Diff (95% CI) [e]	-20.5% (-62.5, 21.6)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	50.0%	55.6%
		OR (95% CI) [b]	0.80 (0.08, 8.47)	
	p-value [b]	0.8530		
	RR (95% CI) [d]	0.90 (0.29, 2.82)		
	Risk Diff (95% CI) [e]	-5.6% (-64.3, 53.2)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 811 von 1722

Table 14.2.18.7.1.5.7 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	44
	Number of subjects with clinically meaningful improvement at Week 24	n	17	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.1%	12.3%
		OR (95% CI) [b]	0.98 (0.40, 2.40)	
		p-value [b]	0.9591	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4840	
		RR (95% CI) [d]	0.98 (0.45, 2.15)	
		Risk Diff (95% CI) [e]	-0.3% (-9.9, 9.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.9%
	OR (95% CI) [b]	0.85 (0.34, 2.14)		
	p-value [b]	0.7306		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8655		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.7 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.87 (0.41, 1.88)	
		Risk Diff (95% CI) [e]	-2.3% (-15.6, 11.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 813 von 1722

Table 14.2.18.7.1.5.7 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	1	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	27.3%
		OR (95% CI) [b]	0.38 (0.03, 4.55)	
		p-value [b]	0.4456	
		RR (95% CI) [d]	0.46 (0.06, 3.64)	
		Risk Diff (95% CI) [e]	-14.8% (-49.7, 20.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.05, 9.47)	
	p-value [b]	0.7646		
	RR (95% CI) [d]	0.75 (0.11, 5.18)		
	Risk Diff (95% CI) [e]	-8.3% (-60.8, 44.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 814 von 1722

Table 14.2.18.7.1.6.7 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	44
	Number of subjects with clinically meaningful improvement at Week 24	n	25	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.7%	20.0%
		OR (95% CI) [b]	0.86 (0.41, 1.82)	
		p-value [b]	0.6965	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5365	
		RR (95% CI) [d]	0.89 (0.49, 1.62)	
		Risk Diff (95% CI) [e]	-2.3% (-13.9, 9.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.4%
	OR (95% CI) [b]	0.73 (0.33, 1.60)		
	p-value [b]	0.4274		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9510		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 815 von 1722

Table 14.2.18.7.1.6.7 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.79 (0.45, 1.40)	
		Risk Diff (95% CI) [e]	-6.2% (-21.9, 9.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 816 von 1722

Table 14.2.18.7.1.6.7 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	1	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	27.3%
		OR (95% CI) [b]	0.38 (0.03, 4.55)	
		p-value [b]	0.4457	
		RR (95% CI) [d]	0.46 (0.06, 3.64)	
		Risk Diff (95% CI) [e]	-14.8% (-49.7, 20.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.05, 9.47)	
	p-value [b]	0.7645		
	RR (95% CI) [d]	0.75 (0.11, 5.18)		
	Risk Diff (95% CI) [e]	-8.3% (-60.8, 44.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 817 von 1722

Table 14.2.18.7.1.7.7 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	106	45
	Number of subjects with clinically meaningful improvement at Week 24	n	30	17
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.3%	26.2%
		OR (95% CI) [b]	0.76 (0.38, 1.51)	
		p-value [b]	0.4389	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2017	
		RR (95% CI) [d]	0.81 (0.48, 1.36)	
		Risk Diff (95% CI) [e]	-4.9% (-17.5, 7.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.3%
	OR (95% CI) [b]	0.65 (0.31, 1.36)		
	p-value [b]	0.2516		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5382		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.7 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.46, 1.21)	
		Risk Diff (95% CI) [e]	-9.5% (-26.0, 7.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.7 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	2	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	63.6%
		OR (95% CI) [b]	0.19 (0.03, 1.43)	
		p-value [b]	0.1072	
		RR (95% CI) [d]	0.39 (0.11, 1.41)	
		Risk Diff (95% CI) [e]	-38.6% (-80.0, 2.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	50.0%	77.8%
		OR (95% CI) [b]	0.29 (0.02, 3.52)	
	p-value [b]	0.3285		
	RR (95% CI) [d]	0.64 (0.23, 1.82)		
	Risk Diff (95% CI) [e]	-27.8% (-83.8, 28.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.7 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	45
	Number of subjects with clinically meaningful improvement at Week 24	n	16	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.3%	4.6%
		OR (95% CI) [b]	2.64 (0.74, 9.42)	
		p-value [b]	0.1334	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3374	
		RR (95% CI) [d]	2.46 (0.74, 8.14)	
		Risk Diff (95% CI) [e]	6.7% (-0.6, 14.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.0%
	OR (95% CI) [b]	2.46 (0.68, 8.91)		
	p-value [b]	0.1699		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6304		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 821 von 1722

Table 14.2.18.7.1.8.7 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.24 (0.69, 7.32)	
		Risk Diff (95% CI) [e]	8.3% (-1.7, 18.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.7 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	1	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	18.2%
		OR (95% CI) [b]	0.64 (0.05, 8.62)	
		p-value [b]	0.7386	
		RR (95% CI) [d]	0.69 (0.07, 6.34)	
		Risk Diff (95% CI) [e]	-5.7% (-38.0, 26.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	22.2%
		OR (95% CI) [b]	1.17 (0.07, 18.35)	
	p-value [b]	0.9127		
	RR (95% CI) [d]	1.13 (0.14, 9.11)		
	Risk Diff (95% CI) [e]	2.8% (-47.6, 53.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.7 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	106	45
	Number of subjects with clinically meaningful improvement at Week 24	n	22	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.6%	21.5%
		OR (95% CI) [b]	0.67 (0.32, 1.42)	
		p-value [b]	0.2990	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9459	
		RR (95% CI) [d]	0.72 (0.40, 1.32)	
		Risk Diff (95% CI) [e]	-5.9% (-17.6, 5.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.8%
	OR (95% CI) [b]	0.58 (0.26, 1.27)		
	p-value [b]	0.1746		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9376		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.7 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.67 (0.38, 1.18)	
		Risk Diff (95% CI) [e]	-10.4% (-25.9, 5.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.7 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	3	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	37.5%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9475	
		RR (95% CI) [d]	9.33 (0.55, 158.85)	
		Risk Diff (95% CI) [e]	34.7% (0.9, 68.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	75.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9400		
	RR (95% CI) [d]	14.00 (0.88, 221.62)		
	Risk Diff (95% CI) [e]	65.0% (22.6, 100.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 826 von 1722

Table 14.2.18.7.1.10.7 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	103	45
	Number of subjects with clinically meaningful improvement at Week 24	n	23	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.3%	20.0%
		OR (95% CI) [b]	0.78 (0.37, 1.66)	
		p-value [b]	0.5178	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5882	
		RR (95% CI) [d]	0.82 (0.44, 1.51)	
		Risk Diff (95% CI) [e]	-3.7% (-15.2, 7.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.3%
	OR (95% CI) [b]	0.71 (0.32, 1.57)		
	p-value [b]	0.3934		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9662		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 827 von 1722

Table 14.2.18.7.1.10.7 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.77 (0.43, 1.38)	
		Risk Diff (95% CI) [e]	-6.6% (-22.1, 8.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 828 von 1722

Table 14.2.18.7.1.10.7 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	1	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	27.3%
		OR (95% CI) [b]	0.38 (0.03, 4.55)	
		p-value [b]	0.4457	
		RR (95% CI) [d]	0.46 (0.06, 3.64)	
		Risk Diff (95% CI) [e]	-14.8% (-49.7, 20.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.05, 9.47)	
	p-value [b]	0.7644		
	RR (95% CI) [d]	0.75 (0.11, 5.18)		
	Risk Diff (95% CI) [e]	-8.3% (-60.8, 44.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 829 von 1722

Table 14.2.18.7.1.11.7 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	105	45
	Number of subjects with clinically meaningful improvement at Week 24	n	27	15
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	23.1%
		OR (95% CI) [b]	0.79 (0.39, 1.61)	
		p-value [b]	0.5160	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5800	
		RR (95% CI) [d]	0.83 (0.47, 1.45)	
		Risk Diff (95% CI) [e]	-3.9% (-16.1, 8.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.7%
	OR (95% CI) [b]	0.69 (0.32, 1.48)		
	p-value [b]	0.3421		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9786		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 830 von 1722

Table 14.2.18.7.1.11.7 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.77 (0.46, 1.31)	
		Risk Diff (95% CI) [e]	-7.6% (-23.7, 8.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 831 von 1722

Table 14.2.18.7.1.11.7 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	1	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	27.3%
		OR (95% CI) [b]	0.38 (0.03, 4.55)	
		p-value [b]	0.4457	
		RR (95% CI) [d]	0.46 (0.06, 3.64)	
		Risk Diff (95% CI) [e]	-14.8% (-49.7, 20.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.05, 9.47)	
	p-value [b]	0.7646		
	RR (95% CI) [d]	0.75 (0.11, 5.18)		
	Risk Diff (95% CI) [e]	-8.3% (-60.8, 44.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 832 von 1722

Table 14.2.18.7.1.12.7 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	106	44
	Number of subjects with clinically meaningful improvement at Week 24	n	20	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.2%	9.2%
		OR (95% CI) [b]	1.63 (0.62, 4.26)	
		p-value [b]	0.3233	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6057	
		RR (95% CI) [d]	1.54 (0.65, 3.64)	
		Risk Diff (95% CI) [e]	5.0% (-4.1, 14.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.9%
	OR (95% CI) [b]	1.47 (0.55, 3.96)		
	p-value [b]	0.4429		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8173		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.7 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.38 (0.60, 3.21)	
		Risk Diff (95% CI) [e]	5.2% (-7.3, 17.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 834 von 1722

Table 14.2.18.7.1.12.7 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	27.3%
		OR (95% CI) [b]	0.89 (0.11, 7.11)	
		p-value [b]	0.9116	
		RR (95% CI) [d]	0.92 (0.20, 4.28)	
		Risk Diff (95% CI) [e]	-2.3% (-42.2, 37.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	50.0%	33.3%
		OR (95% CI) [b]	2.00 (0.18, 22.06)	
	p-value [b]	0.5714		
	RR (95% CI) [d]	1.50 (0.39, 5.77)		
	Risk Diff (95% CI) [e]	16.7% (-41.2, 74.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 835 von 1722

Table 14.2.18.7.1.13.7 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	44
	Number of subjects with clinically meaningful improvement at Week 24	n	29	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.6%	15.4%
		OR (95% CI) [b]	1.42 (0.65, 3.13)	
		p-value [b]	0.3791	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6778	
		RR (95% CI) [d]	1.34 (0.69, 2.58)	
		Risk Diff (95% CI) [e]	5.2% (-5.8, 16.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.1%
	OR (95% CI) [b]	1.26 (0.55, 2.88)		
	p-value [b]	0.5772		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7237		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 836 von 1722

Table 14.2.18.7.1.13.7 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.19 (0.64, 2.23)	
		Risk Diff (95% CI) [e]	4.4% (-10.6, 19.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 837 von 1722

Table 14.2.18.7.1.13.7 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	27.3%
		OR (95% CI) [b]	0.89 (0.11, 7.11)	
		p-value [b]	0.9116	
		RR (95% CI) [d]	0.92 (0.20, 4.28)	
		Risk Diff (95% CI) [e]	-2.3% (-42.2, 37.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	50.0%	33.3%
		OR (95% CI) [b]	2.00 (0.18, 22.03)	
	p-value [b]	0.5720		
	RR (95% CI) [d]	1.50 (0.39, 5.77)		
	Risk Diff (95% CI) [e]	16.7% (-41.2, 74.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 838 von 1722

Table 14.2.18.7.1.14.7 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects at baseline	N	141	65
	Number of HRQoL evaluable subjects at Week 24	N	107	44
	Number of subjects with clinically meaningful improvement at Week 24	n	11	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.8%	6.2%
		OR (95% CI) [b]	1.29 (0.39, 4.22)	
		p-value [b]	0.6730	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9536	
		RR (95% CI) [d]	1.27 (0.42, 3.83)	
		Risk Diff (95% CI) [e]	1.6% (-5.7, 9.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.3%
	OR (95% CI) [b]	1.15 (0.34, 3.81)		
	p-value [b]	0.8244		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9624		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.7 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.13 (0.38, 3.36)	
		Risk Diff (95% CI) [e]	1.2% (-9.1, 11.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.7 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects at baseline	N	8	11
	Number of HRQoL evaluable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	0	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	18.2%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9546	
		RR (95% CI) [d]	0.27 (0.01, 4.90)	
		Risk Diff (95% CI) [e]	-15.3% (-42.7, 12.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	22.2%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9629		
	RR (95% CI) [d]	0.40 (0.02, 6.85)		
	Risk Diff (95% CI) [e]	-15.0% (-52.6, 22.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 841 von 1722

Table 14.2.18.7.1.15.7 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluatable subjects at baseline	N	141	65
	Number of HRQoL evaluatable subjects at Week 24	N	105	44
	Number of subjects with clinically meaningful improvement at Week 24	n	12	6
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	8.5%	9.2%
		OR (95% CI) [b]	0.91 (0.33, 2.56)	
		p-value [b]	0.8650	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5225	
		RR (95% CI) [d]	0.92 (0.36, 2.35)	
		Risk Diff (95% CI) [e]	-0.7% (-9.1, 7.7)	
		Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.4%
	OR (95% CI) [b]	0.82 (0.29, 2.34)		
	p-value [b]	0.7063		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8888		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 842 von 1722

Table 14.2.18.7.1.15.7 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.84 (0.34, 2.09)	
		Risk Diff (95% CI) [e]	-2.2% (-14.0, 9.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 843 von 1722

Table 14.2.18.7.1.15.7 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluatable subjects at baseline	N	8	11
	Number of HRQoL evaluatable subjects at Week 24	N	4	9
	Number of subjects with clinically meaningful improvement at Week 24	n	1	3
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	27.3%
		OR (95% CI) [b]	0.38 (0.03, 4.55)	
		p-value [b]	0.4457	
		RR (95% CI) [d]	0.46 (0.06, 3.64)	
		Risk Diff (95% CI) [e]	-14.8% (-49.7, 20.1)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.05, 9.47)	
	p-value [b]	0.7646		
	RR (95% CI) [d]	0.75 (0.11, 5.18)		
	Risk Diff (95% CI) [e]	-8.3% (-60.8, 44.1)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 844 von 1722

Table 14.2.18.7.1.1.8 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	30
	Number of subjects with clinically meaningful improvement at Week 24	n	19	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.6%	11.6%
		OR (95% CI) [b]	2.22 (0.77, 6.43)	
		p-value [b]	0.1412	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2013	
		RR (95% CI) [d]	1.95 (0.78, 4.85)	
		Risk Diff (95% CI) [e]	11.0% (-2.1, 24.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.9%
	OR (95% CI) [b]	1.94 (0.65, 5.80)		
	p-value [b]	0.2367		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3245		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 845 von 1722

Table 14.2.18.7.1.1.8 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.68 (0.69, 4.07)	
		Risk Diff (95% CI) [e]	11.3% (-5.8, 28.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.8 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	12	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.5%	21.2%
		OR (95% CI) [b]	0.84 (0.30, 2.39)	
		p-value [b]	0.7450	
		RR (95% CI) [d]	0.87 (0.38, 2.00)	
		Risk Diff (95% CI) [e]	-2.8% (-19.6, 14.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.9%	30.4%
		OR (95% CI) [b]	0.88 (0.29, 2.69)	
	p-value [b]	0.8289		
	RR (95% CI) [d]	0.92 (0.42, 2.01)		
	Risk Diff (95% CI) [e]	-2.5% (-25.6, 20.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 847 von 1722

Table 14.2.18.7.1.2.8 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	31
	Number of subjects with clinically meaningful improvement at Week 24	n	16	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.0%	27.9%
		OR (95% CI) [b]	0.61 (0.26, 1.44)	
		p-value [b]	0.2569	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8105	
		RR (95% CI) [d]	0.68 (0.36, 1.31)	
		Risk Diff (95% CI) [e]	-8.9% (-24.7, 7.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.5%
	OR (95% CI) [b]	0.49 (0.20, 1.22)		
	p-value [b]	0.1233		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5723		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 848 von 1722

Table 14.2.18.7.1.2.8 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.61 (0.33, 1.13)	
		Risk Diff (95% CI) [e]	-15.2% (-35.1, 4.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.8 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.8%	18.2%
		OR (95% CI) [b]	0.72 (0.23, 2.24)	
		p-value [b]	0.5742	
		RR (95% CI) [d]	0.76 (0.30, 1.96)	
		Risk Diff (95% CI) [e]	-4.3% (-19.9, 11.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.9%	26.1%
		OR (95% CI) [b]	0.75 (0.23, 2.45)	
	p-value [b]	0.6344		
	RR (95% CI) [d]	0.80 (0.33, 1.97)		
	Risk Diff (95% CI) [e]	-5.2% (-26.8, 16.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 850 von 1722

Table 14.2.18.7.1.3.8 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	31
	Number of subjects with clinically meaningful improvement at Week 24	n	18	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.4%	25.6%
		OR (95% CI) [b]	0.79 (0.34, 1.88)	
		p-value [b]	0.5981	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9326	
		RR (95% CI) [d]	0.84 (0.44, 1.61)	
		Risk Diff (95% CI) [e]	-4.2% (-19.9, 11.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.5%
	OR (95% CI) [b]	0.65 (0.26, 1.63)		
	p-value [b]	0.3623		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6809		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 851 von 1722

Table 14.2.18.7.1.3.8 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.40, 1.38)	
		Risk Diff (95% CI) [e]	-9.0% (-28.9, 10.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 852 von 1722

Table 14.2.18.7.1.3.8 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	12	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.5%	21.2%
		OR (95% CI) [b]	0.84 (0.30, 2.39)	
		p-value [b]	0.7450	
		RR (95% CI) [d]	0.87 (0.38, 2.00)	
		Risk Diff (95% CI) [e]	-2.8% (-19.6, 14.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.9%	30.4%
		OR (95% CI) [b]	0.88 (0.29, 2.69)	
	p-value [b]	0.8289		
	RR (95% CI) [d]	0.92 (0.42, 2.01)		
	Risk Diff (95% CI) [e]	-2.5% (-25.6, 20.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 853 von 1722

Table 14.2.18.7.1.4.8 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	30
	Number of subjects with clinically meaningful improvement at Week 24	n	19	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.6%	16.3%
		OR (95% CI) [b]	1.50 (0.58, 3.92)	
		p-value [b]	0.4040	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2754	
		RR (95% CI) [d]	1.39 (0.63, 3.05)	
		Risk Diff (95% CI) [e]	6.3% (-7.9, 20.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.9%
	OR (95% CI) [b]	1.27 (0.47, 3.46)		
	p-value [b]	0.6344		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4295		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 854 von 1722

Table 14.2.18.7.1.4.8 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.20 (0.56, 2.54)	
		Risk Diff (95% CI) [e]	4.6% (-13.9, 23.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.8 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluatable subjects at baseline	N	65	33
	Number of HRQoL evaluatable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	10	7
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	15.4%	21.2%
		OR (95% CI) [b]	0.68 (0.23, 1.97)	
		p-value [b]	0.4733	
		RR (95% CI) [d]	0.73 (0.30, 1.73)	
		Risk Diff (95% CI) [e]	-5.8% (-22.3, 10.6)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.3%	30.4%
		OR (95% CI) [b]	0.69 (0.22, 2.16)	
	p-value [b]	0.5261		
	RR (95% CI) [d]	0.76 (0.34, 1.74)		
	Risk Diff (95% CI) [e]	-7.2% (-29.8, 15.5)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 856 von 1722

Table 14.2.18.7.1.5.8 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	30
	Number of subjects with clinically meaningful improvement at Week 24	n	10	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.9%	18.6%
		OR (95% CI) [b]	0.59 (0.21, 1.63)	
		p-value [b]	0.3092	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3265	
		RR (95% CI) [d]	0.64 (0.27, 1.50)	
		Risk Diff (95% CI) [e]	-6.7% (-20.2, 6.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.7%
	OR (95% CI) [b]	0.47 (0.17, 1.36)		
	p-value [b]	0.1641		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1986		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.8 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.55 (0.24, 1.26)	
		Risk Diff (95% CI) [e]	-12.0% (-29.9, 6.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 858 von 1722

Table 14.2.18.7.1.5.8 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.3%	9.1%
		OR (95% CI) [b]	1.40 (0.35, 5.68)	
		p-value [b]	0.6348	
		RR (95% CI) [d]	1.35 (0.38, 4.77)	
		Risk Diff (95% CI) [e]	3.2% (-9.4, 15.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.6%	13.0%
		OR (95% CI) [b]	1.52 (0.36, 6.41)	
	p-value [b]	0.5654		
	RR (95% CI) [d]	1.43 (0.42, 4.86)		
	Risk Diff (95% CI) [e]	5.6% (-12.5, 23.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 859 von 1722

Table 14.2.18.7.1.6.8 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	30
	Number of subjects with clinically meaningful improvement at Week 24	n	15	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	25.6%
		OR (95% CI) [b]	0.63 (0.26, 1.53)	
		p-value [b]	0.3095	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4257	
		RR (95% CI) [d]	0.70 (0.35, 1.39)	
		Risk Diff (95% CI) [e]	-7.7% (-23.1, 7.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.1%	36.7%
		OR (95% CI) [b]	0.49 (0.19, 1.25)	
		p-value [b]	0.1348	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2331	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 860 von 1722

Table 14.2.18.7.1.6.8 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.60 (0.31, 1.15)	
		Risk Diff (95% CI) [e]	-14.6% (-34.5, 5.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 861 von 1722

Table 14.2.18.7.1.6.8 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	11	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.9%	15.2%
		OR (95% CI) [b]	1.14 (0.36, 3.61)	
		p-value [b]	0.8227	
		RR (95% CI) [d]	1.12 (0.42, 2.95)	
		Risk Diff (95% CI) [e]	1.8% (-13.5, 17.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.6%	21.7%
		OR (95% CI) [b]	1.24 (0.37, 4.13)	
		p-value [b]	0.7288	
		RR (95% CI) [d]	1.18 (0.47, 2.98)	
		Risk Diff (95% CI) [e]	3.8% (-17.5, 25.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 862 von 1722

Table 14.2.18.7.1.7.8 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	31
	Number of subjects with clinically meaningful improvement at Week 24	n	14	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.7%	32.6%
		OR (95% CI) [b]	0.41 (0.18, 0.98)	
		p-value [b]	0.0441	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2399	
		RR (95% CI) [d]	0.51 (0.27, 0.97)	
		Risk Diff (95% CI) [e]	-15.9% (-32.0, 0.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.6%
	OR (95% CI) [b]	0.31 (0.13, 0.79)		
	p-value [b]	0.0138		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1079		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.8 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.46 (0.25, 0.84)	
		Risk Diff (95% CI) [e]	-24.6% (-44.6, -4.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.8 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	42	23
	Number of subjects with clinically meaningful improvement at Week 24	n	18	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	27.7%	30.3%
		OR (95% CI) [b]	0.88 (0.35, 2.21)	
		p-value [b]	0.7869	
		RR (95% CI) [d]	0.91 (0.48, 1.75)	
		Risk Diff (95% CI) [e]	-2.6% (-21.7, 16.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	42.9%	43.5%
		OR (95% CI) [b]	0.97 (0.35, 2.72)	
	p-value [b]	0.9614		
	RR (95% CI) [d]	0.99 (0.55, 1.76)		
	Risk Diff (95% CI) [e]	-0.6% (-25.8, 24.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 865 von 1722

Table 14.2.18.7.1.8.8 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	31
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.5%	7.0%
		OR (95% CI) [b]	1.40 (0.35, 5.58)	
		p-value [b]	0.6305	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5939	
		RR (95% CI) [d]	1.37 (0.38, 4.89)	
		Risk Diff (95% CI) [e]	2.5% (-7.3, 12.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.8%
	OR (95% CI) [b]	1.24 (0.31, 5.05)		
	p-value [b]	0.7596		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4630		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 866 von 1722

Table 14.2.18.7.1.8.8 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.22 (0.35, 4.27)	
		Risk Diff (95% CI) [e]	2.1% (-10.8, 15.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.8 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	9	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.8%	6.1%
		OR (95% CI) [b]	2.49 (0.51, 12.26)	
		p-value [b]	0.2617	
		RR (95% CI) [d]	2.28 (0.52, 9.97)	
		Risk Diff (95% CI) [e]	7.8% (-3.9, 19.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.9%	8.7%
		OR (95% CI) [b]	2.78 (0.55, 14.13)	
	p-value [b]	0.2178		
	RR (95% CI) [d]	2.41 (0.57, 10.22)		
	Risk Diff (95% CI) [e]	12.2% (-4.5, 29.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 868 von 1722

Table 14.2.18.7.1.9.8 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	31
	Number of subjects with clinically meaningful improvement at Week 24	n	16	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.0%	20.9%
		OR (95% CI) [b]	0.89 (0.36, 2.22)	
		p-value [b]	0.8006	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9868	
		RR (95% CI) [d]	0.91 (0.44, 1.89)	
		Risk Diff (95% CI) [e]	-1.9% (-16.7, 12.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.5%
	OR (95% CI) [b]	0.75 (0.29, 1.96)		
	p-value [b]	0.5595		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7380		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.8 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.81 (0.40, 1.63)	
		Risk Diff (95% CI) [e]	-5.5% (-24.4, 13.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 870 von 1722

Table 14.2.18.7.1.9.8 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	42	23
	Number of subjects with clinically meaningful improvement at Week 24	n	9	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.8%	15.2%
		OR (95% CI) [b]	0.90 (0.28, 2.94)	
		p-value [b]	0.8617	
		RR (95% CI) [d]	0.91 (0.33, 2.51)	
		Risk Diff (95% CI) [e]	-1.3% (-16.1, 13.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.4%	21.7%
		OR (95% CI) [b]	0.98 (0.29, 3.38)	
	p-value [b]	0.9768		
	RR (95% CI) [d]	0.99 (0.37, 2.59)		
	Risk Diff (95% CI) [e]	-0.3% (-21.2, 20.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 871 von 1722

Table 14.2.18.7.1.10.8 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	31
	Number of subjects with clinically meaningful improvement at Week 24	n	12	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	23.3%
		OR (95% CI) [b]	0.55 (0.22, 1.40)	
		p-value [b]	0.2101	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3986	
		RR (95% CI) [d]	0.61 (0.29, 1.31)	
		Risk Diff (95% CI) [e]	-9.0% (-23.6, 5.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.6%	32.3%
	OR (95% CI) [b]	0.45 (0.17, 1.20)		
	p-value [b]	0.1094		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1821		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 872 von 1722

Table 14.2.18.7.1.10.8 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.55 (0.27, 1.13)	
		Risk Diff (95% CI) [e]	-14.6% (-33.4, 4.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 873 von 1722

Table 14.2.18.7.1.10.8 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	39	23
	Number of subjects with clinically meaningful improvement at Week 24	n	12	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.5%	18.2%
		OR (95% CI) [b]	1.02 (0.34, 3.01)	
		p-value [b]	0.9730	
		RR (95% CI) [d]	1.02 (0.42, 2.46)	
		Risk Diff (95% CI) [e]	0.3% (-15.9, 16.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%	26.1%
		OR (95% CI) [b]	1.26 (0.40, 3.99)	
	p-value [b]	0.6951		
	RR (95% CI) [d]	1.18 (0.51, 2.71)		
	Risk Diff (95% CI) [e]	4.7% (-18.4, 27.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 874 von 1722

Table 14.2.18.7.1.11.8 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	67	31
	Number of subjects with clinically meaningful improvement at Week 24	n	13	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.5%	16.3%
		OR (95% CI) [b]	0.94 (0.35, 2.57)	
		p-value [b]	0.9064	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5174	
		RR (95% CI) [d]	0.95 (0.41, 2.21)	
		Risk Diff (95% CI) [e]	-0.8% (-14.3, 12.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.4%	22.6%
	OR (95% CI) [b]	0.83 (0.29, 2.33)		
	p-value [b]	0.7169		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6792		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 875 von 1722

Table 14.2.18.7.1.11.8 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.86 (0.38, 1.94)	
		Risk Diff (95% CI) [e]	-3.2% (-20.7, 14.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 876 von 1722

Table 14.2.18.7.1.11.8 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	42	23
	Number of subjects with clinically meaningful improvement at Week 24	n	15	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.1%	33.3%
		OR (95% CI) [b]	0.60 (0.24, 1.51)	
		p-value [b]	0.2794	
		RR (95% CI) [d]	0.69 (0.36, 1.33)	
		Risk Diff (95% CI) [e]	-10.3% (-29.3, 8.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	35.7%	47.8%
		OR (95% CI) [b]	0.61 (0.22, 1.70)	
		p-value [b]	0.3422	
		RR (95% CI) [d]	0.75 (0.41, 1.35)	
		Risk Diff (95% CI) [e]	-12.1% (-37.1, 12.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 877 von 1722

Table 14.2.18.7.1.12.8 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	67	31
	Number of subjects with clinically meaningful improvement at Week 24	n	13	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.5%	7.0%
		OR (95% CI) [b]	2.44 (0.66, 9.08)	
		p-value [b]	0.1831	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1689	
		RR (95% CI) [d]	2.22 (0.67, 7.37)	
		Risk Diff (95% CI) [e]	8.5% (-2.4, 19.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.4%
	OR (95% CI) [b]	2.25 (0.59, 8.54)		
	p-value [b]	0.2349		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2049		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 878 von 1722

Table 14.2.18.7.1.12.8 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.00 (0.62, 6.53)	
		Risk Diff (95% CI) [e]	9.7% (-4.3, 23.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.8 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	22
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.8%	18.2%
		OR (95% CI) [b]	0.72 (0.23, 2.24)	
		p-value [b]	0.5742	
		RR (95% CI) [d]	0.76 (0.30, 1.96)	
		Risk Diff (95% CI) [e]	-4.3% (-19.9, 11.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.9%	27.3%
		OR (95% CI) [b]	0.71 (0.21, 2.32)	
	p-value [b]	0.5667		
	RR (95% CI) [d]	0.77 (0.31, 1.88)		
	Risk Diff (95% CI) [e]	-6.3% (-28.6, 15.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 880 von 1722

Table 14.2.18.7.1.13.8 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	30
	Number of subjects with clinically meaningful improvement at Week 24	n	15	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	16.3%
		OR (95% CI) [b]	1.12 (0.42, 2.99)	
		p-value [b]	0.8242	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7093	
		RR (95% CI) [d]	1.10 (0.48, 2.49)	
		Risk Diff (95% CI) [e]	1.6% (-12.2, 15.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.1%	23.3%
	OR (95% CI) [b]	0.93 (0.33, 2.58)		
	p-value [b]	0.8893		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4442		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 881 von 1722

Table 14.2.18.7.1.13.8 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.95 (0.43, 2.08)	
		Risk Diff (95% CI) [e]	-1.3% (-19.3, 16.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 882 von 1722

Table 14.2.18.7.1.13.8 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	16	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	24.6%	18.2%
		OR (95% CI) [b]	1.47 (0.51, 4.20)	
		p-value [b]	0.4720	
		RR (95% CI) [d]	1.35 (0.58, 3.13)	
		Risk Diff (95% CI) [e]	6.4% (-10.4, 23.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	37.2%	26.1%
		OR (95% CI) [b]	1.68 (0.55, 5.13)	
	p-value [b]	0.3631		
	RR (95% CI) [d]	1.43 (0.65, 3.14)		
	Risk Diff (95% CI) [e]	11.1% (-11.9, 34.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.8 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	68	30
	Number of subjects with clinically meaningful improvement at Week 24	n	8	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.5%	9.3%
		OR (95% CI) [b]	1.03 (0.29, 3.62)	
		p-value [b]	0.9678	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7829	
		RR (95% CI) [d]	1.02 (0.33, 3.21)	
		Risk Diff (95% CI) [e]	0.2% (-10.5, 10.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.8%	13.3%
	OR (95% CI) [b]	0.87 (0.24, 3.13)		
	p-value [b]	0.8273		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9340		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 884 von 1722

Table 14.2.18.7.1.14.8 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.88 (0.29, 2.71)	
		Risk Diff (95% CI) [e]	-1.6% (-15.9, 12.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 885 von 1722

Table 14.2.18.7.1.14.8 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.6%	6.1%
		OR (95% CI) [b]	0.75 (0.12, 4.72)	
		p-value [b]	0.7593	
		RR (95% CI) [d]	0.76 (0.13, 4.34)	
		Risk Diff (95% CI) [e]	-1.4% (-11.1, 8.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.0%	8.7%
		OR (95% CI) [b]	0.79 (0.12, 5.09)	
	p-value [b]	0.8018		
	RR (95% CI) [d]	0.80 (0.14, 4.46)		
	Risk Diff (95% CI) [e]	-1.7% (-15.5, 12.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 886 von 1722

Table 14.2.18.7.1.15.8 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	84	43
	Number of HRQoL evaluable subjects at Week 24	N	66	30
	Number of subjects with clinically meaningful improvement at Week 24	n	8	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.5%	14.0%
		OR (95% CI) [b]	0.65 (0.21, 2.01)	
		p-value [b]	0.4532	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7940	
		RR (95% CI) [d]	0.68 (0.25, 1.84)	
		Risk Diff (95% CI) [e]	-4.4% (-16.5, 7.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.1%	20.0%
	OR (95% CI) [b]	0.55 (0.17, 1.76)		
	p-value [b]	0.3152		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6361		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 887 von 1722

Table 14.2.18.7.1.15.8 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.61 (0.23, 1.59)	
		Risk Diff (95% CI) [e]	-7.9% (-24.2, 8.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 888 von 1722

Table 14.2.18.7.1.15.8 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.7%	9.1%
		OR (95% CI) [b]	0.83 (0.19, 3.72)	
		p-value [b]	0.8113	
		RR (95% CI) [d]	0.85 (0.22, 3.33)	
		Risk Diff (95% CI) [e]	-1.4% (-13.2, 10.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.6%	13.0%
		OR (95% CI) [b]	0.88 (0.19, 4.05)	
	p-value [b]	0.8667		
	RR (95% CI) [d]	0.89 (0.23, 3.40)		
	Risk Diff (95% CI) [e]	-1.4% (-18.2, 15.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.9 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	12	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.5%	21.2%
		OR (95% CI) [b]	0.84 (0.30, 2.39)	
		p-value [b]	0.7450	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1944	
		RR (95% CI) [d]	0.87 (0.38, 2.00)	
		Risk Diff (95% CI) [e]	-2.8% (-19.6, 14.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.9%
	OR (95% CI) [b]	0.88 (0.29, 2.69)		
	p-value [b]	0.8289		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2999		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 890 von 1722

Table 14.2.18.7.1.1.9 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.92 (0.42, 2.01)	
		Risk Diff (95% CI) [e]	-2.5% (-25.6, 20.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.9 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	13.0%
		OR (95% CI) [b]	1.67 (0.39, 7.03)	
		p-value [b]	0.4868	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5941	
		RR (95% CI) [d]	1.53 (0.45, 5.21)	
		Risk Diff (95% CI) [e]	7.0% (-11.6, 25.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.8%
	OR (95% CI) [b]	1.51 (0.34, 6.69)		
	p-value [b]	0.5897		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6465		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.9 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.38 (0.42, 4.49)	
		Risk Diff (95% CI) [e]	7.1% (-17.5, 31.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 893 von 1722

Table 14.2.18.7.1.1.9 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	14
	Number of subjects with clinically meaningful improvement at Week 24	n	11	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	10.0%
		OR (95% CI) [b]	3.00 (0.60, 15.04)	
		p-value [b]	0.1817	
		RR (95% CI) [d]	2.50 (0.61, 10.25)	
		Risk Diff (95% CI) [e]	15.0% (-3.3, 33.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.7%	14.3%
		OR (95% CI) [b]	2.54 (0.49, 13.28)	
	p-value [b]	0.2698		
	RR (95% CI) [d]	2.08 (0.53, 8.23)		
	Risk Diff (95% CI) [e]	15.4% (-8.1, 39.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 894 von 1722

Table 14.2.18.7.1.2.9 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.8%	18.2%
		OR (95% CI) [b]	0.72 (0.23, 2.24)	
		p-value [b]	0.5742	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8246	
		RR (95% CI) [d]	0.76 (0.30, 1.96)	
		Risk Diff (95% CI) [e]	-4.3% (-19.9, 11.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.9%
	OR (95% CI) [b]	0.75 (0.23, 2.45)		
	p-value [b]	0.6344		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6193		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 895 von 1722

Table 14.2.18.7.1.2.9 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.80 (0.33, 1.97)	
		Risk Diff (95% CI) [e]	-5.2% (-26.8, 16.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.9 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	7	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.5%	26.1%
		OR (95% CI) [b]	0.60 (0.17, 2.07)	
		p-value [b]	0.4200	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9985	
		RR (95% CI) [d]	0.67 (0.26, 1.76)	
		Risk Diff (95% CI) [e]	-8.6% (-30.1, 12.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.6%	37.5%
	OR (95% CI) [b]	0.49 (0.13, 1.81)		
	p-value [b]	0.2829		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9930		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 897 von 1722

Table 14.2.18.7.1.2.9 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.60 (0.24, 1.49)	
		Risk Diff (95% CI) [e]	-14.9% (-42.8, 13.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 898 von 1722

Table 14.2.18.7.1.2.9 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	15
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.5%	30.0%
		OR (95% CI) [b]	0.60 (0.18, 2.00)	
		p-value [b]	0.4059	
		RR (95% CI) [d]	0.68 (0.28, 1.66)	
		Risk Diff (95% CI) [e]	-9.5% (-32.9, 13.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.3%	40.0%
		OR (95% CI) [b]	0.48 (0.13, 1.73)	
	p-value [b]	0.2629		
	RR (95% CI) [d]	0.61 (0.26, 1.41)		
	Risk Diff (95% CI) [e]	-15.7% (-44.1, 12.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 899 von 1722

Table 14.2.18.7.1.3.9 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	12	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.5%	21.2%
		OR (95% CI) [b]	0.84 (0.30, 2.39)	
		p-value [b]	0.7450	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5875	
		RR (95% CI) [d]	0.87 (0.38, 2.00)	
		Risk Diff (95% CI) [e]	-2.8% (-19.6, 14.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.9%
	OR (95% CI) [b]	0.88 (0.29, 2.69)		
	p-value [b]	0.8289		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3867		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 900 von 1722

Table 14.2.18.7.1.3.9 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.92 (0.42, 2.01)	
		Risk Diff (95% CI) [e]	-2.5% (-25.6, 20.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 901 von 1722

Table 14.2.18.7.1.3.9 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	8	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	17.4%
		OR (95% CI) [b]	1.19 (0.31, 4.48)	
		p-value [b]	0.7997	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3876	
		RR (95% CI) [d]	1.15 (0.39, 3.40)	
		Risk Diff (95% CI) [e]	2.6% (-17.2, 22.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.8%
	OR (95% CI) [b]	1.04 (0.26, 4.18)		
	p-value [b]	0.9521		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3433		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 902 von 1722

Table 14.2.18.7.1.3.9 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.03 (0.37, 2.91)	
		Risk Diff (95% CI) [e]	0.8% (-25.4, 27.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.9 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	15
	Number of subjects with clinically meaningful improvement at Week 24	n	10	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.7%	35.0%
		OR (95% CI) [b]	0.55 (0.17, 1.74)	
		p-value [b]	0.3061	
		RR (95% CI) [d]	0.65 (0.29, 1.46)	
		Risk Diff (95% CI) [e]	-12.3% (-36.6, 12.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.0%	46.7%
		OR (95% CI) [b]	0.42 (0.12, 1.47)	
	p-value [b]	0.1766		
	RR (95% CI) [d]	0.58 (0.27, 1.23)		
	Risk Diff (95% CI) [e]	-19.6% (-48.7, 9.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 904 von 1722

Table 14.2.18.7.1.4.9 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	10	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.4%	21.2%
		OR (95% CI) [b]	0.68 (0.23, 1.97)	
		p-value [b]	0.4732	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5192	
		RR (95% CI) [d]	0.73 (0.30, 1.73)	
		Risk Diff (95% CI) [e]	-5.8% (-22.3, 10.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.3%
	OR (95% CI) [b]	0.69 (0.22, 2.16)		
	p-value [b]	0.5262		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7489		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.9 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.76 (0.34, 1.74)	
		Risk Diff (95% CI) [e]	-7.2% (-29.8, 15.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.9 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	9	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.5%	13.0%
		OR (95% CI) [b]	1.94 (0.47, 8.03)	
		p-value [b]	0.3629	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6130	
		RR (95% CI) [d]	1.73 (0.52, 5.74)	
		Risk Diff (95% CI) [e]	9.5% (-9.4, 28.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.0%
	OR (95% CI) [b]	1.77 (0.41, 7.75)		
	p-value [b]	0.4470		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5270		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 907 von 1722

Table 14.2.18.7.1.4.9 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.55 (0.49, 4.93)	
		Risk Diff (95% CI) [e]	10.3% (-14.6, 35.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.9 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	14
	Number of subjects with clinically meaningful improvement at Week 24	n	10	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.7%	20.0%
		OR (95% CI) [b]	1.18 (0.32, 4.33)	
		p-value [b]	0.8069	
		RR (95% CI) [d]	1.14 (0.40, 3.19)	
		Risk Diff (95% CI) [e]	2.7% (-18.7, 24.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.0%	28.6%
		OR (95% CI) [b]	0.93 (0.24, 3.64)	
	p-value [b]	0.9122		
	RR (95% CI) [d]	0.95 (0.35, 2.53)		
	Risk Diff (95% CI) [e]	-1.5% (-29.2, 26.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 909 von 1722

Table 14.2.18.7.1.5.9 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.3%	9.1%
		OR (95% CI) [b]	1.40 (0.35, 5.68)	
		p-value [b]	0.6348	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3518	
		RR (95% CI) [d]	1.35 (0.38, 4.77)	
		Risk Diff (95% CI) [e]	3.2% (-9.4, 15.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.6%
	OR (95% CI) [b]	1.52 (0.36, 6.41)		
	p-value [b]	0.5654		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2029		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 910 von 1722

Table 14.2.18.7.1.5.9 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.43 (0.42, 4.86)	
		Risk Diff (95% CI) [e]	5.6% (-12.5, 23.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 911 von 1722

Table 14.2.18.7.1.5.9 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	3	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.5%	13.0%
		OR (95% CI) [b]	0.54 (0.10, 2.93)	
		p-value [b]	0.4756	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9642	
		RR (95% CI) [d]	0.58 (0.13, 2.62)	
		Risk Diff (95% CI) [e]	-5.5% (-21.5, 10.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	9.7%
	OR (95% CI) [b]	0.46 (0.08, 2.62)		
	p-value [b]	0.3848		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9290		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 912 von 1722

Table 14.2.18.7.1.5.9 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.52 (0.12, 2.27)	
		Risk Diff (95% CI) [e]	-9.1% (-30.8, 12.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 913 von 1722

Table 14.2.18.7.1.5.9 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	14
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.9%	25.0%
		OR (95% CI) [b]	0.57 (0.16, 2.07)	
		p-value [b]	0.3913	
		RR (95% CI) [d]	0.64 (0.23, 1.76)	
		Risk Diff (95% CI) [e]	-9.1% (-30.9, 12.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.9%	35.7%
		OR (95% CI) [b]	0.42 (0.11, 1.65)	
	p-value [b]	0.2140		
	RR (95% CI) [d]	0.53 (0.20, 1.40)		
	Risk Diff (95% CI) [e]	-16.8% (-44.9, 11.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 914 von 1722

Table 14.2.18.7.1.6.9 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	11	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.9%	15.2%
		OR (95% CI) [b]	1.14 (0.36, 3.61)	
		p-value [b]	0.8227	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4498	
		RR (95% CI) [d]	1.12 (0.42, 2.95)	
		Risk Diff (95% CI) [e]	1.8% (-13.5, 17.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.6%	21.7%
		OR (95% CI) [b]	1.24 (0.37, 4.13)	
		p-value [b]	0.7288	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2405	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.9 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.18 (0.47, 2.98)	
		Risk Diff (95% CI) [e]	3.8% (-17.5, 25.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 916 von 1722

Table 14.2.18.7.1.6.9 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	6	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.0%	21.7%
		OR (95% CI) [b]	0.64 (0.17, 2.37)	
		p-value [b]	0.4996	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9500	
		RR (95% CI) [d]	0.69 (0.24, 2.01)	
		Risk Diff (95% CI) [e]	-6.7% (-26.9, 13.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.4%
	OR (95% CI) [b]	0.53 (0.13, 2.10)		
	p-value [b]	0.3653		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8293		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 917 von 1722

Table 14.2.18.7.1.6.9 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.62 (0.22, 1.72)	
		Risk Diff (95% CI) [e]	-11.9% (-38.5, 14.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.9 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	14
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.5%	30.0%
		OR (95% CI) [b]	0.60 (0.18, 2.00)	
		p-value [b]	0.4059	
		RR (95% CI) [d]	0.68 (0.28, 1.66)	
		Risk Diff (95% CI) [e]	-9.5% (-32.9, 13.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.3%	42.9%
		OR (95% CI) [b]	0.43 (0.12, 1.57)	
	p-value [b]	0.2007		
	RR (95% CI) [d]	0.57 (0.25, 1.30)		
	Risk Diff (95% CI) [e]	-18.5% (-47.9, 10.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 919 von 1722

Table 14.2.18.7.1.7.9 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	42	23
	Number of subjects with clinically meaningful improvement at Week 24	n	18	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	27.7%	30.3%
		OR (95% CI) [b]	0.88 (0.35, 2.21)	
		p-value [b]	0.7869	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0822	
		RR (95% CI) [d]	0.91 (0.48, 1.75)	
		Risk Diff (95% CI) [e]	-2.6% (-21.7, 16.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	42.9%	43.5%
		OR (95% CI) [b]	0.97 (0.35, 2.72)	
		p-value [b]	0.9614	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0315	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.9 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.99 (0.55, 1.76)	
		Risk Diff (95% CI) [e]	-0.6% (-25.8, 24.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 921 von 1722

Table 14.2.18.7.1.7.9 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.5%	21.7%
		OR (95% CI) [b]	0.76 (0.21, 2.76)	
		p-value [b]	0.6805	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1818	
		RR (95% CI) [d]	0.81 (0.29, 2.25)	
		Risk Diff (95% CI) [e]	-4.2% (-24.8, 16.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.6%
	OR (95% CI) [b]	0.64 (0.17, 2.48)		
	p-value [b]	0.5199		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1416		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 922 von 1722

Table 14.2.18.7.1.7.9 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.27, 1.92)	
		Risk Diff (95% CI) [e]	-8.7% (-35.7, 18.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.9 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	15
	Number of subjects with clinically meaningful improvement at Week 24	n	7	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.9%	45.0%
		OR (95% CI) [b]	0.23 (0.07, 0.76)	
		p-value [b]	0.0163	
		RR (95% CI) [d]	0.35 (0.15, 0.81)	
		Risk Diff (95% CI) [e]	-29.1% (-53.4, -4.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.9%	60.0%
		OR (95% CI) [b]	0.16 (0.04, 0.58)	
	p-value [b]	0.0058		
	RR (95% CI) [d]	0.32 (0.14, 0.69)		
	Risk Diff (95% CI) [e]	-41.1% (-68.9, -13.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 924 von 1722

Table 14.2.18.7.1.8.9 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	9	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.8%	6.1%
		OR (95% CI) [b]	2.49 (0.51, 12.26)	
		p-value [b]	0.2617	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6853	
		RR (95% CI) [d]	2.28 (0.52, 9.97)	
		Risk Diff (95% CI) [e]	7.8% (-3.9, 19.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.9%
	OR (95% CI) [b]	2.78 (0.55, 14.13)		
	p-value [b]	0.2178		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5777		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.9 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.41 (0.57, 10.22)	
		Risk Diff (95% CI) [e]	12.2% (-4.5, 29.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 926 von 1722

Table 14.2.18.7.1.8.9 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	8.7%
		OR (95% CI) [b]	1.50 (0.27, 8.43)	
		p-value [b]	0.6454	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9590	
		RR (95% CI) [d]	1.44 (0.30, 6.82)	
		Risk Diff (95% CI) [e]	3.8% (-11.6, 19.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.1%
	OR (95% CI) [b]	1.35 (0.23, 7.85)		
	p-value [b]	0.7412		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9542		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 927 von 1722

Table 14.2.18.7.1.8.9 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.29 (0.28, 5.93)	
		Risk Diff (95% CI) [e]	3.6% (-17.1, 24.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.9 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	15
	Number of subjects with clinically meaningful improvement at Week 24	n	3	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.8%	5.0%
		OR (95% CI) [b]	1.39 (0.14, 14.25)	
		p-value [b]	0.7815	
		RR (95% CI) [d]	1.36 (0.15, 12.31)	
		Risk Diff (95% CI) [e]	1.8% (-10.3, 13.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	8.1%	6.7%
		OR (95% CI) [b]	1.24 (0.12, 12.91)	
	p-value [b]	0.8600		
	RR (95% CI) [d]	1.22 (0.14, 10.78)		
	Risk Diff (95% CI) [e]	1.4% (-13.9, 16.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 929 von 1722

Table 14.2.18.7.1.9.9 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	42	23
	Number of subjects with clinically meaningful improvement at Week 24	n	9	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.8%	15.2%
		OR (95% CI) [b]	0.90 (0.28, 2.94)	
		p-value [b]	0.8615	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6065	
		RR (95% CI) [d]	0.91 (0.33, 2.51)	
		Risk Diff (95% CI) [e]	-1.3% (-16.1, 13.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.4%	21.7%
		OR (95% CI) [b]	0.98 (0.29, 3.38)	
		p-value [b]	0.9768	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4260	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.9 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.99 (0.37, 2.59)	
		Risk Diff (95% CI) [e]	-0.3% (-21.2, 20.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 931 von 1722

Table 14.2.18.7.1.9.9 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	9	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.5%	17.4%
		OR (95% CI) [b]	1.38 (0.37, 5.11)	
		p-value [b]	0.6304	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3447	
		RR (95% CI) [d]	1.29 (0.45, 3.74)	
		Risk Diff (95% CI) [e]	5.1% (-15.1, 25.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.0%
	OR (95% CI) [b]	1.23 (0.31, 4.84)		
	p-value [b]	0.7698		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3252		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 932 von 1722

Table 14.2.18.7.1.9.9 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.16 (0.42, 3.19)	
		Risk Diff (95% CI) [e]	4.0% (-22.5, 30.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.9 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	15
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.9%	25.0%
		OR (95% CI) [b]	0.57 (0.16, 2.07)	
		p-value [b]	0.3913	
		RR (95% CI) [d]	0.64 (0.23, 1.76)	
		Risk Diff (95% CI) [e]	-9.1% (-30.9, 12.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.9%	33.3%
		OR (95% CI) [b]	0.47 (0.12, 1.80)	
	p-value [b]	0.2694		
	RR (95% CI) [d]	0.57 (0.21, 1.51)		
	Risk Diff (95% CI) [e]	-14.4% (-41.4, 12.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 934 von 1722

Table 14.2.18.7.1.10.9 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	39	23
	Number of subjects with clinically meaningful improvement at Week 24	n	12	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.5%	18.2%
		OR (95% CI) [b]	1.02 (0.34, 3.01)	
		p-value [b]	0.9730	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5955	
		RR (95% CI) [d]	1.02 (0.42, 2.46)	
		Risk Diff (95% CI) [e]	0.3% (-15.9, 16.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%
	OR (95% CI) [b]	1.26 (0.40, 3.99)		
	p-value [b]	0.6951		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3602		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 935 von 1722

Table 14.2.18.7.1.10.9 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.18 (0.51, 2.71)	
		Risk Diff (95% CI) [e]	4.7% (-18.4, 27.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.9 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	6	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.0%	26.1%
		OR (95% CI) [b]	0.50 (0.14, 1.79)	
		p-value [b]	0.2857	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8083	
		RR (95% CI) [d]	0.58 (0.21, 1.58)	
		Risk Diff (95% CI) [e]	-11.1% (-32.2, 10.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.4%
	OR (95% CI) [b]	0.40 (0.10, 1.54)		
	p-value [b]	0.1829		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7766		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 937 von 1722

Table 14.2.18.7.1.10.9 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.52 (0.20, 1.34)	
		Risk Diff (95% CI) [e]	-18.1% (-45.6, 9.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.9 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	15
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.6%	20.0%
		OR (95% CI) [b]	0.63 (0.16, 2.54)	
		p-value [b]	0.5181	
		RR (95% CI) [d]	0.68 (0.22, 2.15)	
		Risk Diff (95% CI) [e]	-6.4% (-26.6, 13.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.2%	26.7%
		OR (95% CI) [b]	0.53 (0.13, 2.25)	
	p-value [b]	0.3907		
	RR (95% CI) [d]	0.61 (0.20, 1.85)		
	Risk Diff (95% CI) [e]	-10.5% (-35.8, 14.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 939 von 1722

Table 14.2.18.7.1.11.9 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	42	23
	Number of subjects with clinically meaningful improvement at Week 24	n	15	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.1%	33.3%
		OR (95% CI) [b]	0.60 (0.24, 1.51)	
		p-value [b]	0.2794	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2842	
		RR (95% CI) [d]	0.69 (0.36, 1.33)	
		Risk Diff (95% CI) [e]	-10.3% (-29.3, 8.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	35.7%	47.8%
		OR (95% CI) [b]	0.61 (0.22, 1.70)	
		p-value [b]	0.3422	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3664	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 940 von 1722

Table 14.2.18.7.1.11.9 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.41, 1.35)	
		Risk Diff (95% CI) [e]	-12.1% (-37.1, 12.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 941 von 1722

Table 14.2.18.7.1.11.9 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	30	16
	Number of subjects with clinically meaningful improvement at Week 24	n	6	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.0%	21.7%
		OR (95% CI) [b]	0.64 (0.17, 2.37)	
		p-value [b]	0.4996	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3635	
		RR (95% CI) [d]	0.69 (0.24, 2.01)	
		Risk Diff (95% CI) [e]	-6.7% (-26.9, 13.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%
	OR (95% CI) [b]	0.55 (0.14, 2.20)		
	p-value [b]	0.3975		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3647		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 942 von 1722

Table 14.2.18.7.1.11.9 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.64 (0.23, 1.77)	
		Risk Diff (95% CI) [e]	-11.3% (-38.1, 15.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.9 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	15
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.9%	10.0%
		OR (95% CI) [b]	1.70 (0.32, 9.04)	
		p-value [b]	0.5321	
		RR (95% CI) [d]	1.59 (0.36, 6.99)	
		Risk Diff (95% CI) [e]	5.9% (-11.1, 22.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.9%	13.3%
		OR (95% CI) [b]	1.52 (0.28, 8.31)	
	p-value [b]	0.6313		
	RR (95% CI) [d]	1.42 (0.33, 6.06)		
	Risk Diff (95% CI) [e]	5.6% (-15.7, 26.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 944 von 1722

Table 14.2.18.7.1.12.9 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	22
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.8%	18.2%
		OR (95% CI) [b]	0.72 (0.23, 2.24)	
		p-value [b]	0.5742	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3387	
		RR (95% CI) [d]	0.76 (0.30, 1.96)	
		Risk Diff (95% CI) [e]	-4.3% (-19.9, 11.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.9%
	OR (95% CI) [b]	0.71 (0.21, 2.32)		
	p-value [b]	0.5667		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3687		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 945 von 1722

Table 14.2.18.7.1.12.9 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.77 (0.31, 1.88)	
		Risk Diff (95% CI) [e]	-6.3% (-28.6, 15.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 946 von 1722

Table 14.2.18.7.1.12.9 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	8	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	8.7%
		OR (95% CI) [b]	2.62 (0.51, 13.59)	
		p-value [b]	0.2500	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9576	
		RR (95% CI) [d]	2.30 (0.53, 9.92)	
		Risk Diff (95% CI) [e]	11.3% (-5.6, 28.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.8%
	OR (95% CI) [b]	2.43 (0.45, 13.14)		
	p-value [b]	0.3009		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9580		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 947 von 1722

Table 14.2.18.7.1.12.9 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.06 (0.50, 8.60)	
		Risk Diff (95% CI) [e]	13.3% (-9.1, 35.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 948 von 1722

Table 14.2.18.7.1.12.9 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	36	15
	Number of subjects with clinically meaningful improvement at Week 24	n	5	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.4%	5.0%
		OR (95% CI) [b]	2.44 (0.27, 22.34)	
		p-value [b]	0.4310	
		RR (95% CI) [d]	2.27 (0.28, 18.21)	
		Risk Diff (95% CI) [e]	6.4% (-7.0, 19.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.9%	6.7%
		OR (95% CI) [b]	2.26 (0.24, 21.17)	
	p-value [b]	0.4756		
	RR (95% CI) [d]	2.08 (0.27, 16.36)		
	Risk Diff (95% CI) [e]	7.2% (-9.7, 24.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.9 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	16	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	24.6%	18.2%
		OR (95% CI) [b]	1.47 (0.51, 4.20)	
		p-value [b]	0.4722	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3428	
		RR (95% CI) [d]	1.35 (0.58, 3.13)	
		Risk Diff (95% CI) [e]	6.4% (-10.4, 23.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	37.2%	26.1%
		OR (95% CI) [b]	1.68 (0.55, 5.13)	
		p-value [b]	0.3634	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1832	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 950 von 1722

Table 14.2.18.7.1.13.9 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.43 (0.65, 3.14)	
		Risk Diff (95% CI) [e]	11.1% (-11.9, 34.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 951 von 1722

Table 14.2.18.7.1.13.9 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	9	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.5%	13.0%
		OR (95% CI) [b]	1.94 (0.47, 8.03)	
		p-value [b]	0.3629	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2703	
		RR (95% CI) [d]	1.73 (0.52, 5.74)	
		Risk Diff (95% CI) [e]	9.5% (-9.4, 28.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.0%
	OR (95% CI) [b]	1.77 (0.41, 7.75)		
	p-value [b]	0.4470		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2190		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 952 von 1722

Table 14.2.18.7.1.13.9 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.55 (0.49, 4.93)	
		Risk Diff (95% CI) [e]	10.3% (-14.6, 35.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 953 von 1722

Table 14.2.18.7.1.13.9 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	14
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.6%	20.0%
		OR (95% CI) [b]	0.63 (0.16, 2.54)	
		p-value [b]	0.5181	
		RR (95% CI) [d]	0.68 (0.22, 2.15)	
		Risk Diff (95% CI) [e]	-6.4% (-26.6, 13.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.2%	28.6%
		OR (95% CI) [b]	0.48 (0.11, 2.07)	
	p-value [b]	0.3272		
	RR (95% CI) [d]	0.57 (0.19, 1.71)		
	Risk Diff (95% CI) [e]	-12.4% (-38.8, 14.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 954 von 1722

Table 14.2.18.7.1.14.9 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.6%	6.1%
		OR (95% CI) [b]	0.75 (0.12, 4.72)	
		p-value [b]	0.7593	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9227	
		RR (95% CI) [d]	0.76 (0.13, 4.34)	
		Risk Diff (95% CI) [e]	-1.4% (-11.1, 8.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.0%	8.7%
		OR (95% CI) [b]	0.79 (0.12, 5.09)	
		p-value [b]	0.8018	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7704	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 955 von 1722

Table 14.2.18.7.1.14.9 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.80 (0.14, 4.46)	
		Risk Diff (95% CI) [e]	-1.7% (-15.5, 12.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 956 von 1722

Table 14.2.18.7.1.14.9 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	31	16
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	8.7%
		OR (95% CI) [b]	1.50 (0.27, 8.43)	
		p-value [b]	0.6453	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5265	
		RR (95% CI) [d]	1.44 (0.30, 6.82)	
		Risk Diff (95% CI) [e]	3.8% (-11.6, 19.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.1%
	OR (95% CI) [b]	1.35 (0.23, 7.85)		
	p-value [b]	0.7412		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4813		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 957 von 1722

Table 14.2.18.7.1.14.9 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.29 (0.28, 5.93)	
		Risk Diff (95% CI) [e]	3.6% (-17.1, 24.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 958 von 1722

Table 14.2.18.7.1.14.9 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	44	20
	Number of HRQoL evaluable subjects at Week 24	N	37	14
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.8%	10.0%
		OR (95% CI) [b]	0.66 (0.10, 4.29)	
		p-value [b]	0.6620	
		RR (95% CI) [d]	0.68 (0.12, 3.77)	
		Risk Diff (95% CI) [e]	-3.2% (-18.3, 11.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	8.1%	14.3%
		OR (95% CI) [b]	0.53 (0.08, 3.56)	
	p-value [b]	0.5132		
	RR (95% CI) [d]	0.57 (0.11, 3.05)		
	Risk Diff (95% CI) [e]	-6.2% (-26.5, 14.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 959 von 1722

Table 14.2.18.7.1.15.9 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	65	33
	Number of HRQoL evaluable subjects at Week 24	N	43	23
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	7.7%	9.1%
		OR (95% CI) [b]	0.83 (0.19, 3.72)	
		p-value [b]	0.8113	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5511	
		RR (95% CI) [d]	0.85 (0.22, 3.33)	
		Risk Diff (95% CI) [e]	-1.4% (-13.2, 10.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.6%	13.0%
		OR (95% CI) [b]	0.88 (0.19, 4.05)	
		p-value [b]	0.8667	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6762	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.9 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.89 (0.23, 3.40)	
		Risk Diff (95% CI) [e]	-1.4% (-18.2, 15.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.9 (AMNOG)  
EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects at baseline	N	40	23
	Number of HRQoL evaluable subjects at Week 24	N	29	16
	Number of subjects with clinically meaningful improvement at Week 24	n	4	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.0%	21.7%
		OR (95% CI) [b]	0.40 (0.10, 1.67)	
		p-value [b]	0.2094	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2533	
		RR (95% CI) [d]	0.46 (0.14, 1.54)	
		Risk Diff (95% CI) [e]	-11.7% (-31.0, 7.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.8%	31.3%
	OR (95% CI) [b]	0.35 (0.08, 1.57)		
	p-value [b]	0.1707		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2816		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 962 von 1722

Table 14.2.18.7.1.15.9 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.44 (0.14, 1.41)	
		Risk Diff (95% CI) [e]	-17.5% (-43.4, 8.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.9 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluatable subjects at baseline	N	44	20
	Number of HRQoL evaluatable subjects at Week 24	N	37	14
	Number of subjects with clinically meaningful improvement at Week 24	n	4	1
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	9.1%	5.0%
		OR (95% CI) [b]	1.90 (0.20, 18.17)	
		p-value [b]	0.5775	
		RR (95% CI) [d]	1.82 (0.22, 15.25)	
		Risk Diff (95% CI) [e]	4.1% (-8.7, 16.9)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.8%	7.1%
		OR (95% CI) [b]	1.58 (0.16, 15.46)	
	p-value [b]	0.6963		
	RR (95% CI) [d]	1.51 (0.18, 12.40)		
	Risk Diff (95% CI) [e]	3.7% (-13.1, 20.5)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 964 von 1722

Table 14.2.18.7.1.1.10 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	42
	Number of subjects with clinically meaningful improvement at Week 24	n	28	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.8%	14.3%
		OR (95% CI) [b]	1.77 (0.78, 4.02)	
		p-value [b]	0.1740	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1717	
		RR (95% CI) [d]	1.59 (0.80, 3.17)	
		Risk Diff (95% CI) [e]	8.5% (-2.9, 19.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.1%
	OR (95% CI) [b]	1.58 (0.67, 3.73)		
	p-value [b]	0.2975		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2895		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 965 von 1722

Table 14.2.18.7.1.1.10 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.41 (0.73, 2.71)	
		Risk Diff (95% CI) [e]	8.7% (-6.8, 24.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 966 von 1722

Table 14.2.18.7.1.1.10 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	3	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	23.1%
		OR (95% CI) [b]	0.45 (0.08, 2.66)	
		p-value [b]	0.3816	
		RR (95% CI) [d]	0.52 (0.12, 2.22)	
		Risk Diff (95% CI) [e]	-11.1% (-37.3, 15.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.7%	27.3%
		OR (95% CI) [b]	0.53 (0.09, 3.28)	
	p-value [b]	0.4975		
	RR (95% CI) [d]	0.61 (0.15, 2.51)		
	Risk Diff (95% CI) [e]	-10.6% (-42.1, 20.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 967 von 1722

Table 14.2.18.7.1.2.10 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	43
	Number of subjects with clinically meaningful improvement at Week 24	n	20	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.3%	20.6%
		OR (95% CI) [b]	0.75 (0.34, 1.62)	
		p-value [b]	0.4607	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4654	
		RR (95% CI) [d]	0.79 (0.42, 1.48)	
		Risk Diff (95% CI) [e]	-4.4% (-16.3, 7.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.5%
	OR (95% CI) [b]	0.63 (0.28, 1.43)		
	p-value [b]	0.2717		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7274		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.10 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.71 (0.39, 1.29)	
		Risk Diff (95% CI) [e]	-8.7% (-24.8, 7.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.10 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	5	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	38.5%
		OR (95% CI) [b]	0.40 (0.09, 1.77)	
		p-value [b]	0.2269	
		RR (95% CI) [d]	0.52 (0.18, 1.48)	
		Risk Diff (95% CI) [e]	-18.5% (-49.2, 12.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.8%	45.5%
		OR (95% CI) [b]	0.46 (0.10, 2.22)	
	p-value [b]	0.3349		
	RR (95% CI) [d]	0.61 (0.23, 1.64)		
	Risk Diff (95% CI) [e]	-17.7% (-53.6, 18.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 970 von 1722

Table 14.2.18.7.1.3.10 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	43
	Number of subjects with clinically meaningful improvement at Week 24	n	26	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.1%	19.0%
		OR (95% CI) [b]	1.14 (0.53, 2.44)	
		p-value [b]	0.7380	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0606	
		RR (95% CI) [d]	1.11 (0.60, 2.05)	
		Risk Diff (95% CI) [e]	2.1% (-10.0, 14.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.0%	27.9%
		OR (95% CI) [b]	1.00 (0.45, 2.24)	
		p-value [b]	0.9952	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1205		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.10 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.00 (0.56, 1.79)	
		Risk Diff (95% CI) [e]	0.1% (-16.2, 16.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 972 von 1722

Table 14.2.18.7.1.3.10 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	4	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.0%	46.2%
		OR (95% CI) [b]	0.22 (0.05, 1.02)	
		p-value [b]	0.0536	
		RR (95% CI) [d]	0.35 (0.12, 1.01)	
		Risk Diff (95% CI) [e]	-30.2% (-60.8, 0.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	54.5%
		OR (95% CI) [b]	0.24 (0.05, 1.21)	
	p-value [b]	0.0836		
	RR (95% CI) [d]	0.41 (0.15, 1.13)		
	Risk Diff (95% CI) [e]	-32.3% (-67.5, 2.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 973 von 1722

Table 14.2.18.7.1.4.10 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	42
	Number of subjects with clinically meaningful improvement at Week 24	n	24	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.5%	15.9%
		OR (95% CI) [b]	1.28 (0.57, 2.89)	
		p-value [b]	0.5440	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3502	
		RR (95% CI) [d]	1.23 (0.63, 2.41)	
		Risk Diff (95% CI) [e]	3.6% (-7.8, 15.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.8%
	OR (95% CI) [b]	1.11 (0.48, 2.60)		
	p-value [b]	0.8046		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5867		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 974 von 1722

Table 14.2.18.7.1.4.10 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.08 (0.57, 2.06)	
		Risk Diff (95% CI) [e]	2.0% (-13.7, 17.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 975 von 1722

Table 14.2.18.7.1.4.10 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	5	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	30.8%
		OR (95% CI) [b]	0.56 (0.12, 2.60)	
		p-value [b]	0.4617	
		RR (95% CI) [d]	0.65 (0.21, 2.01)	
		Risk Diff (95% CI) [e]	-10.8% (-40.4, 18.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.8%	36.4%
		OR (95% CI) [b]	0.67 (0.14, 3.35)	
	p-value [b]	0.6283		
	RR (95% CI) [d]	0.76 (0.26, 2.25)		
	Risk Diff (95% CI) [e]	-8.6% (-43.7, 26.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 976 von 1722

Table 14.2.18.7.1.5.10 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	42
	Number of subjects with clinically meaningful improvement at Week 24	n	14	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.4%	11.1%
		OR (95% CI) [b]	1.03 (0.39, 2.69)	
		p-value [b]	0.9559	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3566	
		RR (95% CI) [d]	1.02 (0.44, 2.41)	
		Risk Diff (95% CI) [e]	0.3% (-9.3, 9.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.1%
	OR (95% CI) [b]	0.89 (0.33, 2.39)		
	p-value [b]	0.8109		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5612		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.10 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.90 (0.39, 2.07)	
		Risk Diff (95% CI) [e]	-1.6% (-15.0, 11.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.10 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	4	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.0%	30.8%
		OR (95% CI) [b]	0.43 (0.09, 2.10)	
		p-value [b]	0.2964	
		RR (95% CI) [d]	0.52 (0.15, 1.75)	
		Risk Diff (95% CI) [e]	-14.8% (-43.7, 14.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	36.4%
		OR (95% CI) [b]	0.50 (0.10, 2.62)	
	p-value [b]	0.4121		
	RR (95% CI) [d]	0.61 (0.19, 1.96)		
	Risk Diff (95% CI) [e]	-14.1% (-48.4, 20.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 979 von 1722

Table 14.2.18.7.1.6.10 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	42
	Number of subjects with clinically meaningful improvement at Week 24	n	24	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.5%	20.6%
		OR (95% CI) [b]	0.93 (0.44, 1.99)	
		p-value [b]	0.8560	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2709	
		RR (95% CI) [d]	0.95 (0.52, 1.73)	
		Risk Diff (95% CI) [e]	-1.1% (-13.3, 11.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.8%	31.0%
		OR (95% CI) [b]	0.78 (0.35, 1.73)	
		p-value [b]	0.5354	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4384	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.10 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.83 (0.47, 1.47)	
		Risk Diff (95% CI) [e]	-5.1% (-21.7, 11.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 981 von 1722

Table 14.2.18.7.1.6.10 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	23.1%
		OR (95% CI) [b]	0.29 (0.04, 2.01)	
		p-value [b]	0.2103	
		RR (95% CI) [d]	0.35 (0.07, 1.82)	
		Risk Diff (95% CI) [e]	-15.1% (-40.3, 10.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.1%	27.3%
		OR (95% CI) [b]	0.33 (0.05, 2.42)	
	p-value [b]	0.2769		
	RR (95% CI) [d]	0.41 (0.08, 2.07)		
	Risk Diff (95% CI) [e]	-16.2% (-46.2, 13.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 982 von 1722

Table 14.2.18.7.1.7.10 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	92	43
	Number of subjects with clinically meaningful improvement at Week 24	n	29	19
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.6%	30.2%
		OR (95% CI) [b]	0.71 (0.36, 1.41)	
		p-value [b]	0.3327	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1914	
		RR (95% CI) [d]	0.78 (0.48, 1.28)	
		Risk Diff (95% CI) [e]	-6.6% (-20.2, 7.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.5%
	OR (95% CI) [b]	0.58 (0.28, 1.23)		
	p-value [b]	0.1540		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3540		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.10 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.71 (0.45, 1.12)	
		Risk Diff (95% CI) [e]	-12.7% (-30.3, 5.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 984 von 1722

Table 14.2.18.7.1.7.10 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	3	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	38.5%
		OR (95% CI) [b]	0.22 (0.04, 1.13)	
		p-value [b]	0.0696	
		RR (95% CI) [d]	0.31 (0.09, 1.11)	
		Risk Diff (95% CI) [e]	-26.5% (-55.8, 2.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.7%	45.5%
		OR (95% CI) [b]	0.24 (0.04, 1.34)	
	p-value [b]	0.1031		
	RR (95% CI) [d]	0.37 (0.11, 1.24)		
	Risk Diff (95% CI) [e]	-28.8% (-62.9, 5.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 985 von 1722

Table 14.2.18.7.1.8.10 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	43
	Number of subjects with clinically meaningful improvement at Week 24	n	15	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.2%	6.3%
		OR (95% CI) [b]	2.05 (0.65, 6.45)	
		p-value [b]	0.2207	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6308	
		RR (95% CI) [d]	1.92 (0.67, 5.55)	
		Risk Diff (95% CI) [e]	5.8% (-2.5, 14.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.1%	9.3%
		OR (95% CI) [b]	1.87 (0.58, 6.03)	
		p-value [b]	0.2915	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7753		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.10 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.73 (0.61, 4.92)	
		Risk Diff (95% CI) [e]	6.8% (-4.6, 18.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 987 von 1722

Table 14.2.18.7.1.8.10 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	7.7%
		OR (95% CI) [b]	1.04 (0.09, 12.71)	
		p-value [b]	0.9734	
		RR (95% CI) [d]	1.04 (0.10, 10.43)	
		Risk Diff (95% CI) [e]	0.3% (-17.7, 18.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.1%	9.1%
		OR (95% CI) [b]	1.25 (0.10, 15.65)	
	p-value [b]	0.8626		
	RR (95% CI) [d]	1.22 (0.12, 11.95)		
	Risk Diff (95% CI) [e]	2.0% (-20.3, 24.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 988 von 1722

Table 14.2.18.7.1.9.10 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	92	43
	Number of subjects with clinically meaningful improvement at Week 24	n	22	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	20.6%
		OR (95% CI) [b]	0.84 (0.39, 1.80)	
		p-value [b]	0.6501	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5984	
		RR (95% CI) [d]	0.87 (0.47, 1.60)	
		Risk Diff (95% CI) [e]	-2.7% (-14.8, 9.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.9%
	OR (95% CI) [b]	0.73 (0.32, 1.63)		
	p-value [b]	0.4360		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4327		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 989 von 1722

Table 14.2.18.7.1.9.10 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.79 (0.44, 1.42)	
		Risk Diff (95% CI) [e]	-6.3% (-22.6, 9.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 990 von 1722

Table 14.2.18.7.1.9.10 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluatable subjects at baseline	N	25	13
	Number of HRQoL evaluatable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	3	1
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	7.7%
		OR (95% CI) [b]	1.64 (0.15, 17.50)	
		p-value [b]	0.6839	
		RR (95% CI) [d]	1.56 (0.18, 13.55)	
		Risk Diff (95% CI) [e]	4.3% (-15.0, 23.6)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.7%	9.1%
		OR (95% CI) [b]	2.00 (0.18, 22.03)	
	p-value [b]	0.5717		
	RR (95% CI) [d]	1.83 (0.22, 15.51)		
	Risk Diff (95% CI) [e]	7.6% (-16.6, 31.8)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 991 von 1722

Table 14.2.18.7.1.10.10 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	89	43
	Number of subjects with clinically meaningful improvement at Week 24	n	22	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	22.2%
		OR (95% CI) [b]	0.76 (0.36, 1.62)	
		p-value [b]	0.4795	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6806	
		RR (95% CI) [d]	0.80 (0.44, 1.46)	
		Risk Diff (95% CI) [e]	-4.3% (-16.6, 8.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.7%	32.6%
		OR (95% CI) [b]	0.68 (0.31, 1.51)	
		p-value [b]	0.3446	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8698	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 992 von 1722

Table 14.2.18.7.1.10.10 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.76 (0.43, 1.33)	
		Risk Diff (95% CI) [e]	-7.8% (-24.5, 8.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 993 von 1722

Table 14.2.18.7.1.10.10 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	2	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	15.4%
		OR (95% CI) [b]	0.48 (0.06, 3.86)	
		p-value [b]	0.4887	
		RR (95% CI) [d]	0.52 (0.08, 3.28)	
		Risk Diff (95% CI) [e]	-7.4% (-29.7, 14.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.1%	18.2%
		OR (95% CI) [b]	0.56 (0.07, 4.70)	
	p-value [b]	0.5954		
	RR (95% CI) [d]	0.61 (0.10, 3.74)		
	Risk Diff (95% CI) [e]	-7.1% (-34.1, 20.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 994 von 1722

Table 14.2.18.7.1.11.10 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	91	43
	Number of subjects with clinically meaningful improvement at Week 24	n	24	15
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.5%	23.8%
		OR (95% CI) [b]	0.78 (0.37, 1.61)	
		p-value [b]	0.4962	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8302	
		RR (95% CI) [d]	0.82 (0.46, 1.45)	
		Risk Diff (95% CI) [e]	-4.3% (-16.9, 8.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.4%
	OR (95% CI) [b]	0.67 (0.31, 1.46)		
	p-value [b]	0.3127		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8926		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.10 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.76 (0.44, 1.29)	
		Risk Diff (95% CI) [e]	-8.5% (-25.4, 8.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 996 von 1722

Table 14.2.18.7.1.11.10 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.0%	23.1%
		OR (95% CI) [b]	0.64 (0.12, 3.39)	
		p-value [b]	0.5954	
		RR (95% CI) [d]	0.69 (0.18, 2.64)	
		Risk Diff (95% CI) [e]	-7.1% (-34.1, 20.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	27.3%
		OR (95% CI) [b]	0.76 (0.13, 4.30)	
	p-value [b]	0.7582		
	RR (95% CI) [d]	0.81 (0.22, 2.98)		
	Risk Diff (95% CI) [e]	-5.1% (-37.6, 27.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 997 von 1722

Table 14.2.18.7.1.12.10 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	92	43
	Number of subjects with clinically meaningful improvement at Week 24	n	19	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.4%	9.5%
		OR (95% CI) [b]	1.74 (0.66, 4.59)	
		p-value [b]	0.2668	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1928	
		RR (95% CI) [d]	1.62 (0.68, 3.86)	
		Risk Diff (95% CI) [e]	5.9% (-3.7, 15.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.7%	14.0%
		OR (95% CI) [b]	1.61 (0.59, 4.36)	
		p-value [b]	0.3535	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2465	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.10 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.48 (0.64, 3.44)	
		Risk Diff (95% CI) [e]	6.7% (-6.6, 20.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 999 von 1722

Table 14.2.18.7.1.12.10 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	10
	Number of subjects with clinically meaningful improvement at Week 24	n	3	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	23.1%
		OR (95% CI) [b]	0.45 (0.08, 2.66)	
		p-value [b]	0.3816	
		RR (95% CI) [d]	0.52 (0.12, 2.22)	
		Risk Diff (95% CI) [e]	-11.1% (-37.3, 15.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.7%	30.0%
		OR (95% CI) [b]	0.47 (0.07, 2.92)	
	p-value [b]	0.4155		
	RR (95% CI) [d]	0.56 (0.14, 2.26)		
	Risk Diff (95% CI) [e]	-13.3% (-46.5, 19.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1000 von 1722

Table 14.2.18.7.1.13.10 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	42
	Number of subjects with clinically meaningful improvement at Week 24	n	26	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.1%	17.5%
		OR (95% CI) [b]	1.27 (0.58, 2.77)	
		p-value [b]	0.5526	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9350	
		RR (95% CI) [d]	1.21 (0.64, 2.29)	
		Risk Diff (95% CI) [e]	3.7% (-8.2, 15.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.0%
	OR (95% CI) [b]	1.09 (0.48, 2.49)		
	p-value [b]	0.8313		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6564		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.10 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.07 (0.58, 1.95)	
		Risk Diff (95% CI) [e]	1.8% (-14.4, 17.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.10 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	15.4%
		OR (95% CI) [b]	1.37 (0.23, 8.29)	
		p-value [b]	0.7286	
		RR (95% CI) [d]	1.30 (0.29, 5.81)	
		Risk Diff (95% CI) [e]	4.6% (-20.5, 29.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.8%	18.2%
		OR (95% CI) [b]	1.73 (0.27, 10.97)	
	p-value [b]	0.5605		
	RR (95% CI) [d]	1.53 (0.36, 6.57)		
	Risk Diff (95% CI) [e]	9.6% (-21.2, 40.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1003 von 1722

Table 14.2.18.7.1.14.10 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	93	42
	Number of subjects with clinically meaningful improvement at Week 24	n	10	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.1%	7.9%
		OR (95% CI) [b]	1.03 (0.34, 3.14)	
		p-value [b]	0.9634	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6459	
		RR (95% CI) [d]	1.02 (0.37, 2.87)	
		Risk Diff (95% CI) [e]	0.2% (-8.0, 8.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.8%	11.9%
		OR (95% CI) [b]	0.89 (0.28, 2.79)	
		p-value [b]	0.8437	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7925	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1004 von 1722

Table 14.2.18.7.1.14.10 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.90 (0.33, 2.48)	
		Risk Diff (95% CI) [e]	-1.2% (-12.8, 10.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.10 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	1	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.0%	7.7%
		OR (95% CI) [b]	0.50 (0.03, 8.71)	
		p-value [b]	0.6345	
		RR (95% CI) [d]	0.52 (0.04, 7.66)	
		Risk Diff (95% CI) [e]	-3.7% (-20.1, 12.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	5.6%	9.1%
		OR (95% CI) [b]	0.59 (0.03, 10.48)	
	p-value [b]	0.7180		
	RR (95% CI) [d]	0.61 (0.04, 8.81)		
	Risk Diff (95% CI) [e]	-3.5% (-23.6, 16.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.10 (AMNOG)  
EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects at baseline	N	123	63
	Number of HRQoL evaluable subjects at Week 24	N	91	42
	Number of subjects with clinically meaningful improvement at Week 24	n	12	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.8%	11.1%
		OR (95% CI) [b]	0.86 (0.32, 2.32)	
		p-value [b]	0.7729	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3336	
		RR (95% CI) [d]	0.88 (0.36, 2.12)	
		Risk Diff (95% CI) [e]	-1.4% (-10.7, 8.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.2%	16.7%
		OR (95% CI) [b]	0.76 (0.28, 2.09)	
		p-value [b]	0.5947	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4491	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1007 von 1722

Table 14.2.18.7.1.15.10 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.79 (0.34, 1.86)	
		Risk Diff (95% CI) [e]	-3.5% (-16.7, 9.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1008 von 1722

Table 14.2.18.7.1.15.10 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects at baseline	N	25	13
	Number of HRQoL evaluable subjects at Week 24	N	18	11
	Number of subjects with clinically meaningful improvement at Week 24	n	1	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.0%	15.4%
		OR (95% CI) [b]	0.23 (0.02, 2.80)	
		p-value [b]	0.2490	
		RR (95% CI) [d]	0.26 (0.03, 2.61)	
		Risk Diff (95% CI) [e]	-11.4% (-32.4, 9.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	5.6%	18.2%
		OR (95% CI) [b]	0.26 (0.02, 3.33)	
	p-value [b]	0.3038		
	RR (95% CI) [d]	0.31 (0.03, 2.99)		
	Risk Diff (95% CI) [e]	-12.6% (-37.8, 12.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1009 von 1722

Table 14.2.18.7.1.1.11 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	66.7%
		OR (95% CI) [b]	0.17 (0.02, 1.72)	
		p-value [b]	0.1322	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0663	
		RR (95% CI) [d]	0.38 (0.10, 1.41)	
		Risk Diff (95% CI) [e]	-41.7% (-89.9, 6.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%
	OR (95% CI) [b]	0.08 (0.01, 1.26)		
	p-value [b]	0.0727		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0534		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1010 von 1722

Table 14.2.18.7.1.1.11 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.31 (0.09, 1.12)	
		Risk Diff (95% CI) [e]	-55.0% (-100.0, -8.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1011 von 1722

Table 14.2.18.7.1.1.11 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	22	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.6%	9.8%
		OR (95% CI) [b]	1.96 (0.75, 5.11)	
		p-value [b]	0.1702	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7573	
		RR (95% CI) [d]	1.79 (0.77, 4.18)	
		Risk Diff (95% CI) [e]	7.8% (-2.3, 17.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.4%	14.3%
	OR (95% CI) [b]	1.94 (0.72, 5.22)		
	p-value [b]	0.1887		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8725		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1012 von 1722

Table 14.2.18.7.1.1.11 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.71 (0.75, 3.90)	
		Risk Diff (95% CI) [e]	10.2% (-3.7, 24.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1013 von 1722

Table 14.2.18.7.1.1.11 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	6
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	43.8%	22.2%
		OR (95% CI) [b]	2.72 (0.43, 17.42)	
		p-value [b]	0.2903	
		RR (95% CI) [d]	1.97 (0.51, 7.54)	
		Risk Diff (95% CI) [e]	21.5% (-14.9, 58.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	53.8%	33.3%
		OR (95% CI) [b]	2.33 (0.31, 17.54)	
	p-value [b]	0.4104		
	RR (95% CI) [d]	1.62 (0.47, 5.57)		
	Risk Diff (95% CI) [e]	20.5% (-25.9, 67.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1014 von 1722

Table 14.2.18.7.1.2.11 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	16.7%
		OR (95% CI) [b]	1.67 (0.11, 24.26)	
		p-value [b]	0.7085	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8325	
		RR (95% CI) [d]	1.50 (0.17, 12.94)	
		Risk Diff (95% CI) [e]	8.3% (-34.0, 50.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	20.0%
		OR (95% CI) [b]	1.33 (0.09, 20.11)	
		p-value [b]	0.8354	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9159	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.11 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.25 (0.15, 10.46)	
		Risk Diff (95% CI) [e]	5.0% (-41.1, 51.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.11 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	19	15
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.2%	24.6%
		OR (95% CI) [b]	0.55 (0.26, 1.18)	
		p-value [b]	0.1229	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4783	
		RR (95% CI) [d]	0.62 (0.34, 1.13)	
		Risk Diff (95% CI) [e]	-9.4% (-21.9, 3.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.1%	35.7%
	OR (95% CI) [b]	0.48 (0.21, 1.08)		
	p-value [b]	0.0768		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4514		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1017 von 1722

Table 14.2.18.7.1.2.11 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.59 (0.33, 1.04)	
		Risk Diff (95% CI) [e]	-14.6% (-31.4, 2.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1018 von 1722

Table 14.2.18.7.1.2.11 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	7
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	22.2%
		OR (95% CI) [b]	1.17 (0.17, 8.09)	
		p-value [b]	0.8760	
		RR (95% CI) [d]	1.13 (0.25, 4.98)	
		Risk Diff (95% CI) [e]	2.8% (-31.7, 37.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%	28.6%
		OR (95% CI) [b]	1.11 (0.15, 8.37)	
	p-value [b]	0.9185		
	RR (95% CI) [d]	1.08 (0.26, 4.49)		
	Risk Diff (95% CI) [e]	2.2% (-39.6, 44.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1019 von 1722

Table 14.2.18.7.1.3.11 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	1	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	33.3%
		OR (95% CI) [b]	0.29 (0.02, 4.24)	
		p-value [b]	0.3625	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6719	
		RR (95% CI) [d]	0.38 (0.04, 3.23)	
		Risk Diff (95% CI) [e]	-20.8% (-65.0, 23.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.5%
	OR (95% CI) [b]	0.21 (0.01, 3.37)		
	p-value [b]	0.2732		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6449		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.11 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.31 (0.04, 2.62)	
		Risk Diff (95% CI) [e]	-27.5% (-76.2, 21.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.11 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	24	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.2%	19.7%
		OR (95% CI) [b]	0.97 (0.45, 2.10)	
		p-value [b]	0.9390	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5719	
		RR (95% CI) [d]	0.98 (0.52, 1.82)	
		Risk Diff (95% CI) [e]	-0.5% (-12.6, 11.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.7%	28.6%
	OR (95% CI) [b]	0.91 (0.40, 2.06)		
	p-value [b]	0.8190		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5242		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1022 von 1722

Table 14.2.18.7.1.3.11 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.93 (0.52, 1.68)	
		Risk Diff (95% CI) [e]	-1.9% (-18.3, 14.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.11 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	7
	Number of subjects with clinically meaningful improvement at Week 24	n	5	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	31.3%	44.4%
		OR (95% CI) [b]	0.57 (0.11, 3.07)	
		p-value [b]	0.5113	
		RR (95% CI) [d]	0.70 (0.25, 1.97)	
		Risk Diff (95% CI) [e]	-13.2% (-52.8, 26.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	38.5%	57.1%
		OR (95% CI) [b]	0.47 (0.07, 3.04)	
	p-value [b]	0.4266		
	RR (95% CI) [d]	0.67 (0.26, 1.72)		
	Risk Diff (95% CI) [e]	-18.7% (-63.9, 26.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1024 von 1722

Table 14.2.18.7.1.4.11 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.06, 6.87)	
		p-value [b]	0.7334	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4519	
		RR (95% CI) [d]	0.75 (0.14, 3.90)	
		Risk Diff (95% CI) [e]	-8.3% (-56.5, 39.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%
	OR (95% CI) [b]	0.50 (0.05, 5.51)		
	p-value [b]	0.5711		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4409		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Table 14.2.18.7.1.4.11 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.63 (0.12, 3.13)	
		Risk Diff (95% CI) [e]	-15.0% (-67.4, 37.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Table 14.2.18.7.1.4.11 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	21	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.8%	16.4%
		OR (95% CI) [b]	1.03 (0.45, 2.35)	
		p-value [b]	0.9443	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4942	
		RR (95% CI) [d]	1.02 (0.52, 2.04)	
		Risk Diff (95% CI) [e]	0.4% (-11.0, 11.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.3%	23.8%
	OR (95% CI) [b]	0.97 (0.41, 2.31)		
	p-value [b]	0.9521		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6133		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Table 14.2.18.7.1.4.11 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.98 (0.51, 1.89)	
		Risk Diff (95% CI) [e]	-0.5% (-16.0, 15.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Table 14.2.18.7.1.4.11 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	6
	Number of subjects with clinically meaningful improvement at Week 24	n	6	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	37.5%	22.2%
		OR (95% CI) [b]	2.10 (0.32, 13.61)	
		p-value [b]	0.4366	
		RR (95% CI) [d]	1.69 (0.43, 6.68)	
		Risk Diff (95% CI) [e]	15.3% (-20.8, 51.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	46.2%	33.3%
		OR (95% CI) [b]	1.71 (0.23, 12.89)	
	p-value [b]	0.6004		
	RR (95% CI) [d]	1.38 (0.39, 4.95)		
	Risk Diff (95% CI) [e]	12.8% (-33.6, 59.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Table 14.2.18.7.1.5.11 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	16.7%
		OR (95% CI) [b]	1.67 (0.11, 24.26)	
		p-value [b]	0.7085	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8416	
		RR (95% CI) [d]	1.50 (0.17, 12.94)	
		Risk Diff (95% CI) [e]	8.3% (-34.0, 50.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	20.0%
		OR (95% CI) [b]	1.33 (0.09, 20.11)	
		p-value [b]	0.8354	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8423	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.11 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.25 (0.15, 10.46)	
		Risk Diff (95% CI) [e]	5.0% (-41.1, 51.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.11 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	14	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.2%	14.8%
		OR (95% CI) [b]	0.73 (0.30, 1.79)	
		p-value [b]	0.4906	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7445	
		RR (95% CI) [d]	0.76 (0.35, 1.66)	
		Risk Diff (95% CI) [e]	-3.6% (-14.0, 6.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.6%	21.4%
	OR (95% CI) [b]	0.68 (0.27, 1.71)		
	p-value [b]	0.4091		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8343		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1032 von 1722

Table 14.2.18.7.1.5.11 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.73 (0.34, 1.54)	
		Risk Diff (95% CI) [e]	-5.9% (-20.4, 8.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.11 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	6
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	11.1%
		OR (95% CI) [b]	1.14 (0.09, 14.68)	
		p-value [b]	0.9183	
		RR (95% CI) [d]	1.13 (0.12, 10.75)	
		Risk Diff (95% CI) [e]	1.4% (-24.8, 27.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.4%	16.7%
		OR (95% CI) [b]	0.91 (0.07, 12.52)	
	p-value [b]	0.9432		
	RR (95% CI) [d]	0.92 (0.10, 8.31)		
	Risk Diff (95% CI) [e]	-1.3% (-37.0, 34.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1034 von 1722

Table 14.2.18.7.1.6.11 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	0	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	16.7%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9818	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9824	
		RR (95% CI) [d]	0.26 (0.01, 5.44)	
		Risk Diff (95% CI) [e]	-15.9% (-49.8, 18.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%
	OR (95% CI) [b]	NE (NE, NE)		
	p-value [b]	0.9794		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9807		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1035 von 1722

Table 14.2.18.7.1.6.11 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.22 (0.01, 4.60)	
		Risk Diff (95% CI) [e]	-19.4% (-57.2, 18.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1036 von 1722

Table 14.2.18.7.1.6.11 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	22	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.6%	19.7%
		OR (95% CI) [b]	0.87 (0.40, 1.91)	
		p-value [b]	0.7315	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7874	
		RR (95% CI) [d]	0.89 (0.47, 1.69)	
		Risk Diff (95% CI) [e]	-2.1% (-14.1, 9.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.4%	28.6%
	OR (95% CI) [b]	0.81 (0.35, 1.84)		
	p-value [b]	0.6139		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5854		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1037 von 1722

Table 14.2.18.7.1.6.11 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.86 (0.47, 1.56)	
		Risk Diff (95% CI) [e]	-4.1% (-20.4, 12.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1038 von 1722

Table 14.2.18.7.1.6.11 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	6
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.11, 3.99)	
		p-value [b]	0.6569	
		RR (95% CI) [d]	0.75 (0.21, 2.63)	
		Risk Diff (95% CI) [e]	-8.3% (-45.7, 29.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%	50.0%
		OR (95% CI) [b]	0.44 (0.06, 3.24)	
	p-value [b]	0.4238		
	RR (95% CI) [d]	0.62 (0.20, 1.93)		
	Risk Diff (95% CI) [e]	-19.2% (-66.5, 28.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1039 von 1722

Table 14.2.18.7.1.7.11 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	1	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	50.0%
		OR (95% CI) [b]	0.14 (0.01, 1.99)	
		p-value [b]	0.1480	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5062	
		RR (95% CI) [d]	0.25 (0.03, 1.85)	
		Risk Diff (95% CI) [e]	-37.5% (-83.6, 8.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.5%
	OR (95% CI) [b]	0.10 (0.01, 1.50)		
	p-value [b]	0.0944		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4637		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.11 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.21 (0.03, 1.49)	
		Risk Diff (95% CI) [e]	-47.5% (-96.2, 1.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.11 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	89	42
	Number of subjects with clinically meaningful improvement at Week 24	n	27	17
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.6%	27.9%
		OR (95% CI) [b]	0.71 (0.35, 1.44)	
		p-value [b]	0.3461	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5737	
		RR (95% CI) [d]	0.78 (0.46, 1.31)	
		Risk Diff (95% CI) [e]	-6.3% (-19.6, 7.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.3%	40.5%
	OR (95% CI) [b]	0.64 (0.30, 1.38)		
	p-value [b]	0.2530		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5329		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.11 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.46, 1.22)	
		Risk Diff (95% CI) [e]	-10.1% (-27.8, 7.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.11 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	7
	Number of subjects with clinically meaningful improvement at Week 24	n	4	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	44.4%
		OR (95% CI) [b]	0.42 (0.07, 2.36)	
		p-value [b]	0.3226	
		RR (95% CI) [d]	0.56 (0.18, 1.72)	
		Risk Diff (95% CI) [e]	-19.4% (-58.2, 19.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%	57.1%
		OR (95% CI) [b]	0.33 (0.05, 2.24)	
	p-value [b]	0.2583		
	RR (95% CI) [d]	0.54 (0.19, 1.52)		
	Risk Diff (95% CI) [e]	-26.4% (-70.8, 18.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1044 von 1722

Table 14.2.18.7.1.8.11 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9549	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9950	
		RR (95% CI) [d]	3.89 (0.22, 68.67)	
		Risk Diff (95% CI) [e]	20.6% (-14.3, 55.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%
	OR (95% CI) [b]	NE (NE, NE)		
	p-value [b]	0.9538		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9955		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.11 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	3.33 (0.19, 57.93)	
		Risk Diff (95% CI) [e]	19.4% (-17.2, 56.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1046 von 1722

Table 14.2.18.7.1.8.11 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	14	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.2%	8.2%
		OR (95% CI) [b]	1.41 (0.48, 4.12)	
		p-value [b]	0.5271	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9544	
		RR (95% CI) [d]	1.37 (0.52, 3.62)	
		Risk Diff (95% CI) [e]	3.0% (-5.8, 11.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.6%
	OR (95% CI) [b]	1.36 (0.46, 4.07)		
	p-value [b]	0.5789		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9537		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1047 von 1722

Table 14.2.18.7.1.8.11 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.31 (0.50, 3.39)	
		Risk Diff (95% CI) [e]	3.7% (-8.7, 16.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1048 von 1722

Table 14.2.18.7.1.8.11 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	7
	Number of subjects with clinically meaningful improvement at Week 24	n	1	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.3%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9527	
		RR (95% CI) [d]	1.76 (0.08, 39.32)	
		Risk Diff (95% CI) [e]	3.8% (-15.3, 22.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.7%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9522		
	RR (95% CI) [d]	1.71 (0.08, 37.32)		
	Risk Diff (95% CI) [e]	4.5% (-18.9, 27.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1049 von 1722

Table 14.2.18.7.1.9.11 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	16.7%
		OR (95% CI) [b]	1.67 (0.11, 24.26)	
		p-value [b]	0.7085	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8325	
		RR (95% CI) [d]	1.50 (0.17, 12.94)	
		Risk Diff (95% CI) [e]	8.3% (-34.0, 50.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	20.0%
		OR (95% CI) [b]	1.33 (0.09, 20.11)	
		p-value [b]	0.8354	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9159	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.11 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.25 (0.15, 10.46)	
		Risk Diff (95% CI) [e]	5.0% (-41.1, 51.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1051 von 1722

Table 14.2.18.7.1.9.11 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	89	42
	Number of subjects with clinically meaningful improvement at Week 24	n	19	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.2%	18.0%
		OR (95% CI) [b]	0.81 (0.36, 1.84)	
		p-value [b]	0.6223	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7377	
		RR (95% CI) [d]	0.84 (0.43, 1.66)	
		Risk Diff (95% CI) [e]	-2.8% (-14.4, 8.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.3%	26.2%
	OR (95% CI) [b]	0.76 (0.33, 1.80)		
	p-value [b]	0.5388		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7385		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1052 von 1722

Table 14.2.18.7.1.9.11 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.82 (0.43, 1.55)	
		Risk Diff (95% CI) [e]	-4.8% (-20.6, 10.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1053 von 1722

Table 14.2.18.7.1.9.11 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	7
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	22.2%
		OR (95% CI) [b]	1.17 (0.17, 8.09)	
		p-value [b]	0.8760	
		RR (95% CI) [d]	1.13 (0.25, 4.98)	
		Risk Diff (95% CI) [e]	2.8% (-31.7, 37.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%	28.6%
		OR (95% CI) [b]	1.11 (0.15, 8.37)	
	p-value [b]	0.9184		
	RR (95% CI) [d]	1.08 (0.26, 4.49)		
	Risk Diff (95% CI) [e]	2.2% (-39.6, 44.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1054 von 1722

Table 14.2.18.7.1.10.11 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	50.0%
		OR (95% CI) [b]	0.33 (0.03, 3.20)	
		p-value [b]	0.3414	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5667	
		RR (95% CI) [d]	0.50 (0.12, 2.12)	
		Risk Diff (95% CI) [e]	-25.0% (-75.0, 25.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%
	OR (95% CI) [b]	0.22 (0.02, 2.45)		
	p-value [b]	0.2194		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4535		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.11 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.42 (0.10, 1.69)	
		Risk Diff (95% CI) [e]	-35.0% (-87.4, 17.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1056 von 1722

Table 14.2.18.7.1.10.11 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	86	42
	Number of subjects with clinically meaningful improvement at Week 24	n	19	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.2%	18.0%
		OR (95% CI) [b]	0.81 (0.36, 1.84)	
		p-value [b]	0.6223	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9937	
		RR (95% CI) [d]	0.84 (0.43, 1.66)	
		Risk Diff (95% CI) [e]	-2.8% (-14.4, 8.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.1%	26.2%
	OR (95% CI) [b]	0.80 (0.34, 1.88)		
	p-value [b]	0.6078		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9560		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1057 von 1722

Table 14.2.18.7.1.10.11 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.84 (0.44, 1.61)	
		Risk Diff (95% CI) [e]	-4.1% (-20.0, 11.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1058 von 1722

Table 14.2.18.7.1.10.11 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	7
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.8%	22.2%
		OR (95% CI) [b]	0.81 (0.11, 6.04)	
		p-value [b]	0.8351	
		RR (95% CI) [d]	0.84 (0.17, 4.15)	
		Risk Diff (95% CI) [e]	-3.5% (-36.7, 29.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.1%	28.6%
		OR (95% CI) [b]	0.75 (0.09, 6.04)	
	p-value [b]	0.7870		
	RR (95% CI) [d]	0.81 (0.17, 3.75)		
	Risk Diff (95% CI) [e]	-5.5% (-46.0, 35.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.11 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.06, 6.87)	
		p-value [b]	0.7334	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5788	
		RR (95% CI) [d]	0.75 (0.14, 3.90)	
		Risk Diff (95% CI) [e]	-8.3% (-56.5, 39.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%
	OR (95% CI) [b]	0.50 (0.05, 5.51)		
	p-value [b]	0.5714		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6147		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.11 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.63 (0.12, 3.13)	
		Risk Diff (95% CI) [e]	-15.0% (-67.4, 37.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1061 von 1722

Table 14.2.18.7.1.11.11 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	88	42
	Number of subjects with clinically meaningful improvement at Week 24	n	23	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.4%	19.7%
		OR (95% CI) [b]	0.92 (0.42, 2.00)	
		p-value [b]	0.8349	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2498	
		RR (95% CI) [d]	0.94 (0.50, 1.75)	
		Risk Diff (95% CI) [e]	-1.3% (-13.3, 10.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.1%	28.6%
	OR (95% CI) [b]	0.88 (0.39, 2.01)		
	p-value [b]	0.7698		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2099		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1062 von 1722

Table 14.2.18.7.1.11.11 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.91 (0.51, 1.66)	
		Risk Diff (95% CI) [e]	-2.4% (-18.9, 14.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.11 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	7
	Number of subjects with clinically meaningful improvement at Week 24	n	3	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.8%	44.4%
		OR (95% CI) [b]	0.29 (0.05, 1.78)	
		p-value [b]	0.1801	
		RR (95% CI) [d]	0.42 (0.12, 1.48)	
		Risk Diff (95% CI) [e]	-25.7% (-63.4, 12.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.1%	57.1%
		OR (95% CI) [b]	0.23 (0.03, 1.62)	
	p-value [b]	0.1390		
	RR (95% CI) [d]	0.40 (0.12, 1.32)		
	Risk Diff (95% CI) [e]	-34.1% (-77.3, 9.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.11 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	1	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	33.3%
		OR (95% CI) [b]	0.29 (0.02, 4.24)	
		p-value [b]	0.3625	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5448	
		RR (95% CI) [d]	0.38 (0.04, 3.23)	
		Risk Diff (95% CI) [e]	-20.8% (-65.0, 23.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.5%	40.0%
		OR (95% CI) [b]	0.21 (0.01, 3.37)	
		p-value [b]	0.2732	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4774	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.11 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.31 (0.04, 2.62)	
		Risk Diff (95% CI) [e]	-27.5% (-76.2, 21.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.11 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	89	41
	Number of subjects with clinically meaningful improvement at Week 24	n	18	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.4%	8.2%
		OR (95% CI) [b]	1.88 (0.66, 5.34)	
		p-value [b]	0.2336	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4637	
		RR (95% CI) [d]	1.76 (0.68, 4.51)	
		Risk Diff (95% CI) [e]	6.2% (-3.0, 15.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.2%
	OR (95% CI) [b]	1.83 (0.63, 5.32)		
	p-value [b]	0.2698		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4571		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.11 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.66 (0.66, 4.16)	
		Risk Diff (95% CI) [e]	8.0% (-5.0, 21.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1068 von 1722

Table 14.2.18.7.1.12.11 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	7
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.8%	22.2%
		OR (95% CI) [b]	0.81 (0.11, 6.04)	
		p-value [b]	0.8351	
		RR (95% CI) [d]	0.84 (0.17, 4.15)	
		Risk Diff (95% CI) [e]	-3.5% (-36.7, 29.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.1%	28.6%
		OR (95% CI) [b]	0.75 (0.09, 6.04)	
	p-value [b]	0.7870		
	RR (95% CI) [d]	0.81 (0.17, 3.75)		
	Risk Diff (95% CI) [e]	-5.5% (-46.0, 35.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.11 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	16.7%
		OR (95% CI) [b]	1.67 (0.11, 24.26)	
		p-value [b]	0.7085	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4928	
		RR (95% CI) [d]	1.50 (0.17, 12.94)	
		Risk Diff (95% CI) [e]	8.3% (-34.0, 50.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	20.0%
		OR (95% CI) [b]	1.33 (0.09, 20.11)	
		p-value [b]	0.8354	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4716	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.11 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.25 (0.15, 10.46)	
		Risk Diff (95% CI) [e]	5.0% (-41.1, 51.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.11 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	27	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.6%	16.4%
		OR (95% CI) [b]	1.41 (0.63, 3.13)	
		p-value [b]	0.4050	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3793	
		RR (95% CI) [d]	1.32 (0.68, 2.54)	
		Risk Diff (95% CI) [e]	5.2% (-6.6, 17.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.0%	23.8%
	OR (95% CI) [b]	1.37 (0.59, 3.18)		
	p-value [b]	0.4617		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2824		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.11 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.26 (0.67, 2.36)	
		Risk Diff (95% CI) [e]	6.2% (-9.8, 22.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.11 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	6
	Number of subjects with clinically meaningful improvement at Week 24	n	2	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	22.2%
		OR (95% CI) [b]	0.50 (0.06, 4.33)	
		p-value [b]	0.5293	
		RR (95% CI) [d]	0.56 (0.09, 3.34)	
		Risk Diff (95% CI) [e]	-9.7% (-41.4, 21.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.4%	33.3%
		OR (95% CI) [b]	0.36 (0.04, 3.52)	
	p-value [b]	0.3823		
	RR (95% CI) [d]	0.46 (0.08, 2.54)		
	Risk Diff (95% CI) [e]	-17.9% (-60.5, 24.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.11 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24		0	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
		p-value [b]	1.0000	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	1.0000	
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	0.0%
	OR (95% CI) [b]	1.00 (NE, NE)		
	p-value [b]	1.0000		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	1.0000		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.11 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.11 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	90	42
	Number of subjects with clinically meaningful improvement at Week 24	n	11	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.8%	9.8%
		OR (95% CI) [b]	0.88 (0.31, 2.52)	
		p-value [b]	0.8180	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9996	
		RR (95% CI) [d]	0.89 (0.35, 2.31)	
		Risk Diff (95% CI) [e]	-1.0% (-10.0, 7.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.2%	14.3%
	OR (95% CI) [b]	0.84 (0.29, 2.44)		
	p-value [b]	0.7419		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9990		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1077 von 1722

Table 14.2.18.7.1.14.11 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.86 (0.34, 2.16)	
		Risk Diff (95% CI) [e]	-2.1% (-14.6, 10.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1078 von 1722

Table 14.2.18.7.1.14.11 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	6
	Number of subjects with clinically meaningful improvement at Week 24		0	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
		p-value [b]	1.0000	
		RR (95% CI) [d]	NE (NE, NE)	
		Risk Diff (95% CI) [e]	NE (NE, NE)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	0.0%
		OR (95% CI) [b]	1.00 (NE, NE)	
	p-value [b]	1.0000		
	RR (95% CI) [d]	NE (NE, NE)		
	Risk Diff (95% CI) [e]	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.11 (AMNOG)  
EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	8	6
	Number of HRQoL evaluable subjects at Week 24	N	8	5
	Number of subjects with clinically meaningful improvement at Week 24	n	1	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	16.7%
		OR (95% CI) [b]	0.71 (0.04, 14.35)	
		p-value [b]	0.8260	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8909	
		RR (95% CI) [d]	0.75 (0.06, 9.72)	
		Risk Diff (95% CI) [e]	-4.2% (-41.8, 33.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.5%	20.0%
		OR (95% CI) [b]	0.57 (0.03, 11.85)	
		p-value [b]	0.7175	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8839	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.11 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.63 (0.05, 7.90)	
		Risk Diff (95% CI) [e]	-7.5% (-49.4, 34.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1081 von 1722

Table 14.2.18.7.1.15.11 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	125	61
	Number of HRQoL evaluable subjects at Week 24	N	88	42
	Number of subjects with clinically meaningful improvement at Week 24	n	11	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.8%	11.5%
		OR (95% CI) [b]	0.74 (0.27, 2.03)	
		p-value [b]	0.5634	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8314	
		RR (95% CI) [d]	0.77 (0.31, 1.88)	
		Risk Diff (95% CI) [e]	-2.7% (-12.1, 6.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.5%	16.7%
	OR (95% CI) [b]	0.71 (0.26, 2.00)		
	p-value [b]	0.5214		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7362		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.11 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.31, 1.80)	
		Risk Diff (95% CI) [e]	-4.2% (-17.4, 9.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.11 (AMNOG)  
EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects at baseline	N	16	9
	Number of HRQoL evaluable subjects at Week 24	N	13	6
	Number of subjects with clinically meaningful improvement at Week 24	n	1	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.3%	11.1%
		OR (95% CI) [b]	0.53 (0.03, 9.71)	
		p-value [b]	0.6711	
		RR (95% CI) [d]	0.56 (0.04, 7.95)	
		Risk Diff (95% CI) [e]	-4.9% (-28.6, 18.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.7%	16.7%
		OR (95% CI) [b]	0.42 (0.02, 8.05)	
	p-value [b]	0.5623		
	RR (95% CI) [d]	0.46 (0.03, 6.20)		
	Risk Diff (95% CI) [e]	-9.0% (-42.1, 24.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1084 von 1722

Table 14.2.18.7.1.1.12 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	9	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.1%	23.1%
		OR (95% CI) [b]	1.00 (0.23, 4.44)	
		p-value [b]	1.0000	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6583	
		RR (95% CI) [d]	1.00 (0.32, 3.15)	
		Risk Diff (95% CI) [e]	0.0% (-26.4, 26.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.1%
	OR (95% CI) [b]	0.65 (0.13, 3.31)		
	p-value [b]	0.6059		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3609		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1085 von 1722

Table 14.2.18.7.1.1.12 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.26, 2.15)	
		Risk Diff (95% CI) [e]	-9.4% (-46.4, 27.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.12 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	45
	Number of subjects with clinically meaningful improvement at Week 24	n	22	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	14.5%
		OR (95% CI) [b]	1.47 (0.63, 3.43)	
		p-value [b]	0.3708	
		RR (95% CI) [d]	1.38 (0.68, 2.80)	
		Risk Diff (95% CI) [e]	5.5% (-6.0, 17.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.8%	20.0%
		OR (95% CI) [b]	1.54 (0.64, 3.72)	
	p-value [b]	0.3340		
	RR (95% CI) [d]	1.39 (0.70, 2.76)		
	Risk Diff (95% CI) [e]	7.8% (-7.5, 23.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1087 von 1722

Table 14.2.18.7.1.2.12 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	11	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.2%	23.1%
		OR (95% CI) [b]	1.31 (0.30, 5.68)	
		p-value [b]	0.7186	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2178	
		RR (95% CI) [d]	1.22 (0.40, 3.71)	
		Risk Diff (95% CI) [e]	5.1% (-21.8, 32.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.4%	37.5%
	OR (95% CI) [b]	0.87 (0.18, 4.35)		
	p-value [b]	0.8684		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4670		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.12 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.92 (0.33, 2.53)	
		Risk Diff (95% CI) [e]	-3.1% (-40.5, 34.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.12 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	46
	Number of subjects with clinically meaningful improvement at Week 24	n	14	15
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.7%	24.2%
		OR (95% CI) [b]	0.46 (0.20, 1.02)	
		p-value [b]	0.0573	
		RR (95% CI) [d]	0.53 (0.27, 1.02)	
		Risk Diff (95% CI) [e]	-11.5% (-23.8, 0.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.7%	32.6%
		OR (95% CI) [b]	0.45 (0.19, 1.04)	
	p-value [b]	0.0604		
	RR (95% CI) [d]	0.54 (0.29, 1.02)		
	Risk Diff (95% CI) [e]	-14.9% (-30.8, 1.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1090 von 1722

Table 14.2.18.7.1.3.12 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	8	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.5%	15.4%
		OR (95% CI) [b]	1.42 (0.26, 7.72)	
		p-value [b]	0.6861	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4709	
		RR (95% CI) [d]	1.33 (0.32, 5.50)	
		Risk Diff (95% CI) [e]	5.1% (-18.2, 28.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%	25.0%
		OR (95% CI) [b]	1.00 (0.17, 5.98)	
		p-value [b]	1.0000	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7454	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1091 von 1722

Table 14.2.18.7.1.3.12 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.00 (0.26, 3.83)	
		Risk Diff (95% CI) [e]	0.0% (-33.5, 33.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.12 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	46
	Number of subjects with clinically meaningful improvement at Week 24	n	22	16
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	25.8%
		OR (95% CI) [b]	0.72 (0.34, 1.50)	
		p-value [b]	0.3791	
		RR (95% CI) [d]	0.78 (0.44, 1.36)	
		Risk Diff (95% CI) [e]	-5.8% (-19.0, 7.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.8%	34.8%
		OR (95% CI) [b]	0.72 (0.33, 1.58)	
	p-value [b]	0.4171		
	RR (95% CI) [d]	0.80 (0.47, 1.36)		
	Risk Diff (95% CI) [e]	-6.9% (-23.9, 10.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1093 von 1722

Table 14.2.18.7.1.4.12 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	6	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.4%	7.7%
		OR (95% CI) [b]	2.18 (0.24, 20.02)	
		p-value [b]	0.4907	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5128	
		RR (95% CI) [d]	2.00 (0.26, 15.10)	
		Risk Diff (95% CI) [e]	7.7% (-10.7, 26.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.8%
	OR (95% CI) [b]	1.62 (0.17, 15.72)		
	p-value [b]	0.6796		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7037		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1094 von 1722

Table 14.2.18.7.1.4.12 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.50 (0.21, 10.76)	
		Risk Diff (95% CI) [e]	6.3% (-20.4, 32.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.12 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance ≥ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	45
	Number of subjects with clinically meaningful improvement at Week 24	n	23	13
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.9%	21.0%
		OR (95% CI) [b]	1.00 (0.46, 2.14)	
		p-value [b]	0.9928	
		RR (95% CI) [d]	1.00 (0.54, 1.83)	
		Risk Diff (95% CI) [e]	-0.1% (-12.7, 12.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.1%	28.9%
		OR (95% CI) [b]	1.01 (0.45, 2.27)	
	p-value [b]	0.9788		
	RR (95% CI) [d]	1.01 (0.57, 1.79)		
	Risk Diff (95% CI) [e]	0.2% (-16.4, 16.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.12 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	2	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.1%	15.4%
		OR (95% CI) [b]	0.30 (0.04, 2.36)	
		p-value [b]	0.2513	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2903	
		RR (95% CI) [d]	0.33 (0.05, 2.13)	
		Risk Diff (95% CI) [e]	-10.3% (-31.1, 10.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	6.3%	25.0%
	OR (95% CI) [b]	0.20 (0.02, 1.71)		
	p-value [b]	0.1418		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1722		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1097 von 1722

Table 14.2.18.7.1.5.12 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.25 (0.04, 1.51)	
		Risk Diff (95% CI) [e]	-18.8% (-49.9, 12.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.12 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	45
	Number of subjects with clinically meaningful improvement at Week 24	n	16	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.5%	14.5%
		OR (95% CI) [b]	1.00 (0.41, 2.42)	
		p-value [b]	0.9958	
		RR (95% CI) [d]	1.00 (0.47, 2.13)	
		Risk Diff (95% CI) [e]	0.0% (-10.9, 11.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.3%	20.0%
		OR (95% CI) [b]	1.02 (0.41, 2.53)	
	p-value [b]	0.9730		
	RR (95% CI) [d]	1.01 (0.49, 2.10)		
	Risk Diff (95% CI) [e]	0.3% (-14.4, 14.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1099 von 1722

Table 14.2.18.7.1.6.12 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	15.4%
		OR (95% CI) [b]	1.20 (0.22, 6.68)	
		p-value [b]	0.8326	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5885	
		RR (95% CI) [d]	1.17 (0.28, 4.93)	
		Risk Diff (95% CI) [e]	2.6% (-20.5, 25.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%
	OR (95% CI) [b]	0.84 (0.14, 5.11)		
	p-value [b]	0.8500		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8582		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.12 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.88 (0.22, 3.43)	
		Risk Diff (95% CI) [e]	-3.1% (-36.4, 30.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1101 von 1722

Table 14.2.18.7.1.6.12 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	45
	Number of subjects with clinically meaningful improvement at Week 24	n	19	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.3%	22.6%
		OR (95% CI) [b]	0.72 (0.33, 1.55)	
		p-value [b]	0.3971	
		RR (95% CI) [d]	0.76 (0.41, 1.42)	
		Risk Diff (95% CI) [e]	-5.3% (-17.9, 7.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.1%	31.1%
		OR (95% CI) [b]	0.70 (0.31, 1.58)	
	p-value [b]	0.3933		
	RR (95% CI) [d]	0.77 (0.43, 1.39)		
	Risk Diff (95% CI) [e]	-7.1% (-23.5, 9.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1102 von 1722

Table 14.2.18.7.1.7.12 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	10	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.6%	30.8%
		OR (95% CI) [b]	0.78 (0.20, 3.08)	
		p-value [b]	0.7184	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6217	
		RR (95% CI) [d]	0.83 (0.31, 2.21)	
		Risk Diff (95% CI) [e]	-5.1% (-33.7, 23.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.3%
	OR (95% CI) [b]	0.45 (0.09, 2.20)		
	p-value [b]	0.3264		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8962		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1103 von 1722

Table 14.2.18.7.1.7.12 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.63 (0.26, 1.48)	
		Risk Diff (95% CI) [e]	-18.8% (-56.9, 19.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1104 von 1722

Table 14.2.18.7.1.7.12 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	78	46
	Number of subjects with clinically meaningful improvement at Week 24	n	22	20
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	32.3%
		OR (95% CI) [b]	0.52 (0.26, 1.07)	
		p-value [b]	0.0746	
		RR (95% CI) [d]	0.62 (0.37, 1.04)	
		Risk Diff (95% CI) [e]	-12.3% (-26.1, 1.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.2%	43.5%
		OR (95% CI) [b]	0.51 (0.24, 1.10)	
	p-value [b]	0.0846		
	RR (95% CI) [d]	0.65 (0.40, 1.05)		
	Risk Diff (95% CI) [e]	-15.3% (-32.7, 2.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1105 von 1722

Table 14.2.18.7.1.8.12 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	15.4%
		OR (95% CI) [b]	1.20 (0.22, 6.68)	
		p-value [b]	0.8325	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6572	
		RR (95% CI) [d]	1.17 (0.28, 4.93)	
		Risk Diff (95% CI) [e]	2.6% (-20.5, 25.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%
	OR (95% CI) [b]	0.84 (0.14, 5.11)		
	p-value [b]	0.8499		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4308		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1106 von 1722

Table 14.2.18.7.1.8.12 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.88 (0.22, 3.43)	
		Risk Diff (95% CI) [e]	-3.1% (-36.4, 30.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.12 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance ≥ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	46
	Number of subjects with clinically meaningful improvement at Week 24	n	10	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.1%	4.8%
		OR (95% CI) [b]	1.97 (0.52, 7.43)	
		p-value [b]	0.3189	
		RR (95% CI) [d]	1.88 (0.54, 6.57)	
		Risk Diff (95% CI) [e]	4.3% (-3.3, 11.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.7%	6.5%
		OR (95% CI) [b]	2.08 (0.54, 7.97)	
	p-value [b]	0.2869		
	RR (95% CI) [d]	1.94 (0.56, 6.69)		
	Risk Diff (95% CI) [e]	6.1% (-4.1, 16.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.12 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.5%	23.1%
		OR (95% CI) [b]	0.86 (0.19, 3.88)	
		p-value [b]	0.8442	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9868	
		RR (95% CI) [d]	0.89 (0.28, 2.86)	
		Risk Diff (95% CI) [e]	-2.6% (-28.7, 23.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.0%
	OR (95% CI) [b]	0.56 (0.11, 2.86)		
	p-value [b]	0.4823		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6211		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1109 von 1722

Table 14.2.18.7.1.9.12 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.67 (0.23, 1.96)	
		Risk Diff (95% CI) [e]	-12.5% (-49.2, 24.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.12 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	78	46
	Number of subjects with clinically meaningful improvement at Week 24	n	17	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.5%	17.7%
		OR (95% CI) [b]	0.85 (0.37, 1.95)	
		p-value [b]	0.6966	
		RR (95% CI) [d]	0.87 (0.44, 1.74)	
		Risk Diff (95% CI) [e]	-2.3% (-14.0, 9.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.8%	23.9%
		OR (95% CI) [b]	0.89 (0.37, 2.11)	
	p-value [b]	0.7853		
	RR (95% CI) [d]	0.91 (0.47, 1.77)		
	Risk Diff (95% CI) [e]	-2.1% (-17.5, 13.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1111 von 1722

Table 14.2.18.7.1.10.12 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	31	8
	Number of subjects with clinically meaningful improvement at Week 24	n	8	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.5%	30.8%
		OR (95% CI) [b]	0.58 (0.14, 2.38)	
		p-value [b]	0.4502	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8104	
		RR (95% CI) [d]	0.67 (0.24, 1.85)	
		Risk Diff (95% CI) [e]	-10.3% (-38.4, 17.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.8%
	OR (95% CI) [b]	0.35 (0.07, 1.73)		
	p-value [b]	0.1965		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4030		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1112 von 1722

Table 14.2.18.7.1.10.12 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.52 (0.21, 1.29)	
		Risk Diff (95% CI) [e]	-24.2% (-62.1, 13.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.12 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance ≥ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	76	46
	Number of subjects with clinically meaningful improvement at Week 24	n	16	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.5%	19.4%
		OR (95% CI) [b]	0.71 (0.31, 1.62)	
		p-value [b]	0.4134	
		RR (95% CI) [d]	0.75 (0.38, 1.48)	
		Risk Diff (95% CI) [e]	-4.8% (-16.6, 7.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.1%	26.1%
		OR (95% CI) [b]	0.76 (0.32, 1.78)	
	p-value [b]	0.5223		
	RR (95% CI) [d]	0.81 (0.42, 1.55)		
	Risk Diff (95% CI) [e]	-5.0% (-20.7, 10.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1114 von 1722

Table 14.2.18.7.1.11.12 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	31	8
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.8%	15.4%
		OR (95% CI) [b]	0.81 (0.14, 4.77)	
		p-value [b]	0.8148	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9494	
		RR (95% CI) [d]	0.83 (0.18, 3.79)	
		Risk Diff (95% CI) [e]	-2.6% (-24.8, 19.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.1%
	OR (95% CI) [b]	0.58 (0.09, 3.72)		
	p-value [b]	0.5632		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7660		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1115 von 1722

Table 14.2.18.7.1.11.12 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.65 (0.15, 2.73)	
		Risk Diff (95% CI) [e]	-8.9% (-41.6, 23.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1116 von 1722

Table 14.2.18.7.1.11.12 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	78	46
	Number of subjects with clinically meaningful improvement at Week 24	n	23	16
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.9%	25.8%
		OR (95% CI) [b]	0.76 (0.37, 1.58)	
		p-value [b]	0.4621	
		RR (95% CI) [d]	0.81 (0.46, 1.41)	
		Risk Diff (95% CI) [e]	-4.9% (-18.2, 8.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	29.5%	34.8%
		OR (95% CI) [b]	0.78 (0.36, 1.71)	
		p-value [b]	0.5399	
		RR (95% CI) [d]	0.85 (0.50, 1.43)	
		Risk Diff (95% CI) [e]	-5.3% (-22.4, 11.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1117 von 1722

Table 14.2.18.7.1.12.12 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.9%	15.4%
		OR (95% CI) [b]	1.20 (0.22, 6.68)	
		p-value [b]	0.8326	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9756	
		RR (95% CI) [d]	1.17 (0.28, 4.93)	
		Risk Diff (95% CI) [e]	2.6% (-20.5, 25.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.9%
	OR (95% CI) [b]	0.84 (0.14, 5.11)		
	p-value [b]	0.8497		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6812		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1118 von 1722

Table 14.2.18.7.1.12.12 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.88 (0.22, 3.43)	
		Risk Diff (95% CI) [e]	-3.1% (-36.4, 30.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.12 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	78	45
	Number of subjects with clinically meaningful improvement at Week 24	n	15	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.6%	11.3%
		OR (95% CI) [b]	1.24 (0.48, 3.23)	
		p-value [b]	0.6587	
		RR (95% CI) [d]	1.21 (0.52, 2.80)	
		Risk Diff (95% CI) [e]	2.3% (-7.8, 12.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.2%	15.6%
		OR (95% CI) [b]	1.29 (0.48, 3.45)	
	p-value [b]	0.6092		
	RR (95% CI) [d]	1.24 (0.55, 2.80)		
	Risk Diff (95% CI) [e]	3.7% (-10.1, 17.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1120 von 1722

Table 14.2.18.7.1.13.12 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	10	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.6%	15.4%
		OR (95% CI) [b]	1.90 (0.36, 10.07)	
		p-value [b]	0.4523	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5607	
		RR (95% CI) [d]	1.67 (0.42, 6.64)	
		Risk Diff (95% CI) [e]	10.3% (-13.7, 34.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	31.3%
	OR (95% CI) [b]	1.36 (0.23, 7.98)		
	p-value [b]	0.7307		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8431		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1121 von 1722

Table 14.2.18.7.1.13.12 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.25 (0.34, 4.61)	
		Risk Diff (95% CI) [e]	6.3% (-27.8, 40.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1122 von 1722

Table 14.2.18.7.1.13.12 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance ≥ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	45
	Number of subjects with clinically meaningful improvement at Week 24	n	21	11
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	17.7%
		OR (95% CI) [b]	1.09 (0.49, 2.45)	
		p-value [b]	0.8273	
		RR (95% CI) [d]	1.08 (0.56, 2.08)	
		Risk Diff (95% CI) [e]	1.3% (-10.7, 13.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.6%	24.4%
		OR (95% CI) [b]	1.12 (0.48, 2.60)	
	p-value [b]	0.7937		
	RR (95% CI) [d]	1.09 (0.58, 2.04)		
	Risk Diff (95% CI) [e]	2.1% (-13.8, 18.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1123 von 1722

Table 14.2.18.7.1.14.12 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	32	8
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.1%	7.7%
		OR (95% CI) [b]	0.65 (0.05, 7.80)	
		p-value [b]	0.7330	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7480	
		RR (95% CI) [d]	0.67 (0.07, 6.76)	
		Risk Diff (95% CI) [e]	-2.6% (-18.6, 13.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	6.3%
	OR (95% CI) [b]	0.47 (0.04, 5.90)		
	p-value [b]	0.5561		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5788		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.12 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.50 (0.05, 4.85)	
		Risk Diff (95% CI) [e]	-6.3% (-30.7, 18.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.12 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance ≥ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	79	45
	Number of subjects with clinically meaningful improvement at Week 24	n	9	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.2%	8.1%
		OR (95% CI) [b]	1.02 (0.32, 3.18)	
		p-value [b]	0.9784	
		RR (95% CI) [d]	1.01 (0.36, 2.89)	
		Risk Diff (95% CI) [e]	0.1% (-8.4, 8.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.4%	11.1%
		OR (95% CI) [b]	1.03 (0.32, 3.28)	
		p-value [b]	0.9620	
		RR (95% CI) [d]	1.03 (0.37, 2.87)	
		Risk Diff (95% CI) [e]	0.3% (-11.3, 11.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.12 (AMNOG)  
EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance <60	Number of HRQoL evaluable subjects at baseline	N	39	13
	Number of HRQoL evaluable subjects at Week 24	N	31	8
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.1%	23.1%
		OR (95% CI) [b]	0.18 (0.03, 1.23)	
		p-value [b]	0.0802	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1168	
		RR (95% CI) [d]	0.22 (0.04, 1.19)	
		Risk Diff (95% CI) [e]	-17.9% (-41.9, 6.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	6.5%
	OR (95% CI) [b]	0.11 (0.02, 0.87)		
	p-value [b]	0.0363		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0566		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.12 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.17 (0.03, 0.86)	
		Risk Diff (95% CI) [e]	-31.0% (-65.7, 3.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1128 von 1722

Table 14.2.18.7.1.15.12 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance ≥ 60	Number of HRQoL evaluable subjects at baseline	N	110	62
	Number of HRQoL evaluable subjects at Week 24	N	78	45
	Number of subjects with clinically meaningful improvement at Week 24	n	11	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.0%	9.7%
		OR (95% CI) [b]	1.04 (0.36, 2.96)	
		p-value [b]	0.9457	
		RR (95% CI) [d]	1.03 (0.40, 2.66)	
		Risk Diff (95% CI) [e]	0.3% (-8.9, 9.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.1%	13.3%
		OR (95% CI) [b]	1.07 (0.37, 3.11)	
	p-value [b]	0.9052		
	RR (95% CI) [d]	1.06 (0.42, 2.67)		
	Risk Diff (95% CI) [e]	0.8% (-11.8, 13.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.13 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	11	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.0%	16.1%
		OR (95% CI) [b]	1.47 (0.46, 4.72)	
		p-value [b]	0.5204	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7729	
		RR (95% CI) [d]	1.36 (0.52, 3.55)	
		Risk Diff (95% CI) [e]	5.9% (-11.4, 23.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.9%
	OR (95% CI) [b]	1.47 (0.44, 4.94)		
	p-value [b]	0.5363		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6465		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.13 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.33 (0.53, 3.35)	
		Risk Diff (95% CI) [e]	7.2% (-15.0, 29.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.13 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	7	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.4%	21.1%
		OR (95% CI) [b]	0.91 (0.23, 3.59)	
		p-value [b]	0.8873	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8951	
		RR (95% CI) [d]	0.92 (0.31, 2.76)	
		Risk Diff (95% CI) [e]	-1.6% (-24.0, 20.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.9%
	OR (95% CI) [b]	0.64 (0.14, 2.90)		
	p-value [b]	0.5672		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8206		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.13 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.74 (0.27, 2.02)	
		Risk Diff (95% CI) [e]	-9.4% (-42.6, 23.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1133 von 1722

Table 14.2.18.7.1.1.13 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	8	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.5%	6.7%
		OR (95% CI) [b]	3.39 (0.39, 29.75)	
		p-value [b]	0.2699	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4266	
		RR (95% CI) [d]	2.93 (0.40, 21.47)	
		Risk Diff (95% CI) [e]	12.8% (-4.7, 30.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.6%	9.1%
		OR (95% CI) [b]	4.00 (0.44, 36.51)	
		p-value [b]	0.2197	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3060	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1134 von 1722

Table 14.2.18.7.1.1.13 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	3.14 (0.44, 22.28)	
		Risk Diff (95% CI) [e]	19.5% (-4.4, 43.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1135 von 1722

Table 14.2.18.7.1.1.13 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.0%	18.2%
		OR (95% CI) [b]	1.06 (0.16, 6.94)	
		p-value [b]	0.9525	
		RR (95% CI) [d]	1.05 (0.23, 4.85)	
		Risk Diff (95% CI) [e]	0.9% (-27.4, 29.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	25.0%
		OR (95% CI) [b]	0.86 (0.12, 6.01)	
	p-value [b]	0.8768		
	RR (95% CI) [d]	0.89 (0.20, 3.90)		
	Risk Diff (95% CI) [e]	-2.8% (-38.4, 32.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1136 von 1722

Table 14.2.18.7.1.2.13 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	9	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.0%	9.7%
		OR (95% CI) [b]	2.05 (0.51, 8.24)	
		p-value [b]	0.3126	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3125	
		RR (95% CI) [d]	1.86 (0.55, 6.35)	
		Risk Diff (95% CI) [e]	8.3% (-6.6, 23.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.7%
	OR (95% CI) [b]	2.07 (0.50, 8.61)		
	p-value [b]	0.3175		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2092		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.13 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.82 (0.55, 6.03)	
		Risk Diff (95% CI) [e]	10.6% (-8.7, 29.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.13 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	12
	Number of subjects with clinically meaningful improvement at Week 24	n	5	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.9%	36.8%
		OR (95% CI) [b]	0.28 (0.07, 1.04)	
		p-value [b]	0.0576	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3724	
		RR (95% CI) [d]	0.38 (0.14, 1.03)	
		Risk Diff (95% CI) [e]	-23.0% (-47.4, 1.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.2%
	OR (95% CI) [b]	0.17 (0.04, 0.77)		
	p-value [b]	0.0211		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3516		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1139 von 1722

Table 14.2.18.7.1.2.13 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.33 (0.13, 0.83)	
		Risk Diff (95% CI) [e]	-39.1% (-70.8, -7.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1140 von 1722

Table 14.2.18.7.1.2.13 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	4	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.8%	26.7%
		OR (95% CI) [b]	0.30 (0.06, 1.39)	
		p-value [b]	0.1228	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4427	
		RR (95% CI) [d]	0.37 (0.10, 1.28)	
		Risk Diff (95% CI) [e]	-16.9% (-41.1, 7.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.3%
	OR (95% CI) [b]	0.29 (0.06, 1.48)		
	p-value [b]	0.1364		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6527		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1141 von 1722

Table 14.2.18.7.1.2.13 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.39 (0.12, 1.30)	
		Risk Diff (95% CI) [e]	-22.1% (-53.3, 9.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.13 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	28.6%	36.4%
		OR (95% CI) [b]	0.70 (0.15, 3.30)	
		p-value [b]	0.6522	
		RR (95% CI) [d]	0.79 (0.28, 2.21)	
		Risk Diff (95% CI) [e]	-7.8% (-42.2, 26.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	33.3%	50.0%
		OR (95% CI) [b]	0.50 (0.09, 2.73)	
	p-value [b]	0.4235		
	RR (95% CI) [d]	0.67 (0.26, 1.73)		
	Risk Diff (95% CI) [e]	-16.7% (-57.6, 24.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1143 von 1722

Table 14.2.18.7.1.3.13 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	9	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.0%	16.1%
		OR (95% CI) [b]	1.14 (0.34, 3.78)	
		p-value [b]	0.8287	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2333	
		RR (95% CI) [d]	1.12 (0.41, 3.03)	
		Risk Diff (95% CI) [e]	1.9% (-14.9, 18.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.7%
	OR (95% CI) [b]	1.12 (0.32, 3.87)		
	p-value [b]	0.8610		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1699		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1144 von 1722

Table 14.2.18.7.1.3.13 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.09 (0.42, 2.85)	
		Risk Diff (95% CI) [e]	1.9% (-19.7, 23.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1145 von 1722

Table 14.2.18.7.1.3.13 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	12
	Number of subjects with clinically meaningful improvement at Week 24	n	8	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.2%	42.1%
		OR (95% CI) [b]	0.39 (0.12, 1.31)	
		p-value [b]	0.1279	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7755	
		RR (95% CI) [d]	0.53 (0.24, 1.18)	
		Risk Diff (95% CI) [e]	-19.9% (-45.9, 6.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%
	OR (95% CI) [b]	0.22 (0.05, 0.96)		
	p-value [b]	0.0436		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9602		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1146 von 1722

Table 14.2.18.7.1.3.13 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.46 (0.23, 0.93)	
		Risk Diff (95% CI) [e]	-35.9% (-67.9, -3.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.13 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	10	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	24.4%	13.3%
		OR (95% CI) [b]	2.10 (0.40, 10.92)	
		p-value [b]	0.3793	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1252	
		RR (95% CI) [d]	1.83 (0.45, 7.41)	
		Risk Diff (95% CI) [e]	11.1% (-10.6, 32.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	35.7%	18.2%
		OR (95% CI) [b]	2.50 (0.45, 13.91)	
		p-value [b]	0.2953	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0687	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1148 von 1722

Table 14.2.18.7.1.3.13 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.96 (0.51, 7.57)	
		Risk Diff (95% CI) [e]	17.5% (-11.4, 46.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1149 von 1722

Table 14.2.18.7.1.3.13 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.5%	27.3%
		OR (95% CI) [b]	0.28 (0.04, 2.01)	
		p-value [b]	0.2064	
		RR (95% CI) [d]	0.35 (0.07, 1.79)	
		Risk Diff (95% CI) [e]	-17.7% (-46.9, 11.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.1%	37.5%
		OR (95% CI) [b]	0.21 (0.03, 1.62)	
	p-value [b]	0.1341		
	RR (95% CI) [d]	0.30 (0.06, 1.44)		
	Risk Diff (95% CI) [e]	-26.4% (-62.9, 10.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1150 von 1722

Table 14.2.18.7.1.4.13 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	10	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	16.1%
		OR (95% CI) [b]	1.30 (0.40, 4.24)	
		p-value [b]	0.6634	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6559	
		RR (95% CI) [d]	1.24 (0.47, 3.29)	
		Risk Diff (95% CI) [e]	3.9% (-13.2, 20.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.3%
	OR (95% CI) [b]	1.29 (0.38, 4.38)		
	p-value [b]	0.6879		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7457		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1151 von 1722

Table 14.2.18.7.1.4.13 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.21 (0.47, 3.10)	
		Risk Diff (95% CI) [e]	4.6% (-17.3, 26.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.13 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	9	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	26.3%
		OR (95% CI) [b]	0.93 (0.26, 3.32)	
		p-value [b]	0.9152	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4941	
		RR (95% CI) [d]	0.95 (0.37, 2.44)	
		Risk Diff (95% CI) [e]	-1.3% (-25.6, 23.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.6%	45.5%
		OR (95% CI) [b]	0.64 (0.15, 2.67)	
		p-value [b]	0.5357	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4175	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1153 von 1722

Table 14.2.18.7.1.4.13 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.76 (0.33, 1.76)	
		Risk Diff (95% CI) [e]	-10.8% (-45.5, 23.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.13 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.2%	20.0%
		OR (95% CI) [b]	0.56 (0.12, 2.68)	
		p-value [b]	0.4641	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3138	
		RR (95% CI) [d]	0.61 (0.17, 2.25)	
		Risk Diff (95% CI) [e]	-7.8% (-30.4, 14.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.9%
	OR (95% CI) [b]	0.58 (0.11, 2.99)		
	p-value [b]	0.5152		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4001		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1155 von 1722

Table 14.2.18.7.1.4.13 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.65 (0.19, 2.29)	
		Risk Diff (95% CI) [e]	-9.4% (-39.3, 20.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.13 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	4	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.0%	9.1%
		OR (95% CI) [b]	2.35 (0.23, 24.09)	
		p-value [b]	0.4710	
		RR (95% CI) [d]	2.10 (0.27, 16.54)	
		Risk Diff (95% CI) [e]	10.0% (-13.9, 33.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	12.5%
		OR (95% CI) [b]	2.00 (0.19, 21.43)	
	p-value [b]	0.5668		
	RR (95% CI) [d]	1.78 (0.23, 13.49)		
	Risk Diff (95% CI) [e]	9.7% (-20.2, 39.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.13 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.0%	9.7%
		OR (95% CI) [b]	1.78 (0.43, 7.28)	
		p-value [b]	0.4240	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6656	
		RR (95% CI) [d]	1.65 (0.47, 5.76)	
		Risk Diff (95% CI) [e]	6.3% (-8.2, 20.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.1%
	OR (95% CI) [b]	1.78 (0.42, 7.52)		
	p-value [b]	0.4344		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5553		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.13 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.61 (0.48, 5.48)	
		Risk Diff (95% CI) [e]	8.0% (-10.9, 26.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1159 von 1722

Table 14.2.18.7.1.5.13 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	3	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.3%	26.3%
		OR (95% CI) [b]	0.25 (0.05, 1.21)	
		p-value [b]	0.0860	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2530	
		RR (95% CI) [d]	0.32 (0.08, 1.18)	
		Risk Diff (95% CI) [e]	-18.0% (-39.7, 3.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.5%
	OR (95% CI) [b]	0.16 (0.03, 0.85)		
	p-value [b]	0.0315		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1963		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.13 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.25 (0.07, 0.88)	
		Risk Diff (95% CI) [e]	-33.9% (-65.8, -2.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.13 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.9%	6.7%
		OR (95% CI) [b]	0.72 (0.06, 8.55)	
		p-value [b]	0.7932	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8065	
		RR (95% CI) [d]	0.73 (0.07, 7.49)	
		Risk Diff (95% CI) [e]	-1.8% (-16.0, 12.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.1%
	OR (95% CI) [b]	0.77 (0.06, 9.46)		
	p-value [b]	0.8378		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9469		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.13 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.79 (0.08, 7.81)	
		Risk Diff (95% CI) [e]	-1.9% (-21.4, 17.5)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1163 von 1722

Table 14.2.18.7.1.5.13 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.0%	18.2%
		OR (95% CI) [b]	1.06 (0.16, 6.94)	
		p-value [b]	0.9525	
		RR (95% CI) [d]	1.05 (0.23, 4.85)	
		Risk Diff (95% CI) [e]	0.9% (-27.4, 29.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	25.0%
		OR (95% CI) [b]	0.86 (0.12, 6.01)	
	p-value [b]	0.8768		
	RR (95% CI) [d]	0.89 (0.20, 3.90)		
	Risk Diff (95% CI) [e]	-2.8% (-38.4, 32.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1164 von 1722

Table 14.2.18.7.1.6.13 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	7	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.0%	12.9%
		OR (95% CI) [b]	1.10 (0.29, 4.11)	
		p-value [b]	0.8886	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2326	
		RR (95% CI) [d]	1.09 (0.35, 3.41)	
		Risk Diff (95% CI) [e]	1.1% (-14.1, 16.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.4%
	OR (95% CI) [b]	1.07 (0.28, 4.16)		
	p-value [b]	0.9193		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1525		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.13 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.06 (0.35, 3.23)	
		Risk Diff (95% CI) [e]	1.0% (-18.8, 20.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.13 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.4%	26.3%
		OR (95% CI) [b]	0.68 (0.18, 2.51)	
		p-value [b]	0.5587	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4487	
		RR (95% CI) [d]	0.74 (0.27, 2.02)	
		Risk Diff (95% CI) [e]	-6.9% (-30.5, 16.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.9%
	OR (95% CI) [b]	0.44 (0.10, 1.92)		
	p-value [b]	0.2763		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5118		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1167 von 1722

Table 14.2.18.7.1.6.13 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.59 (0.24, 1.47)	
		Risk Diff (95% CI) [e]	-18.5% (-52.5, 15.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1168 von 1722

Table 14.2.18.7.1.6.13 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.5%	20.0%
		OR (95% CI) [b]	0.97 (0.22, 4.27)	
		p-value [b]	0.9675	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3018	
		RR (95% CI) [d]	0.98 (0.30, 3.20)	
		Risk Diff (95% CI) [e]	-0.5% (-24.1, 23.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.6%	27.3%
		OR (95% CI) [b]	1.07 (0.22, 5.07)	
		p-value [b]	0.9354	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1764	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1169 von 1722

Table 14.2.18.7.1.6.13 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.05 (0.34, 3.24)	
		Risk Diff (95% CI) [e]	1.3% (-29.9, 32.5)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1170 von 1722

Table 14.2.18.7.1.6.13 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluatable subjects at baseline	N	21	11
	Number of HRQoL evaluatable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	3	4
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	36.4%
		OR (95% CI) [b]	0.29 (0.05, 1.65)	
		p-value [b]	0.1634	
		RR (95% CI) [d]	0.39 (0.11, 1.45)	
		Risk Diff (95% CI) [e]	-22.1% (-54.2, 10.0)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.7%	50.0%
		OR (95% CI) [b]	0.20 (0.03, 1.28)	
	p-value [b]	0.0898		
	RR (95% CI) [d]	0.33 (0.10, 1.16)		
	Risk Diff (95% CI) [e]	-33.3% (-72.0, 5.4)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1171 von 1722

Table 14.2.18.7.1.7.13 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	11	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.0%	25.8%
		OR (95% CI) [b]	0.81 (0.28, 2.31)	
		p-value [b]	0.6946	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3222	
		RR (95% CI) [d]	0.85 (0.39, 1.88)	
		Risk Diff (95% CI) [e]	-3.8% (-23.0, 15.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.9%
	OR (95% CI) [b]	0.76 (0.25, 2.31)		
	p-value [b]	0.6338		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2250		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1172 von 1722

Table 14.2.18.7.1.7.13 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.83 (0.39, 1.76)	
		Risk Diff (95% CI) [e]	-5.8% (-30.1, 18.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1173 von 1722

Table 14.2.18.7.1.7.13 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	12
	Number of subjects with clinically meaningful improvement at Week 24	n	10	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	27.8%	47.4%
		OR (95% CI) [b]	0.43 (0.13, 1.36)	
		p-value [b]	0.1505	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7195	
		RR (95% CI) [d]	0.59 (0.29, 1.19)	
		Risk Diff (95% CI) [e]	-19.6% (-46.4, 7.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	38.5%	75.0%
	OR (95% CI) [b]	0.21 (0.05, 0.96)		
	p-value [b]	0.0441		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9735		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1174 von 1722

Table 14.2.18.7.1.7.13 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.51 (0.29, 0.92)	
		Risk Diff (95% CI) [e]	-36.5% (-67.4, -5.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.13 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	27	11
	Number of subjects with clinically meaningful improvement at Week 24	n	7	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.1%	20.0%
		OR (95% CI) [b]	0.82 (0.18, 3.71)	
		p-value [b]	0.8003	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3753	
		RR (95% CI) [d]	0.85 (0.25, 2.88)	
		Risk Diff (95% CI) [e]	-2.9% (-26.2, 20.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.9%
	OR (95% CI) [b]	0.93 (0.19, 4.54)		
	p-value [b]	0.9319		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2162		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1176 von 1722

Table 14.2.18.7.1.7.13 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.95 (0.30, 3.02)	
		Risk Diff (95% CI) [e]	-1.3% (-32.4, 29.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1177 von 1722

Table 14.2.18.7.1.7.13 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluatable subjects at baseline	N	21	11
	Number of HRQoL evaluatable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	3	4
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	36.4%
		OR (95% CI) [b]	0.29 (0.05, 1.65)	
		p-value [b]	0.1634	
		RR (95% CI) [d]	0.39 (0.11, 1.45)	
		Risk Diff (95% CI) [e]	-22.1% (-54.2, 10.0)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.7%	50.0%
		OR (95% CI) [b]	0.20 (0.03, 1.28)	
	p-value [b]	0.0898		
	RR (95% CI) [d]	0.33 (0.10, 1.16)		
	Risk Diff (95% CI) [e]	-33.3% (-72.0, 5.4)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1178 von 1722

Table 14.2.18.7.1.8.13 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	6	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	6.5%
		OR (95% CI) [b]	1.98 (0.37, 10.48)	
		p-value [b]	0.4230	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6826	
		RR (95% CI) [d]	1.86 (0.40, 8.64)	
		Risk Diff (95% CI) [e]	5.5% (-6.9, 18.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.8%
	OR (95% CI) [b]	1.97 (0.36, 10.69)		
	p-value [b]	0.4327		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6045		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.13 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.82 (0.40, 8.25)	
		Risk Diff (95% CI) [e]	7.1% (-9.2, 23.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1180 von 1722

Table 14.2.18.7.1.8.13 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	12
	Number of subjects with clinically meaningful improvement at Week 24	n	4	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.1%	5.3%
		OR (95% CI) [b]	2.25 (0.23, 21.69)	
		p-value [b]	0.4831	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6604	
		RR (95% CI) [d]	2.11 (0.25, 17.58)	
		Risk Diff (95% CI) [e]	5.8% (-8.5, 20.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.4%
	OR (95% CI) [b]	2.00 (0.20, 20.10)		
	p-value [b]	0.5561		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6382		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1181 von 1722

Table 14.2.18.7.1.8.13 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.85 (0.23, 14.80)	
		Risk Diff (95% CI) [e]	7.1% (-13.9, 28.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1182 von 1722

Table 14.2.18.7.1.8.13 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	4	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.8%	6.7%
		OR (95% CI) [b]	1.51 (0.16, 14.74)	
		p-value [b]	0.7212	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8340	
		RR (95% CI) [d]	1.46 (0.18, 12.07)	
		Risk Diff (95% CI) [e]	3.1% (-12.5, 18.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.3%	9.1%
		OR (95% CI) [b]	1.67 (0.17, 16.83)	
		p-value [b]	0.6650	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7143	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.13 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.57 (0.20, 12.55)	
		Risk Diff (95% CI) [e]	5.2% (-16.2, 26.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1184 von 1722

Table 14.2.18.7.1.8.13 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.5%	9.1%
		OR (95% CI) [b]	1.05 (0.08, 13.08)	
		p-value [b]	0.9682	
		RR (95% CI) [d]	1.05 (0.11, 10.31)	
		Risk Diff (95% CI) [e]	0.4% (-20.7, 21.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.1%	12.5%
		OR (95% CI) [b]	0.88 (0.07, 11.31)	
	p-value [b]	0.9186		
	RR (95% CI) [d]	0.89 (0.09, 8.44)		
	Risk Diff (95% CI) [e]	-1.4% (-28.5, 25.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1185 von 1722

Table 14.2.18.7.1.9.13 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	6	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.0%	29.0%
		OR (95% CI) [b]	0.33 (0.11, 1.06)	
		p-value [b]	0.0618	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4846	
		RR (95% CI) [d]	0.41 (0.16, 1.05)	
		Risk Diff (95% CI) [e]	-17.0% (-35.4, 1.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.8%	39.1%
		OR (95% CI) [b]	0.29 (0.09, 0.98)	
		p-value [b]	0.0458	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5488	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.13 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.40 (0.17, 0.99)	
		Risk Diff (95% CI) [e]	-23.3% (-46.4, -0.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1187 von 1722

Table 14.2.18.7.1.9.13 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	12
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.9%	10.5%
		OR (95% CI) [b]	1.37 (0.24, 7.84)	
		p-value [b]	0.7228	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6522	
		RR (95% CI) [d]	1.32 (0.28, 6.17)	
		Risk Diff (95% CI) [e]	3.4% (-14.5, 21.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.2%	16.7%
	OR (95% CI) [b]	1.19 (0.20, 7.23)		
	p-value [b]	0.8498		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6204		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1188 von 1722

Table 14.2.18.7.1.9.13 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.15 (0.26, 5.12)	
		Risk Diff (95% CI) [e]	2.6% (-23.4, 28.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1189 von 1722

Table 14.2.18.7.1.9.13 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	27	11
	Number of subjects with clinically meaningful improvement at Week 24	n	10	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	24.4%	6.7%
		OR (95% CI) [b]	4.51 (0.53, 38.74)	
		p-value [b]	0.1695	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2266	
		RR (95% CI) [d]	3.66 (0.51, 26.20)	
		Risk Diff (95% CI) [e]	17.7% (-0.5, 35.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	37.0%	9.1%
		OR (95% CI) [b]	5.88 (0.65, 52.99)	
		p-value [b]	0.1143	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1345	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1190 von 1722

Table 14.2.18.7.1.9.13 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	4.07 (0.59, 28.14)	
		Risk Diff (95% CI) [e]	27.9% (3.0, 52.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1191 von 1722

Table 14.2.18.7.1.9.13 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	18.2%
		OR (95% CI) [b]	0.75 (0.11, 5.32)	
		p-value [b]	0.7736	
		RR (95% CI) [d]	0.79 (0.15, 4.03)	
		Risk Diff (95% CI) [e]	-3.9% (-31.2, 23.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.7%	25.0%
		OR (95% CI) [b]	0.60 (0.08, 4.54)	
	p-value [b]	0.6209		
	RR (95% CI) [d]	0.67 (0.14, 3.25)		
	Risk Diff (95% CI) [e]	-8.3% (-42.9, 26.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1192 von 1722

Table 14.2.18.7.1.10.13 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	10	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	19.4%
		OR (95% CI) [b]	1.04 (0.34, 3.22)	
		p-value [b]	0.9435	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7759	
		RR (95% CI) [d]	1.03 (0.42, 2.56)	
		Risk Diff (95% CI) [e]	0.6% (-17.1, 18.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.3%	26.1%
		OR (95% CI) [b]	1.01 (0.31, 3.29)	
		p-value [b]	0.9843	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7047	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1193 von 1722

Table 14.2.18.7.1.10.13 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.01 (0.42, 2.41)	
		Risk Diff (95% CI) [e]	0.2% (-22.5, 23.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.13 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	25	12
	Number of subjects with clinically meaningful improvement at Week 24	n	2	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.6%	26.3%
		OR (95% CI) [b]	0.16 (0.03, 0.95)	
		p-value [b]	0.0439	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2586	
		RR (95% CI) [d]	0.21 (0.05, 0.99)	
		Risk Diff (95% CI) [e]	-20.8% (-41.9, 0.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	8.0%	41.7%
		OR (95% CI) [b]	0.12 (0.02, 0.77)	
		p-value [b]	0.0253	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2343	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.13 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.19 (0.04, 0.85)	
		Risk Diff (95% CI) [e]	-33.7% (-63.5, -3.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1196 von 1722

Table 14.2.18.7.1.10.13 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.5%	20.0%
		OR (95% CI) [b]	0.97 (0.22, 4.27)	
		p-value [b]	0.9675	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8376	
		RR (95% CI) [d]	0.98 (0.30, 3.20)	
		Risk Diff (95% CI) [e]	-0.5% (-24.1, 23.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.8%
	OR (95% CI) [b]	1.19 (0.25, 5.68)		
	p-value [b]	0.8317		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6400		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1197 von 1722

Table 14.2.18.7.1.10.13 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.13 (0.37, 3.47)	
		Risk Diff (95% CI) [e]	3.5% (-28.2, 35.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1198 von 1722

Table 14.2.18.7.1.10.13 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	17	8
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	18.2%
		OR (95% CI) [b]	0.75 (0.11, 5.32)	
		p-value [b]	0.7736	
		RR (95% CI) [d]	0.79 (0.15, 4.03)	
		Risk Diff (95% CI) [e]	-3.9% (-31.2, 23.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.6%	25.0%
		OR (95% CI) [b]	0.64 (0.08, 4.89)	
	p-value [b]	0.6695		
	RR (95% CI) [d]	0.71 (0.15, 3.43)		
	Risk Diff (95% CI) [e]	-7.4% (-42.4, 27.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1199 von 1722

Table 14.2.18.7.1.11.13 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	37	23
	Number of subjects with clinically meaningful improvement at Week 24	n	10	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.0%	12.9%
		OR (95% CI) [b]	1.69 (0.48, 5.94)	
		p-value [b]	0.4150	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3878	
		RR (95% CI) [d]	1.55 (0.53, 4.52)	
		Risk Diff (95% CI) [e]	7.1% (-9.1, 23.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.0%
	OR (95% CI) [b]	1.76 (0.48, 6.45)		
	p-value [b]	0.3943		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2487		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.13 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.55 (0.55, 4.38)	
		Risk Diff (95% CI) [e]	9.6% (-11.5, 30.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.13 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	12
	Number of subjects with clinically meaningful improvement at Week 24	n	5	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.9%	26.3%
		OR (95% CI) [b]	0.45 (0.11, 1.81)	
		p-value [b]	0.2627	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6801	
		RR (95% CI) [d]	0.53 (0.17, 1.60)	
		Risk Diff (95% CI) [e]	-12.4% (-35.2, 10.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.2%
	OR (95% CI) [b]	0.33 (0.07, 1.50)		
	p-value [b]	0.1528		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7262		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1202 von 1722

Table 14.2.18.7.1.11.13 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.46 (0.16, 1.30)	
		Risk Diff (95% CI) [e]	-22.4% (-54.2, 9.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.13 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	27	11
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.1%	33.3%
		OR (95% CI) [b]	0.41 (0.11, 1.58)	
		p-value [b]	0.1967	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6127	
		RR (95% CI) [d]	0.51 (0.19, 1.37)	
		Risk Diff (95% CI) [e]	-16.3% (-42.8, 10.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.9%
	OR (95% CI) [b]	0.42 (0.10, 1.82)		
	p-value [b]	0.2462		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8789		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.13 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.57 (0.23, 1.42)	
		Risk Diff (95% CI) [e]	-19.5% (-53.3, 14.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1205 von 1722

Table 14.2.18.7.1.11.13 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluatable subjects at baseline	N	21	11
	Number of HRQoL evaluatable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	28.6%	36.4%
		OR (95% CI) [b]	0.70 (0.15, 3.30)	
		p-value [b]	0.6522	
		RR (95% CI) [d]	0.79 (0.28, 2.21)	
		Risk Diff (95% CI) [e]	-7.8% (-42.2, 26.6)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	33.3%	50.0%
		OR (95% CI) [b]	0.50 (0.09, 2.73)	
	p-value [b]	0.4235		
	RR (95% CI) [d]	0.67 (0.26, 1.73)		
	Risk Diff (95% CI) [e]	-16.7% (-57.6, 24.3)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1206 von 1722

Table 14.2.18.7.1.12.13 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	8	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.0%	9.7%
		OR (95% CI) [b]	1.78 (0.43, 7.28)	
		p-value [b]	0.4240	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4836	
		RR (95% CI) [d]	1.65 (0.47, 5.76)	
		Risk Diff (95% CI) [e]	6.3% (-8.2, 20.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.1%
	OR (95% CI) [b]	1.78 (0.42, 7.52)		
	p-value [b]	0.4344		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3222		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1207 von 1722

Table 14.2.18.7.1.12.13 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.61 (0.48, 5.48)	
		Risk Diff (95% CI) [e]	8.0% (-10.9, 26.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.13 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	25	12
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.1%	15.8%
		OR (95% CI) [b]	0.67 (0.13, 3.34)	
		p-value [b]	0.6222	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9275	
		RR (95% CI) [d]	0.70 (0.18, 2.83)	
		Risk Diff (95% CI) [e]	-4.7% (-24.0, 14.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.0%	25.0%
		OR (95% CI) [b]	0.57 (0.11, 3.09)	
		p-value [b]	0.5159	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9216	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1209 von 1722

Table 14.2.18.7.1.12.13 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.64 (0.17, 2.42)	
		Risk Diff (95% CI) [e]	-9.0% (-37.4, 19.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.13 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	6	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.6%	6.7%
		OR (95% CI) [b]	2.40 (0.26, 21.78)	
		p-value [b]	0.4367	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4398	
		RR (95% CI) [d]	2.20 (0.29, 16.76)	
		Risk Diff (95% CI) [e]	8.0% (-8.7, 24.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.4%	9.1%
		OR (95% CI) [b]	2.73 (0.29, 25.75)	
		p-value [b]	0.3811	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2747	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1211 von 1722

Table 14.2.18.7.1.12.13 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.36 (0.32, 17.40)	
		Risk Diff (95% CI) [e]	12.3% (-10.5, 35.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1212 von 1722

Table 14.2.18.7.1.12.13 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	7
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.3%	18.2%
		OR (95% CI) [b]	0.75 (0.11, 5.32)	
		p-value [b]	0.7736	
		RR (95% CI) [d]	0.79 (0.15, 4.03)	
		Risk Diff (95% CI) [e]	-3.9% (-31.2, 23.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.7%	28.6%
		OR (95% CI) [b]	0.50 (0.06, 3.91)	
	p-value [b]	0.5087		
	RR (95% CI) [d]	0.58 (0.12, 2.78)		
	Risk Diff (95% CI) [e]	-11.9% (-49.5, 25.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1213 von 1722

Table 14.2.18.7.1.13.13 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.0%	16.1%
		OR (95% CI) [b]	0.85 (0.24, 2.94)	
		p-value [b]	0.7933	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9490	
		RR (95% CI) [d]	0.87 (0.30, 2.50)	
		Risk Diff (95% CI) [e]	-2.1% (-18.3, 14.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.4%
	OR (95% CI) [b]	0.81 (0.22, 2.94)		
	p-value [b]	0.7523		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9506		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.13 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.85 (0.30, 2.36)	
		Risk Diff (95% CI) [e]	-3.3% (-24.2, 17.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1215 von 1722

Table 14.2.18.7.1.13.13 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	10	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	27.8%	10.5%
		OR (95% CI) [b]	3.27 (0.64, 16.80)	
		p-value [b]	0.1560	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9550	
		RR (95% CI) [d]	2.64 (0.64, 10.84)	
		Risk Diff (95% CI) [e]	17.3% (-2.9, 37.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	38.5%
	OR (95% CI) [b]	2.81 (0.50, 15.77)		
	p-value [b]	0.2397		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9560		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.13 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.12 (0.55, 8.12)	
		Risk Diff (95% CI) [e]	20.3% (-9.2, 49.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.13 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.0%	40.0%
		OR (95% CI) [b]	0.42 (0.12, 1.50)	
		p-value [b]	0.1830	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9459	
		RR (95% CI) [d]	0.55 (0.24, 1.28)	
		Risk Diff (95% CI) [e]	-18.0% (-45.9, 9.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	32.1%
	OR (95% CI) [b]	0.39 (0.09, 1.65)		
	p-value [b]	0.2018		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9475		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1218 von 1722

Table 14.2.18.7.1.13.13 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.59 (0.28, 1.26)	
		Risk Diff (95% CI) [e]	-22.4% (-56.5, 11.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.13 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	4	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9497	
		RR (95% CI) [d]	4.91 (0.29, 83.67)	
		Risk Diff (95% CI) [e]	16.3% (-4.0, 36.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9515		
	RR (95% CI) [d]	4.26 (0.26, 70.96)		
	Risk Diff (95% CI) [e]	18.1% (-6.1, 42.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1220 von 1722

Table 14.2.18.7.1.14.13 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects at baseline	N	50	31
	Number of HRQoL evaluable subjects at Week 24	N	38	23
	Number of subjects with clinically meaningful improvement at Week 24	n	4	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	3.2%
		OR (95% CI) [b]	2.61 (0.28, 24.48)	
		p-value [b]	0.4013	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9584	
		RR (95% CI) [d]	2.48 (0.29, 21.18)	
		Risk Diff (95% CI) [e]	4.8% (-5.0, 14.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.5%
	OR (95% CI) [b]	2.59 (0.27, 24.70)		
	p-value [b]	0.4087		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9594		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1221 von 1722

Table 14.2.18.7.1.14.13 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.42 (0.29, 20.36)	
		Risk Diff (95% CI) [e]	6.2% (-6.7, 19.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1222 von 1722

Table 14.2.18.7.1.14.13 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.6%	15.8%
		OR (95% CI) [b]	0.31 (0.05, 2.07)	
		p-value [b]	0.2281	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9483	
		RR (95% CI) [d]	0.35 (0.06, 1.93)	
		Risk Diff (95% CI) [e]	-10.2% (-28.3, 7.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.7%	27.3%
	OR (95% CI) [b]	0.22 (0.03, 1.58)		
	p-value [b]	0.1326		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9478		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1223 von 1722

Table 14.2.18.7.1.14.13 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.28 (0.05, 1.46)	
		Risk Diff (95% CI) [e]	-19.6% (-47.8, 8.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1224 von 1722

Table 14.2.18.7.1.14.13 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.8%	13.3%
		OR (95% CI) [b]	0.70 (0.11, 4.30)	
		p-value [b]	0.7026	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9521	
		RR (95% CI) [d]	0.73 (0.15, 3.59)	
		Risk Diff (95% CI) [e]	-3.6% (-23.0, 15.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.3%
	OR (95% CI) [b]	0.75 (0.12, 4.83)		
	p-value [b]	0.7621		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9535		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.13 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.79 (0.17, 3.69)	
		Risk Diff (95% CI) [e]	-3.9% (-30.1, 22.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1226 von 1722

Table 14.2.18.7.1.14.13 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects at baseline	N	21	11
	Number of HRQoL evaluable subjects at Week 24	N	18	8
	Number of subjects with clinically meaningful improvement at Week 24	n	1	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.8%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9538	
		RR (95% CI) [d]	1.64 (0.07, 37.15)	
		Risk Diff (95% CI) [e]	2.7% (-12.8, 18.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	5.6%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9549		
	RR (95% CI) [d]	1.42 (0.06, 31.57)		
	Risk Diff (95% CI) [e]	2.3% (-16.9, 21.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1227 von 1722

Table 14.2.18.7.1.15.13 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluatable subjects at baseline	N	50	31
	Number of HRQoL evaluatable subjects at Week 24	N	37	23
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	6.5%
		OR (95% CI) [b]	1.26 (0.22, 7.33)	
		p-value [b]	0.7963	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9084	
		RR (95% CI) [d]	1.24 (0.24, 6.37)	
		Risk Diff (95% CI) [e]	1.5% (-9.9, 13.0)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.8%	8.7%
		OR (95% CI) [b]	1.27 (0.21, 7.57)	
		p-value [b]	0.7910	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8456	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1228 von 1722

Table 14.2.18.7.1.15.13 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.24 (0.25, 6.25)	
		Risk Diff (95% CI) [e]	2.1% (-13.1, 17.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.13 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects at baseline	N	36	19
	Number of HRQoL evaluable subjects at Week 24	N	26	11
	Number of subjects with clinically meaningful improvement at Week 24	n	4	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.1%	26.3%
		OR (95% CI) [b]	0.35 (0.08, 1.50)	
		p-value [b]	0.1579	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4584	
		RR (95% CI) [d]	0.42 (0.13, 1.39)	
		Risk Diff (95% CI) [e]	-15.2% (-37.5, 7.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.4%
	OR (95% CI) [b]	0.22 (0.04, 1.08)		
	p-value [b]	0.0613		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3453		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.13 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.34 (0.11, 1.03)	
		Risk Diff (95% CI) [e]	-30.1% (-62.6, 2.5)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.13 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects at baseline	N	41	15
	Number of HRQoL evaluable subjects at Week 24	N	28	11
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	4.9%	6.7%
		OR (95% CI) [b]	0.72 (0.06, 8.55)	
		p-value [b]	0.7932	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8319	
		RR (95% CI) [d]	0.73 (0.07, 7.49)	
		Risk Diff (95% CI) [e]	-1.8% (-16.0, 12.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.1%
	OR (95% CI) [b]	0.77 (0.06, 9.45)		
	p-value [b]	0.8376		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9158		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.13 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.79 (0.08, 7.81)	
		Risk Diff (95% CI) [e]	-1.9% (-21.4, 17.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1233 von 1722

Table 14.2.18.7.1.15.13 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluatable subjects at baseline	N	21	11
	Number of HRQoL evaluatable subjects at Week 24	N	17	8
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 [a]	%	9.5%	9.1%
		OR (95% CI) [b]	1.05 (0.08, 13.08)	
		p-value [b]	0.9682	
		RR (95% CI) [d]	1.05 (0.11, 10.31)	
		Risk Diff (95% CI) [e]	0.4% (-20.7, 21.6)	
	Percent of HRQoL evaluatable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.8%	12.5%
		OR (95% CI) [b]	0.93 (0.07, 12.11)	
	p-value [b]	0.9579		
	RR (95% CI) [d]	0.94 (0.10, 8.92)		
	Risk Diff (95% CI) [e]	-0.7% (-28.3, 26.8)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1234 von 1722

Table 14.2.18.7.1.1.14 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	46
	Number of subjects with clinically meaningful improvement at Week 24	n	29	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.0%	15.4%
		OR (95% CI) [b]	1.46 (0.67, 3.22)	
		p-value [b]	0.3442	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6734	
		RR (95% CI) [d]	1.37 (0.71, 2.63)	
		Risk Diff (95% CI) [e]	5.6% (-5.5, 16.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.4%	21.7%
	OR (95% CI) [b]	1.43 (0.63, 3.25)		
	p-value [b]	0.3938		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4639		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.14 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.31 (0.70, 2.45)	
		Risk Diff (95% CI) [e]	6.7% (-8.1, 21.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.14 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	2	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.2%	20.0%
		OR (95% CI) [b]	0.89 (0.10, 7.86)	
		p-value [b]	0.9156	
		RR (95% CI) [d]	0.91 (0.16, 5.30)	
		Risk Diff (95% CI) [e]	-1.8% (-35.5, 31.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	33.3%
		OR (95% CI) [b]	0.57 (0.06, 5.77)	
	p-value [b]	0.6353		
	RR (95% CI) [d]	0.67 (0.13, 3.53)		
	Risk Diff (95% CI) [e]	-11.1% (-57.6, 35.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1237 von 1722

Table 14.2.18.7.1.2.14 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	47
	Number of subjects with clinically meaningful improvement at Week 24	n	23	17
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.7%	26.2%
		OR (95% CI) [b]	0.56 (0.28, 1.15)	
		p-value [b]	0.1155	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3530	
		RR (95% CI) [d]	0.64 (0.37, 1.11)	
		Risk Diff (95% CI) [e]	-9.5% (-21.8, 2.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.5%	36.2%
	OR (95% CI) [b]	0.51 (0.24, 1.09)		
	p-value [b]	0.0837		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4686		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.14 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.62 (0.37, 1.05)	
		Risk Diff (95% CI) [e]	-13.6% (-29.6, 2.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.14 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.2%	10.0%
		OR (95% CI) [b]	2.00 (0.15, 26.19)	
		p-value [b]	0.5974	
		RR (95% CI) [d]	1.82 (0.19, 17.12)	
		Risk Diff (95% CI) [e]	8.2% (-21.2, 37.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	16.7%
		OR (95% CI) [b]	1.43 (0.10, 20.44)	
	p-value [b]	0.7928		
	RR (95% CI) [d]	1.33 (0.15, 11.64)		
	Risk Diff (95% CI) [e]	5.6% (-34.8, 45.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1240 von 1722

Table 14.2.18.7.1.3.14 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	47
	Number of subjects with clinically meaningful improvement at Week 24	n	28	18
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.3%	27.7%
		OR (95% CI) [b]	0.66 (0.34, 1.32)	
		p-value [b]	0.2415	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9487	
		RR (95% CI) [d]	0.73 (0.44, 1.22)	
		Risk Diff (95% CI) [e]	-7.4% (-20.2, 5.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.5%	38.3%
	OR (95% CI) [b]	0.61 (0.29, 1.27)		
	p-value [b]	0.1848		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9548		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1241 von 1722

Table 14.2.18.7.1.3.14 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.44, 1.16)	
		Risk Diff (95% CI) [e]	-10.8% (-27.2, 5.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1242 von 1722

Table 14.2.18.7.1.3.14 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	2	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.2%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9505	
		RR (95% CI) [d]	4.58 (0.25, 85.33)	
		Risk Diff (95% CI) [e]	16.3% (-9.8, 42.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9567		
	RR (95% CI) [d]	3.50 (0.20, 62.27)		
	Risk Diff (95% CI) [e]	17.9% (-15.1, 50.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.14 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	46
	Number of subjects with clinically meaningful improvement at Week 24	n	29	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.0%	18.5%
		OR (95% CI) [b]	1.18 (0.56, 2.48)	
		p-value [b]	0.6727	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9487	
		RR (95% CI) [d]	1.14 (0.62, 2.08)	
		Risk Diff (95% CI) [e]	2.6% (-9.1, 14.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.4%	26.1%
	OR (95% CI) [b]	1.13 (0.51, 2.47)		
	p-value [b]	0.7681		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9455		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1244 von 1722

Table 14.2.18.7.1.4.14 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.09 (0.61, 1.94)	
		Risk Diff (95% CI) [e]	2.3% (-13.1, 17.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.14 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	0	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	20.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9495	
		RR (95% CI) [d]	0.18 (0.01, 3.41)	
		Risk Diff (95% CI) [e]	-18.6% (-45.8, 8.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	33.3%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9460		
	RR (95% CI) [d]	0.14 (0.01, 2.49)		
	Risk Diff (95% CI) [e]	-30.7% (-68.7, 7.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.14 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	46
	Number of subjects with clinically meaningful improvement at Week 24	n	15	10
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.9%	15.4%
		OR (95% CI) [b]	0.67 (0.28, 1.59)	
		p-value [b]	0.3633	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2235	
		RR (95% CI) [d]	0.71 (0.34, 1.49)	
		Risk Diff (95% CI) [e]	-4.5% (-14.7, 5.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.7%
	OR (95% CI) [b]	0.62 (0.26, 1.51)		
	p-value [b]	0.2933		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3129		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.14 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.68 (0.33, 1.39)	
		Risk Diff (95% CI) [e]	-7.0% (-20.8, 6.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.14 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	3	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	27.3%	10.0%
		OR (95% CI) [b]	3.38 (0.29, 39.33)	
		p-value [b]	0.3315	
		RR (95% CI) [d]	2.73 (0.34, 22.16)	
		Risk Diff (95% CI) [e]	17.3% (-15.0, 49.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	33.3%	16.7%
		OR (95% CI) [b]	2.50 (0.19, 32.19)	
	p-value [b]	0.4822		
	RR (95% CI) [d]	2.00 (0.27, 14.98)		
	Risk Diff (95% CI) [e]	16.7% (-26.2, 59.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1249 von 1722

Table 14.2.18.7.1.6.14 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	46
	Number of subjects with clinically meaningful improvement at Week 24	n	26	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.8%	21.5%
		OR (95% CI) [b]	0.85 (0.41, 1.75)	
		p-value [b]	0.6523	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9506	
		RR (95% CI) [d]	0.87 (0.49, 1.56)	
		Risk Diff (95% CI) [e]	-2.7% (-14.6, 9.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	25.5%	30.4%
	OR (95% CI) [b]	0.78 (0.36, 1.69)		
	p-value [b]	0.5312		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9476		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1250 von 1722

Table 14.2.18.7.1.6.14 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.84 (0.48, 1.45)	
		Risk Diff (95% CI) [e]	-4.9% (-20.7, 10.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1251 von 1722

Table 14.2.18.7.1.6.14 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	0	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	20.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9499	
		RR (95% CI) [d]	0.18 (0.01, 3.41)	
		Risk Diff (95% CI) [e]	-18.6% (-45.8, 8.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	33.3%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9464		
	RR (95% CI) [d]	0.14 (0.01, 2.49)		
	Risk Diff (95% CI) [e]	-30.7% (-68.7, 7.3)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1252 von 1722

Table 14.2.18.7.1.7.14 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	101	47
	Number of subjects with clinically meaningful improvement at Week 24	n	31	23
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.5%	35.4%
		OR (95% CI) [b]	0.53 (0.28, 1.01)	
		p-value [b]	0.0537	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7270	
		RR (95% CI) [d]	0.63 (0.40, 1.00)	
		Risk Diff (95% CI) [e]	-12.9% (-26.5, 0.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.7%	48.9%
	OR (95% CI) [b]	0.46 (0.23, 0.94)		
	p-value [b]	0.0334		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8467		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.14 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.63 (0.41, 0.95)	
		Risk Diff (95% CI) [e]	-18.2% (-35.1, -1.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.14 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	1	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.1%	10.0%
		OR (95% CI) [b]	0.90 (0.05, 16.59)	
		p-value [b]	0.9437	
		RR (95% CI) [d]	0.91 (0.07, 12.69)	
		Risk Diff (95% CI) [e]	-0.9% (-26.1, 24.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.1%	16.7%
		OR (95% CI) [b]	0.63 (0.03, 12.42)	
	p-value [b]	0.7584		
	RR (95% CI) [d]	0.67 (0.05, 8.73)		
	Risk Diff (95% CI) [e]	-5.6% (-41.8, 30.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1255 von 1722

Table 14.2.18.7.1.8.14 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	47
	Number of subjects with clinically meaningful improvement at Week 24	n	16	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.6%	7.7%
		OR (95% CI) [b]	1.57 (0.55, 4.50)	
		p-value [b]	0.3976	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9503	
		RR (95% CI) [d]	1.51 (0.58, 3.94)	
		Risk Diff (95% CI) [e]	3.9% (-4.5, 12.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.7%	10.6%
		OR (95% CI) [b]	1.56 (0.54, 4.56)	
		p-value [b]	0.4134	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9558		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1256 von 1722

Table 14.2.18.7.1.8.14 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.47 (0.57, 3.79)	
		Risk Diff (95% CI) [e]	5.0% (-6.2, 16.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1257 von 1722

Table 14.2.18.7.1.8.14 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	1	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.1%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9480	
		RR (95% CI) [d]	2.75 (0.12, 60.70)	
		Risk Diff (95% CI) [e]	8.0% (-14.4, 30.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.1%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9538		
	RR (95% CI) [d]	2.10 (0.10, 44.40)		
	Risk Diff (95% CI) [e]	7.9% (-21.4, 37.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1258 von 1722

Table 14.2.18.7.1.9.14 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	101	47
	Number of subjects with clinically meaningful improvement at Week 24	n	23	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.7%	18.5%
		OR (95% CI) [b]	0.88 (0.41, 1.91)	
		p-value [b]	0.7522	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9958	
		RR (95% CI) [d]	0.90 (0.48, 1.70)	
		Risk Diff (95% CI) [e]	-1.8% (-13.1, 9.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.8%	25.5%
	OR (95% CI) [b]	0.86 (0.38, 1.92)		
	p-value [b]	0.7132		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7432		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1259 von 1722

Table 14.2.18.7.1.9.14 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.89 (0.49, 1.64)	
		Risk Diff (95% CI) [e]	-2.8% (-17.7, 12.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.14 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	2	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.2%	20.0%
		OR (95% CI) [b]	0.89 (0.10, 7.86)	
		p-value [b]	0.9156	
		RR (95% CI) [d]	0.91 (0.16, 5.30)	
		Risk Diff (95% CI) [e]	-1.8% (-35.5, 31.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	33.3%
		OR (95% CI) [b]	0.57 (0.06, 5.77)	
	p-value [b]	0.6351		
	RR (95% CI) [d]	0.67 (0.13, 3.53)		
	Risk Diff (95% CI) [e]	-11.1% (-57.6, 35.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1261 von 1722

Table 14.2.18.7.1.10.14 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	98	47
	Number of subjects with clinically meaningful improvement at Week 24	n	23	14
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.7%	21.5%
		OR (95% CI) [b]	0.73 (0.35, 1.53)	
		p-value [b]	0.4027	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6609	
		RR (95% CI) [d]	0.77 (0.43, 1.40)	
		Risk Diff (95% CI) [e]	-4.9% (-16.6, 6.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.5%	29.8%
	OR (95% CI) [b]	0.72 (0.33, 1.58)		
	p-value [b]	0.4150		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4566		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.14 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.79 (0.45, 1.39)	
		Risk Diff (95% CI) [e]	-6.3% (-21.9, 9.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1263 von 1722

Table 14.2.18.7.1.10.14 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	1	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.1%	20.0%
		OR (95% CI) [b]	0.40 (0.03, 5.25)	
		p-value [b]	0.4854	
		RR (95% CI) [d]	0.45 (0.05, 4.28)	
		Risk Diff (95% CI) [e]	-10.9% (-41.0, 19.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.1%	33.3%
		OR (95% CI) [b]	0.25 (0.02, 3.66)	
	p-value [b]	0.3114		
	RR (95% CI) [d]	0.33 (0.04, 2.91)		
	Risk Diff (95% CI) [e]	-22.2% (-65.2, 20.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1264 von 1722

Table 14.2.18.7.1.11.14 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	100	47
	Number of subjects with clinically meaningful improvement at Week 24	n	26	17
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.8%	26.2%
		OR (95% CI) [b]	0.66 (0.33, 1.32)	
		p-value [b]	0.2360	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4120	
		RR (95% CI) [d]	0.72 (0.42, 1.23)	
		Risk Diff (95% CI) [e]	-7.3% (-19.8, 5.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.0%	36.2%
	OR (95% CI) [b]	0.62 (0.29, 1.30)		
	p-value [b]	0.2080		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5538		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1265 von 1722

Table 14.2.18.7.1.11.14 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.43, 1.19)	
		Risk Diff (95% CI) [e]	-10.2% (-26.4, 6.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1266 von 1722

Table 14.2.18.7.1.11.14 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.2%	10.0%
		OR (95% CI) [b]	2.00 (0.15, 26.18)	
		p-value [b]	0.5974	
		RR (95% CI) [d]	1.82 (0.19, 17.12)	
		Risk Diff (95% CI) [e]	8.2% (-21.2, 37.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	16.7%
		OR (95% CI) [b]	1.43 (0.10, 20.44)	
	p-value [b]	0.7928		
	RR (95% CI) [d]	1.33 (0.15, 11.64)		
	Risk Diff (95% CI) [e]	5.6% (-34.8, 45.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1267 von 1722

Table 14.2.18.7.1.12.14 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	46
	Number of subjects with clinically meaningful improvement at Week 24	n	18	9
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	13.0%	13.8%
		OR (95% CI) [b]	0.93 (0.39, 2.21)	
		p-value [b]	0.8752	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9539	
		RR (95% CI) [d]	0.94 (0.45, 1.98)	
		Risk Diff (95% CI) [e]	-0.8% (-10.9, 9.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.6%	19.6%
	OR (95% CI) [b]	0.88 (0.36, 2.14)		
	p-value [b]	0.7798		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9358		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1268 von 1722

Table 14.2.18.7.1.12.14 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.90 (0.44, 1.85)	
		Risk Diff (95% CI) [e]	-1.9% (-15.6, 11.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.14 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	8	6
	Number of subjects with clinically meaningful improvement at Week 24	n	4	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	36.4%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9541	
		RR (95% CI) [d]	8.25 (0.50, 136.33)	
		Risk Diff (95% CI) [e]	33.0% (2.9, 63.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	50.0%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9365		
	RR (95% CI) [d]	7.00 (0.45, 109.47)		
	Risk Diff (95% CI) [e]	42.9% (5.0, 80.7)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1270 von 1722

Table 14.2.18.7.1.13.14 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	46
	Number of subjects with clinically meaningful improvement at Week 24	n	28	12
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.3%	18.5%
		OR (95% CI) [b]	1.12 (0.53, 2.38)	
		p-value [b]	0.7600	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4015	
		RR (95% CI) [d]	1.10 (0.60, 2.02)	
		Risk Diff (95% CI) [e]	1.8% (-9.7, 13.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	27.5%	26.1%
	OR (95% CI) [b]	1.07 (0.49, 2.36)		
	p-value [b]	0.8627		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5349		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1271 von 1722

Table 14.2.18.7.1.13.14 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.05 (0.59, 1.88)	
		Risk Diff (95% CI) [e]	1.4% (-14.0, 16.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1272 von 1722

Table 14.2.18.7.1.13.14 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	3	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	27.3%	10.0%
		OR (95% CI) [b]	3.37 (0.29, 39.32)	
		p-value [b]	0.3316	
		RR (95% CI) [d]	2.73 (0.34, 22.16)	
		Risk Diff (95% CI) [e]	17.3% (-15.0, 49.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	33.3%	16.7%
		OR (95% CI) [b]	2.50 (0.19, 32.19)	
	p-value [b]	0.4822		
	RR (95% CI) [d]	2.00 (0.27, 14.98)		
	Risk Diff (95% CI) [e]	16.7% (-26.2, 59.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1273 von 1722

Table 14.2.18.7.1.14.14 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	102	46
	Number of subjects with clinically meaningful improvement at Week 24	n	11	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	7.7%
		OR (95% CI) [b]	1.04 (0.35, 3.12)	
		p-value [b]	0.9452	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9498	
		RR (95% CI) [d]	1.04 (0.38, 2.86)	
		Risk Diff (95% CI) [e]	0.3% (-7.6, 8.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.8%	10.9%
	OR (95% CI) [b]	0.99 (0.32, 3.04)		
	p-value [b]	0.9877		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9461		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1274 von 1722

Table 14.2.18.7.1.14.14 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.99 (0.37, 2.69)	
		Risk Diff (95% CI) [e]	-0.1% (-10.9, 10.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1275 von 1722

Table 14.2.18.7.1.14.14 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	0	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	0.0%	10.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9500	
		RR (95% CI) [d]	0.31 (0.01, 6.74)	
		Risk Diff (95% CI) [e]	-9.5% (-32.7, 13.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	0.0%	16.7%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9460		
	RR (95% CI) [d]	0.23 (0.01, 4.93)		
	Risk Diff (95% CI) [e]	-16.4% (-49.7, 16.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.14 (AMNOG)  
EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects at baseline	N	138	65
	Number of HRQoL evaluable subjects at Week 24	N	100	46
	Number of subjects with clinically meaningful improvement at Week 24	n	11	8
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.0%	12.3%
		OR (95% CI) [b]	0.62 (0.24, 1.62)	
		p-value [b]	0.3258	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4014	
		RR (95% CI) [d]	0.65 (0.27, 1.53)	
		Risk Diff (95% CI) [e]	-4.3% (-13.5, 4.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	11.0%	17.4%
		OR (95% CI) [b]	0.59 (0.22, 1.57)	
		p-value [b]	0.2901	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5391	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1277 von 1722

Table 14.2.18.7.1.15.14 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.63 (0.27, 1.47)	
		Risk Diff (95% CI) [e]	-6.4% (-18.9, 6.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1278 von 1722

Table 14.2.18.7.1.15.14 (AMNOG)  
EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects at baseline	N	11	10
	Number of HRQoL evaluable subjects at Week 24	N	9	6
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.2%	10.0%
		OR (95% CI) [b]	2.00 (0.15, 26.19)	
		p-value [b]	0.5973	
		RR (95% CI) [d]	1.82 (0.19, 17.12)	
		Risk Diff (95% CI) [e]	8.2% (-21.2, 37.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.2%	16.7%
		OR (95% CI) [b]	1.43 (0.10, 20.44)	
	p-value [b]	0.7927		
	RR (95% CI) [d]	1.33 (0.15, 11.64)		
	Risk Diff (95% CI) [e]	5.6% (-34.8, 45.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1279 von 1722

Table 14.2.18.7.1.1.15 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	13
	Number of subjects with clinically meaningful improvement at Week 24	n	10	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.3%	25.0%
		OR (95% CI) [b]	0.81 (0.24, 2.77)	
		p-value [b]	0.7382	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2133	
		RR (95% CI) [d]	0.85 (0.33, 2.17)	
		Risk Diff (95% CI) [e]	-3.7% (-26.0, 18.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.3%
	OR (95% CI) [b]	0.57 (0.15, 2.16)		
	p-value [b]	0.4097		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1092		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.15 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.68 (0.29, 1.63)	
		Risk Diff (95% CI) [e]	-12.1% (-42.1, 17.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.15 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	3	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.8%	44.4%
		OR (95% CI) [b]	0.12 (0.02, 0.71)	
		p-value [b]	0.0193	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0077	
		RR (95% CI) [d]	0.20 (0.05, 0.73)	
		Risk Diff (95% CI) [e]	-35.6% (-69.5, -1.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	12.0%
	OR (95% CI) [b]	0.17 (0.03, 1.02)		
	p-value [b]	0.0520		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0170		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.15 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.27 (0.07, 0.98)	
		Risk Diff (95% CI) [e]	-32.4% (-67.3, 2.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1283 von 1722

Table 14.2.18.7.1.1.15 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	7	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.9%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9531	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9568	
		RR (95% CI) [d]	10.00 (0.60, 166.34)	
		Risk Diff (95% CI) [e]	20.5% (4.9, 36.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.4%
	OR (95% CI) [b]	NE (NE, NE)		
	p-value [b]	0.9544		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9584		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1284 von 1722

Table 14.2.18.7.1.1.15 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	10.00 (0.61, 163.13)	
		Risk Diff (95% CI) [e]	28.1% (7.7, 48.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.1.15 (AMNOG)  
 EORTC QLQ-C30 Global QoL - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	10	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	32.3%	15.0%
		OR (95% CI) [b]	2.70 (0.64, 11.39)	
		p-value [b]	0.1767	
		RR (95% CI) [d]	2.15 (0.67, 6.87)	
		Risk Diff (95% CI) [e]	17.3% (-5.5, 40.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	47.6%	23.1%
		OR (95% CI) [b]	3.03 (0.64, 14.26)	
	p-value [b]	0.1606		
	RR (95% CI) [d]	2.06 (0.69, 6.13)		
	Risk Diff (95% CI) [e]	24.5% (-6.8, 55.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1286 von 1722

Table 14.2.18.7.1.2.15 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	14
	Number of subjects with clinically meaningful improvement at Week 24	n	10	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.3%	35.0%
		OR (95% CI) [b]	0.50 (0.16, 1.59)	
		p-value [b]	0.2418	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8645	
		RR (95% CI) [d]	0.61 (0.27, 1.37)	
		Risk Diff (95% CI) [e]	-13.7% (-37.7, 10.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	26.3%
	OR (95% CI) [b]	0.36 (0.10, 1.27)		
	p-value [b]	0.1127		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7063		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1287 von 1722

Table 14.2.18.7.1.2.15 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.53 (0.25, 1.11)	
		Risk Diff (95% CI) [e]	-23.7% (-53.4, 6.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1288 von 1722

Table 14.2.18.7.1.2.15 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	6	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.6%	55.6%
		OR (95% CI) [b]	0.17 (0.04, 0.83)	
		p-value [b]	0.0290	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2677	
		RR (95% CI) [d]	0.32 (0.13, 0.81)	
		Risk Diff (95% CI) [e]	-37.9% (-72.8, -3.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.0%
	OR (95% CI) [b]	0.25 (0.05, 1.26)		
	p-value [b]	0.0926		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5226		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.15 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.43 (0.17, 1.07)	
		Risk Diff (95% CI) [e]	-31.6% (-68.1, 5.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.15 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.6%	9.5%
		OR (95% CI) [b]	1.76 (0.31, 10.04)	
		p-value [b]	0.5250	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3559	
		RR (95% CI) [d]	1.64 (0.35, 7.69)	
		Risk Diff (95% CI) [e]	6.1% (-11.7, 23.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.7%
	OR (95% CI) [b]	1.81 (0.30, 10.79)		
	p-value [b]	0.5173		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3169		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.2.15 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.63 (0.36, 7.35)	
		Risk Diff (95% CI) [e]	8.4% (-15.7, 32.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1292 von 1722

Table 14.2.18.7.1.2.15 (AMNOG)  
 EORTC QLQ-C30 Physical Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	4	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.9%	20.0%
		OR (95% CI) [b]	0.59 (0.13, 2.70)	
		p-value [b]	0.4992	
		RR (95% CI) [d]	0.65 (0.18, 2.29)	
		Risk Diff (95% CI) [e]	-7.1% (-28.2, 14.0)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.0%	30.8%
		OR (95% CI) [b]	0.53 (0.11, 2.63)	
	p-value [b]	0.4371		
	RR (95% CI) [d]	0.62 (0.19, 2.06)		
	Risk Diff (95% CI) [e]	-11.7% (-41.9, 18.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1293 von 1722

Table 14.2.18.7.1.3.15 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	14
	Number of subjects with clinically meaningful improvement at Week 24	n	9	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	25.0%
		OR (95% CI) [b]	0.71 (0.20, 2.47)	
		p-value [b]	0.5908	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8219	
		RR (95% CI) [d]	0.77 (0.29, 2.00)	
		Risk Diff (95% CI) [e]	-5.9% (-27.9, 16.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.7%
	OR (95% CI) [b]	0.56 (0.15, 2.10)		
	p-value [b]	0.3889		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7189		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1294 von 1722

Table 14.2.18.7.1.3.15 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.66 (0.27, 1.64)	
		Risk Diff (95% CI) [e]	-12.0% (-40.5, 16.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1295 von 1722

Table 14.2.18.7.1.3.15 (AMNOG)  
EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.7%	22.2%
		OR (95% CI) [b]	0.60 (0.10, 3.78)	
		p-value [b]	0.5897	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7472	
		RR (95% CI) [d]	0.66 (0.15, 2.87)	
		Risk Diff (95% CI) [e]	-7.5% (-37.2, 22.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%
	OR (95% CI) [b]	0.88 (0.14, 5.58)		
	p-value [b]	0.8876		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9403		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.15 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.90 (0.21, 3.84)	
		Risk Diff (95% CI) [e]	-2.2% (-33.6, 29.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.3.15 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	8	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	28.6%
		OR (95% CI) [b]	0.83 (0.24, 2.88)	
		p-value [b]	0.7731	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9578	
		RR (95% CI) [d]	0.88 (0.35, 2.16)	
		Risk Diff (95% CI) [e]	-3.6% (-28.0, 20.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.8%
	OR (95% CI) [b]	0.80 (0.21, 3.06)		
	p-value [b]	0.7447		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	1.0000		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1298 von 1722

Table 14.2.18.7.1.3.15 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.87 (0.38, 2.00)	
		Risk Diff (95% CI) [e]	-5.2% (-36.7, 26.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1299 von 1722

Table 14.2.18.7.1.3.15 (AMNOG)  
 EORTC QLQ-C30 Role Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.6%	25.0%
		OR (95% CI) [b]	0.87 (0.23, 3.26)	
		p-value [b]	0.8424	
		RR (95% CI) [d]	0.90 (0.33, 2.46)	
		Risk Diff (95% CI) [e]	-2.4% (-26.4, 21.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	33.3%	38.5%
		OR (95% CI) [b]	0.80 (0.19, 3.37)	
	p-value [b]	0.7612		
	RR (95% CI) [d]	0.87 (0.35, 2.17)		
	Risk Diff (95% CI) [e]	-5.1% (-38.4, 28.1)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1300 von 1722

Table 14.2.18.7.1.4.15 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	13
	Number of subjects with clinically meaningful improvement at Week 24	n	7	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.9%	25.0%
		OR (95% CI) [b]	0.53 (0.14, 1.91)	
		p-value [b]	0.3283	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2541	
		RR (95% CI) [d]	0.60 (0.21, 1.65)	
		Risk Diff (95% CI) [e]	-10.1% (-31.6, 11.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	18.4%
	OR (95% CI) [b]	0.36 (0.09, 1.44)		
	p-value [b]	0.1500		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1535		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1301 von 1722

Table 14.2.18.7.1.4.15 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.48 (0.18, 1.25)	
		Risk Diff (95% CI) [e]	-20.0% (-49.2, 9.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.15 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	7	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.6%	22.2%
		OR (95% CI) [b]	0.91 (0.15, 5.37)	
		p-value [b]	0.9147	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6122	
		RR (95% CI) [d]	0.93 (0.23, 3.72)	
		Risk Diff (95% CI) [e]	-1.6% (-32.0, 28.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.0%
	OR (95% CI) [b]	1.36 (0.23, 8.22)		
	p-value [b]	0.7368		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8682		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1303 von 1722

Table 14.2.18.7.1.4.15 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.26 (0.32, 4.98)	
		Risk Diff (95% CI) [e]	5.8% (-26.6, 38.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1304 von 1722

Table 14.2.18.7.1.4.15 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	8	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	19.0%
		OR (95% CI) [b]	1.42 (0.37, 5.47)	
		p-value [b]	0.6135	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8805	
		RR (95% CI) [d]	1.31 (0.45, 3.81)	
		Risk Diff (95% CI) [e]	6.0% (-16.6, 28.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.8%
	OR (95% CI) [b]	1.47 (0.35, 6.13)		
	p-value [b]	0.5997		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9063		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1305 von 1722

Table 14.2.18.7.1.4.15 (AMNOG)  
 EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.30 (0.48, 3.58)	
		Risk Diff (95% CI) [e]	8.1% (-21.5, 37.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.4.15 (AMNOG)  
EORTC QLQ-C30 Cognitive Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	7	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.6%	15.0%
		OR (95% CI) [b]	1.65 (0.37, 7.32)	
		p-value [b]	0.5082	
		RR (95% CI) [d]	1.51 (0.44, 5.15)	
		Risk Diff (95% CI) [e]	7.6% (-13.9, 29.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	33.3%	23.1%
		OR (95% CI) [b]	1.67 (0.34, 8.07)	
	p-value [b]	0.5256		
	RR (95% CI) [d]	1.44 (0.45, 4.62)		
	Risk Diff (95% CI) [e]	10.3% (-20.3, 40.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.15 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	13
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	8.5%	10.0%
		OR (95% CI) [b]	0.84 (0.14, 4.99)	
		p-value [b]	0.8453	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9982	
		RR (95% CI) [d]	0.85 (0.17, 4.28)	
		Risk Diff (95% CI) [e]	-1.5% (-16.9, 13.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	10.5%
	OR (95% CI) [b]	0.65 (0.10, 4.03)		
	p-value [b]	0.6408		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8796		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.15 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.68 (0.14, 3.31)	
		Risk Diff (95% CI) [e]	-4.9% (-26.8, 17.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1309 von 1722

Table 14.2.18.7.1.5.15 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.7%	33.3%
		OR (95% CI) [b]	0.34 (0.06, 1.85)	
		p-value [b]	0.2141	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4542	
		RR (95% CI) [d]	0.44 (0.13, 1.51)	
		Risk Diff (95% CI) [e]	-18.6% (-51.6, 14.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%
	OR (95% CI) [b]	0.50 (0.09, 2.73)		
	p-value [b]	0.4235		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7125		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1310 von 1722

Table 14.2.18.7.1.5.15 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.60 (0.18, 2.02)	
		Risk Diff (95% CI) [e]	-13.3% (-47.9, 21.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.15 (AMNOG)  
EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.6%	14.3%
		OR (95% CI) [b]	1.11 (0.24, 5.24)	
		p-value [b]	0.8941	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8062	
		RR (95% CI) [d]	1.09 (0.29, 4.10)	
		Risk Diff (95% CI) [e]	1.3% (-18.2, 20.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.7%
	OR (95% CI) [b]	1.11 (0.22, 5.54)		
	p-value [b]	0.8977		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7696		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1312 von 1722

Table 14.2.18.7.1.5.15 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.09 (0.30, 3.89)	
		Risk Diff (95% CI) [e]	1.7% (-24.6, 28.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.5.15 (AMNOG)  
 EORTC QLQ-C30 Emotional Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.9%	15.0%
		OR (95% CI) [b]	0.84 (0.17, 4.22)	
		p-value [b]	0.8319	
		RR (95% CI) [d]	0.86 (0.21, 3.44)	
		Risk Diff (95% CI) [e]	-2.1% (-21.7, 17.5)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.0%	23.1%
		OR (95% CI) [b]	0.78 (0.14, 4.24)	
	p-value [b]	0.7779		
	RR (95% CI) [d]	0.83 (0.22, 3.11)		
	Risk Diff (95% CI) [e]	-4.0% (-32.4, 24.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1314 von 1722

Table 14.2.18.7.1.6.15 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	13
	Number of subjects with clinically meaningful improvement at Week 24	n	6	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.8%	25.0%
		OR (95% CI) [b]	0.44 (0.12, 1.65)	
		p-value [b]	0.2237	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2202	
		RR (95% CI) [d]	0.51 (0.18, 1.48)	
		Risk Diff (95% CI) [e]	-12.2% (-33.5, 9.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	15.8%
	OR (95% CI) [b]	0.30 (0.07, 1.24)		
	p-value [b]	0.0959		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1368		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1315 von 1722

Table 14.2.18.7.1.6.15 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.41 (0.15, 1.12)	
		Risk Diff (95% CI) [e]	-22.7% (-51.5, 6.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1316 von 1722

Table 14.2.18.7.1.6.15 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	6	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.6%	22.2%
		OR (95% CI) [b]	0.75 (0.12, 4.55)	
		p-value [b]	0.7543	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5137	
		RR (95% CI) [d]	0.79 (0.19, 3.29)	
		Risk Diff (95% CI) [e]	-4.6% (-34.6, 25.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.0%
	OR (95% CI) [b]	1.11 (0.18, 6.82)		
	p-value [b]	0.9142		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7361		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.15 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.08 (0.26, 4.41)	
		Risk Diff (95% CI) [e]	1.8% (-30.1, 33.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1318 von 1722

Table 14.2.18.7.1.6.15 (AMNOG)  
EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	8	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.20, 2.24)	
		p-value [b]	0.5112	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3788	
		RR (95% CI) [d]	0.75 (0.32, 1.76)	
		Risk Diff (95% CI) [e]	-8.3% (-33.5, 16.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.8%
	OR (95% CI) [b]	0.61 (0.16, 2.30)		
	p-value [b]	0.4652		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3656		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1319 von 1722

Table 14.2.18.7.1.6.15 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.75 (0.34, 1.62)	
		Risk Diff (95% CI) [e]	-11.9% (-43.8, 20.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.6.15 (AMNOG)  
 EORTC QLQ-C30 Social Functioning - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	5	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.1%	10.0%
		OR (95% CI) [b]	1.73 (0.30, 9.92)	
		p-value [b]	0.5382	
		RR (95% CI) [d]	1.61 (0.35, 7.53)	
		Risk Diff (95% CI) [e]	6.1% (-12.3, 24.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.8%	15.4%
		OR (95% CI) [b]	1.72 (0.28, 10.51)	
	p-value [b]	0.5577		
	RR (95% CI) [d]	1.55 (0.35, 6.85)		
	Risk Diff (95% CI) [e]	8.4% (-18.3, 35.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1321 von 1722

Table 14.2.18.7.1.7.15 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	14
	Number of subjects with clinically meaningful improvement at Week 24	n	11	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	23.4%	35.0%
		OR (95% CI) [b]	0.57 (0.18, 1.77)	
		p-value [b]	0.3301	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3993	
		RR (95% CI) [d]	0.67 (0.30, 1.47)	
		Risk Diff (95% CI) [e]	-11.6% (-35.8, 12.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.9%
	OR (95% CI) [b]	0.41 (0.12, 1.44)		
	p-value [b]	0.1627		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4831		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.15 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.58 (0.28, 1.19)	
		Risk Diff (95% CI) [e]	-21.1% (-51.0, 8.8)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1323 von 1722

Table 14.2.18.7.1.7.15 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	10	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	29.4%	44.4%
		OR (95% CI) [b]	0.52 (0.12, 2.35)	
		p-value [b]	0.3964	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5026	
		RR (95% CI) [d]	0.66 (0.27, 1.62)	
		Risk Diff (95% CI) [e]	-15.0% (-50.9, 20.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	40.0%
	OR (95% CI) [b]	0.83 (0.18, 3.88)		
	p-value [b]	0.8164		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2045		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.15 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.90 (0.38, 2.16)	
		Risk Diff (95% CI) [e]	-4.4% (-42.2, 33.3)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1325 von 1722

Table 14.2.18.7.1.7.15 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	22	15
	Number of subjects with clinically meaningful improvement at Week 24	n	8	7
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	33.3%
		OR (95% CI) [b]	0.67 (0.20, 2.24)	
		p-value [b]	0.5112	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3238	
		RR (95% CI) [d]	0.75 (0.32, 1.76)	
		Risk Diff (95% CI) [e]	-8.3% (-33.5, 16.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	36.4%
	OR (95% CI) [b]	0.65 (0.17, 2.48)		
	p-value [b]	0.5318		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2613		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.15 (AMNOG)  
 EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.78 (0.36, 1.69)	
		Risk Diff (95% CI) [e]	-10.3% (-42.6, 22.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.7.15 (AMNOG)  
EORTC QLQ-C30 Fatigue - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	3	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.7%	30.0%
		OR (95% CI) [b]	0.25 (0.05, 1.15)	
		p-value [b]	0.0752	
		RR (95% CI) [d]	0.32 (0.09, 1.14)	
		Risk Diff (95% CI) [e]	-20.3% (-42.9, 2.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.3%	46.2%
		OR (95% CI) [b]	0.19 (0.04, 1.00)	
	p-value [b]	0.0501		
	RR (95% CI) [d]	0.31 (0.09, 1.03)		
	Risk Diff (95% CI) [e]	-31.9% (-62.8, -0.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1328 von 1722

Table 14.2.18.7.1.8.15 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	14
	Number of subjects with clinically meaningful improvement at Week 24	n	9	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	5.0%
		OR (95% CI) [b]	4.50 (0.53, 38.18)	
		p-value [b]	0.1680	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9585	
		RR (95% CI) [d]	3.83 (0.52, 28.26)	
		Risk Diff (95% CI) [e]	14.1% (-0.6, 28.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.7%
	OR (95% CI) [b]	4.03 (0.46, 35.23)		
	p-value [b]	0.2071		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9605		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1329 von 1722

Table 14.2.18.7.1.8.15 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	3.32 (0.46, 23.85)	
		Risk Diff (95% CI) [e]	16.5% (-2.6, 35.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1330 von 1722

Table 14.2.18.7.1.8.15 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.9%	11.1%
		OR (95% CI) [b]	0.50 (0.04, 6.23)	
		p-value [b]	0.5902	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9486	
		RR (95% CI) [d]	0.53 (0.05, 5.20)	
		Risk Diff (95% CI) [e]	-5.2% (-27.2, 16.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	8.0%
	OR (95% CI) [b]	0.70 (0.06, 8.75)		
	p-value [b]	0.7787		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9532		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.15 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.07, 7.01)	
		Risk Diff (95% CI) [e]	-3.1% (-26.2, 20.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1332 von 1722

Table 14.2.18.7.1.8.15 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.5%	14.3%
		OR (95% CI) [b]	0.86 (0.17, 4.29)	
		p-value [b]	0.8511	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9511	
		RR (95% CI) [d]	0.88 (0.22, 3.52)	
		Risk Diff (95% CI) [e]	-1.8% (-20.6, 17.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	17.4%
	OR (95% CI) [b]	0.84 (0.16, 4.44)		
	p-value [b]	0.8394		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9540		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1333 von 1722

Table 14.2.18.7.1.8.15 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.87 (0.23, 3.35)	
		Risk Diff (95% CI) [e]	-2.6% (-28.1, 22.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.8.15 (AMNOG)  
 EORTC QLQ-C30 Nausea Vomiting - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	2	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.5%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9517	
		RR (95% CI) [d]	3.28 (0.17, 64.99)	
		Risk Diff (95% CI) [e]	5.4% (-5.9, 16.8)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	9.5%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
	p-value [b]	0.9547		
	RR (95% CI) [d]	3.18 (0.16, 61.49)		
	Risk Diff (95% CI) [e]	7.8% (-8.7, 24.2)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.15 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	14
	Number of subjects with clinically meaningful improvement at Week 24	n	9	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	15.0%
		OR (95% CI) [b]	1.34 (0.32, 5.59)	
		p-value [b]	0.6860	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5340	
		RR (95% CI) [d]	1.28 (0.39, 4.23)	
		Risk Diff (95% CI) [e]	4.1% (-15.1, 23.4)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.7%
	OR (95% CI) [b]	1.14 (0.26, 5.00)		
	p-value [b]	0.8641		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5879		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1336 von 1722

Table 14.2.18.7.1.9.15 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.11 (0.35, 3.50)	
		Risk Diff (95% CI) [e]	2.3% (-23.1, 27.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.15 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	4	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.8%	33.3%
		OR (95% CI) [b]	0.27 (0.05, 1.51)	
		p-value [b]	0.1353	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3756	
		RR (95% CI) [d]	0.35 (0.10, 1.30)	
		Risk Diff (95% CI) [e]	-21.6% (-54.2, 11.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.0%
	OR (95% CI) [b]	0.38 (0.07, 2.19)		
	p-value [b]	0.2799		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6559		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.15 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.48 (0.13, 1.74)	
		Risk Diff (95% CI) [e]	-17.3% (-51.3, 16.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.15 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	22	15
	Number of subjects with clinically meaningful improvement at Week 24	n	5	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	15.6%	14.3%
		OR (95% CI) [b]	1.11 (0.24, 5.24)	
		p-value [b]	0.8941	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6790	
		RR (95% CI) [d]	1.09 (0.29, 4.10)	
		Risk Diff (95% CI) [e]	1.3% (-18.2, 20.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	22.7%
	OR (95% CI) [b]	1.18 (0.23, 5.89)		
	p-value [b]	0.8432		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5836		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1340 von 1722

Table 14.2.18.7.1.9.15 (AMNOG)  
 EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.14 (0.32, 4.05)	
		Risk Diff (95% CI) [e]	2.7% (-24.0, 29.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.9.15 (AMNOG)  
EORTC QLQ-C30 Pain - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	6	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.4%	25.0%
		OR (95% CI) [b]	0.72 (0.19, 2.77)	
		p-value [b]	0.6330	
		RR (95% CI) [d]	0.77 (0.27, 2.20)	
		Risk Diff (95% CI) [e]	-5.6% (-29.2, 17.9)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.6%	38.5%
		OR (95% CI) [b]	0.64 (0.15, 2.77)	
	p-value [b]	0.5503		
	RR (95% CI) [d]	0.74 (0.28, 1.95)		
	Risk Diff (95% CI) [e]	-9.9% (-42.6, 22.9)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1342 von 1722

Table 14.2.18.7.1.10.15 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	14
	Number of subjects with clinically meaningful improvement at Week 24	n	9	6
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	30.0%
		OR (95% CI) [b]	0.55 (0.17, 1.84)	
		p-value [b]	0.3331	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2616	
		RR (95% CI) [d]	0.64 (0.26, 1.56)	
		Risk Diff (95% CI) [e]	-10.9% (-33.9, 12.2)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.7%
	OR (95% CI) [b]	0.41 (0.11, 1.51)		
	p-value [b]	0.1821		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1810		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.15 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.55 (0.24, 1.27)	
		Risk Diff (95% CI) [e]	-19.2% (-48.4, 10.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1344 von 1722

Table 14.2.18.7.1.10.15 (AMNOG)  
EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	24	9
	Number of subjects with clinically meaningful improvement at Week 24	n	2	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	5.9%	33.3%
		OR (95% CI) [b]	0.13 (0.02, 0.91)	
		p-value [b]	0.0406	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0417	
		RR (95% CI) [d]	0.18 (0.03, 0.90)	
		Risk Diff (95% CI) [e]	-27.5% (-59.2, 4.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	8.3%
	OR (95% CI) [b]	0.18 (0.02, 1.35)		
	p-value [b]	0.0955		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0886		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1345 von 1722

Table 14.2.18.7.1.10.15 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.25 (0.05, 1.26)	
		Risk Diff (95% CI) [e]	-25.0% (-57.7, 7.7)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1346 von 1722

Table 14.2.18.7.1.10.15 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	20	15
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	18.8%	19.0%
		OR (95% CI) [b]	0.98 (0.24, 4.00)	
		p-value [b]	0.9784	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6173	
		RR (95% CI) [d]	0.98 (0.32, 3.08)	
		Risk Diff (95% CI) [e]	-0.3% (-21.9, 21.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.0%
	OR (95% CI) [b]	1.18 (0.27, 5.24)		
	p-value [b]	0.8290		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7544		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1347 von 1722

Table 14.2.18.7.1.10.15 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.13 (0.38, 3.29)	
		Risk Diff (95% CI) [e]	3.3% (-26.7, 33.4)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.10.15 (AMNOG)  
 EORTC QLQ-C30 Dyspnea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	7	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	22.6%	15.0%
		OR (95% CI) [b]	1.65 (0.37, 7.32)	
		p-value [b]	0.5082	
		RR (95% CI) [d]	1.51 (0.44, 5.15)	
		Risk Diff (95% CI) [e]	7.6% (-13.9, 29.1)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	33.3%	23.1%
		OR (95% CI) [b]	1.67 (0.34, 8.07)	
	p-value [b]	0.5256		
	RR (95% CI) [d]	1.44 (0.45, 4.62)		
	Risk Diff (95% CI) [e]	10.3% (-20.3, 40.8)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1349 von 1722

Table 14.2.18.7.1.11.15 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	37	14
	Number of subjects with clinically meaningful improvement at Week 24	n	9	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.1%	20.0%
		OR (95% CI) [b]	0.95 (0.25, 3.53)	
		p-value [b]	0.9358	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9893	
		RR (95% CI) [d]	0.96 (0.33, 2.75)	
		Risk Diff (95% CI) [e]	-0.9% (-21.7, 20.0)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	24.3%
	OR (95% CI) [b]	0.80 (0.20, 3.20)		
	p-value [b]	0.7564		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9136		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.15 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.85 (0.31, 2.32)	
		Risk Diff (95% CI) [e]	-4.2% (-31.7, 23.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1351 von 1722

Table 14.2.18.7.1.11.15 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	5	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.7%	44.4%
		OR (95% CI) [b]	0.22 (0.04, 1.09)	
		p-value [b]	0.0636	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1733	
		RR (95% CI) [d]	0.33 (0.11, 0.98)	
		Risk Diff (95% CI) [e]	-29.7% (-64.3, 4.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.0%
	OR (95% CI) [b]	0.31 (0.06, 1.61)		
	p-value [b]	0.1645		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3525		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.11.15 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.45 (0.15, 1.31)	
		Risk Diff (95% CI) [e]	-24.4% (-60.5, 11.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1353 von 1722

Table 14.2.18.7.1.11.15 (AMNOG)  
EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	22	15
	Number of subjects with clinically meaningful improvement at Week 24	n	8	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	25.0%	23.8%
		OR (95% CI) [b]	1.07 (0.30, 3.85)	
		p-value [b]	0.9215	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9138	
		RR (95% CI) [d]	1.05 (0.40, 2.78)	
		Risk Diff (95% CI) [e]	1.2% (-22.4, 24.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	36.4%
	OR (95% CI) [b]	1.14 (0.29, 4.55)		
	p-value [b]	0.8497		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8191		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1354 von 1722

Table 14.2.18.7.1.11.15 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.09 (0.44, 2.69)	
		Risk Diff (95% CI) [e]	3.0% (-28.2, 34.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1355 von 1722

Table 14.2.18.7.1.11.15 (AMNOG)  
 EORTC QLQ-C30 Insomnia - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	6	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	19.4%	20.0%
		OR (95% CI) [b]	0.96 (0.23, 3.94)	
		p-value [b]	0.9548	
		RR (95% CI) [d]	0.97 (0.31, 3.01)	
		Risk Diff (95% CI) [e]	-0.6% (-23.0, 21.7)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.6%	30.8%
		OR (95% CI) [b]	0.90 (0.20, 4.08)	
	p-value [b]	0.8913		
	RR (95% CI) [d]	0.93 (0.32, 2.68)		
	Risk Diff (95% CI) [e]	-2.2% (-33.9, 29.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1356 von 1722

Table 14.2.18.7.1.12.15 (AMNOG)  
EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	14
	Number of subjects with clinically meaningful improvement at Week 24	n	8	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	17.0%	20.0%
		OR (95% CI) [b]	0.82 (0.22, 3.11)	
		p-value [b]	0.7713	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2594	
		RR (95% CI) [d]	0.85 (0.29, 2.51)	
		Risk Diff (95% CI) [e]	-3.0% (-23.5, 17.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	21.1%
	OR (95% CI) [b]	0.67 (0.16, 2.70)		
	p-value [b]	0.5696		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2047		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1357 von 1722

Table 14.2.18.7.1.12.15 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.74 (0.26, 2.07)	
		Risk Diff (95% CI) [e]	-7.5% (-34.5, 19.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.15 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	24	9
	Number of subjects with clinically meaningful improvement at Week 24	n	5	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	14.7%	11.1%
		OR (95% CI) [b]	1.38 (0.14, 13.56)	
		p-value [b]	0.7827	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5496	
		RR (95% CI) [d]	1.32 (0.18, 9.95)	
		Risk Diff (95% CI) [e]	3.6% (-20.1, 27.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	20.8%
	OR (95% CI) [b]	2.11 (0.21, 21.01)		
	p-value [b]	0.5259		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7267		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.15 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.88 (0.25, 13.93)	
		Risk Diff (95% CI) [e]	9.7% (-16.5, 35.9)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1360 von 1722

Table 14.2.18.7.1.12.15 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	3	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.4%	14.3%
		OR (95% CI) [b]	0.62 (0.11, 3.41)	
		p-value [b]	0.5835	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2154	
		RR (95% CI) [d]	0.66 (0.15, 2.95)	
		Risk Diff (95% CI) [e]	-4.9% (-23.0, 13.1)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.0%
	OR (95% CI) [b]	0.60 (0.10, 3.46)		
	p-value [b]	0.5679		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2110		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1361 von 1722

Table 14.2.18.7.1.12.15 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.65 (0.15, 2.81)	
		Risk Diff (95% CI) [e]	-7.0% (-31.4, 17.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.12.15 (AMNOG)  
 EORTC QLQ-C30 Appetite Loss - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	5	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	16.1%	5.0%
		OR (95% CI) [b]	3.65 (0.39, 33.88)	
		p-value [b]	0.2541	
		RR (95% CI) [d]	3.23 (0.41, 25.62)	
		Risk Diff (95% CI) [e]	11.1% (-5.0, 27.2)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	23.8%	7.7%
		OR (95% CI) [b]	3.75 (0.39, 36.40)	
	p-value [b]	0.2547		
	RR (95% CI) [d]	3.10 (0.41, 23.63)		
	Risk Diff (95% CI) [e]	16.1% (-7.2, 39.4)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1363 von 1722

Table 14.2.18.7.1.13.15 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	13
	Number of subjects with clinically meaningful improvement at Week 24	n	13	5
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	27.7%	25.0%
		OR (95% CI) [b]	1.15 (0.35, 3.80)	
		p-value [b]	0.8223	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8789	
		RR (95% CI) [d]	1.11 (0.45, 2.69)	
		Risk Diff (95% CI) [e]	2.7% (-20.2, 25.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	34.2%
	OR (95% CI) [b]	0.83 (0.23, 3.06)		
	p-value [b]	0.7820		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9352		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.15 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.89 (0.39, 2.01)	
		Risk Diff (95% CI) [e]	-4.3% (-34.7, 26.2)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.15 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	7	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	20.6%	33.3%
		OR (95% CI) [b]	0.52 (0.10, 2.61)	
		p-value [b]	0.4257	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6243	
		RR (95% CI) [d]	0.62 (0.20, 1.92)	
		Risk Diff (95% CI) [e]	-12.7% (-46.4, 20.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	28.0%
	OR (95% CI) [b]	0.78 (0.15, 4.00)		
	p-value [b]	0.7636		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8991		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.15 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.84 (0.27, 2.57)	
		Risk Diff (95% CI) [e]	-5.3% (-40.8, 30.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.15 (AMNOG)  
EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	7	3
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	21.9%	14.3%
		OR (95% CI) [b]	1.68 (0.38, 7.39)	
		p-value [b]	0.4926	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6500	
		RR (95% CI) [d]	1.53 (0.45, 5.27)	
		Risk Diff (95% CI) [e]	7.6% (-13.1, 28.3)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	30.4%
	OR (95% CI) [b]	1.75 (0.37, 8.21)		
	p-value [b]	0.4780		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6094		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.13.15 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.52 (0.46, 4.98)	
		Risk Diff (95% CI) [e]	10.4% (-17.2, 38.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1369 von 1722

Table 14.2.18.7.1.13.15 (AMNOG)  
 EORTC QLQ-C30 Constipation - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	3	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.7%	10.0%
		OR (95% CI) [b]	0.96 (0.15, 6.35)	
		p-value [b]	0.9699	
		RR (95% CI) [d]	0.97 (0.18, 5.29)	
		Risk Diff (95% CI) [e]	-0.3% (-17.1, 16.4)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	14.3%	15.4%
		OR (95% CI) [b]	0.92 (0.13, 6.38)	
	p-value [b]	0.9300		
	RR (95% CI) [d]	0.93 (0.18, 4.84)		
	Risk Diff (95% CI) [e]	-1.1% (-25.8, 23.6)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1370 von 1722

Table 14.2.18.7.1.14.15 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	38	13
	Number of subjects with clinically meaningful improvement at Week 24	n	3	0
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.4%	0.0%
		OR (95% CI) [b]	NE (NE, NE)	
		p-value [b]	0.9573	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9584	
		RR (95% CI) [d]	3.06 (0.17, 56.70)	
		Risk Diff (95% CI) [e]	4.9% (-4.9, 14.7)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	7.9%
	OR (95% CI) [b]	NE (NE, NE)		
	p-value [b]	0.9604		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9613		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.15 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	2.51 (0.14, 45.65)	
		Risk Diff (95% CI) [e]	5.4% (-7.8, 18.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.15 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	25	9
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	11.8%	22.2%
		OR (95% CI) [b]	0.47 (0.07, 3.08)	
		p-value [b]	0.4284	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5149	
		RR (95% CI) [d]	0.53 (0.11, 2.45)	
		Risk Diff (95% CI) [e]	-10.5% (-39.7, 18.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	16.0%
	OR (95% CI) [b]	0.67 (0.10, 4.46)		
	p-value [b]	0.6759		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6905		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.15 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.72 (0.16, 3.28)	
		Risk Diff (95% CI) [e]	-6.2% (-37.0, 24.5)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.15 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	2	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.3%	9.5%
		OR (95% CI) [b]	0.63 (0.08, 4.88)	
		p-value [b]	0.6612	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6565	
		RR (95% CI) [d]	0.66 (0.10, 4.31)	
		Risk Diff (95% CI) [e]	-3.3% (-18.4, 11.8)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	8.7%
	OR (95% CI) [b]	0.62 (0.08, 4.95)		
	p-value [b]	0.6511		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6677		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.14.15 (AMNOG)  
 EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.65 (0.10, 4.14)	
		Risk Diff (95% CI) [e]	-4.6% (-25.3, 16.1)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1376 von 1722

Table 14.2.18.7.1.14.15 (AMNOG)  
EORTC QLQ-C30 Diarrhea - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	2	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	6.5%	5.0%
		OR (95% CI) [b]	1.31 (0.11, 15.48)	
		p-value [b]	0.8301	
		RR (95% CI) [d]	1.29 (0.13, 13.31)	
		Risk Diff (95% CI) [e]	1.5% (-11.4, 14.3)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	9.5%	7.7%
		OR (95% CI) [b]	1.26 (0.10, 15.49)	
	p-value [b]	0.8551		
	RR (95% CI) [d]	1.24 (0.12, 12.33)		
	Risk Diff (95% CI) [e]	1.8% (-17.3, 21.0)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1377 von 1722

Table 14.2.18.7.1.15.15 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects at baseline	N	47	20
	Number of HRQoL evaluable subjects at Week 24	N	37	13
	Number of subjects with clinically meaningful improvement at Week 24	n	5	1
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	10.6%	5.0%
		OR (95% CI) [b]	2.26 (0.25, 20.71)	
		p-value [b]	0.4700	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7165	
		RR (95% CI) [d]	2.13 (0.27, 17.07)	
		Risk Diff (95% CI) [e]	5.6% (-7.4, 18.6)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.5%
	OR (95% CI) [b]	1.87 (0.20, 17.74)		
	p-value [b]	0.5835		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8032		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1378 von 1722

Table 14.2.18.7.1.15.15 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	1.76 (0.23, 13.67)	
		Risk Diff (95% CI) [e]	5.8% (-12.4, 24.0)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1379 von 1722

Table 14.2.18.7.1.15.15 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects at baseline	N	34	9
	Number of HRQoL evaluable subjects at Week 24	N	24	9
	Number of subjects with clinically meaningful improvement at Week 24	n	1	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	2.9%	22.2%
		OR (95% CI) [b]	0.11 (0.01, 1.34)	
		p-value [b]	0.0828	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1105	
		RR (95% CI) [d]	0.13 (0.01, 1.30)	
		Risk Diff (95% CI) [e]	-19.3% (-47.0, 8.5)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	4.2%
	OR (95% CI) [b]	0.15 (0.01, 1.94)		
	p-value [b]	0.1471		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1832		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1380 von 1722

Table 14.2.18.7.1.15.15 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.19 (0.02, 1.82)	
		Risk Diff (95% CI) [e]	-18.1% (-46.4, 10.3)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.15 (AMNOG)  
EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects at baseline	N	32	21
	Number of HRQoL evaluable subjects at Week 24	N	23	15
	Number of subjects with clinically meaningful improvement at Week 24	n	3	4
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	9.4%	19.0%
		OR (95% CI) [b]	0.44 (0.09, 2.20)	
		p-value [b]	0.3178	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3681	
		RR (95% CI) [d]	0.49 (0.12, 1.98)	
		Risk Diff (95% CI) [e]	-9.7% (-29.3, 9.9)	
		Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	13.0%
	OR (95% CI) [b]	0.41 (0.08, 2.19)		
	p-value [b]	0.2981		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3696		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.7.1.15.15 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
		RR (95% CI) [d]	0.49 (0.13, 1.88)	
		Risk Diff (95% CI) [e]	-13.6% (-39.9, 12.6)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1383 von 1722

Table 14.2.18.7.1.15.15 (AMNOG)  
 EORTC QLQ-C30 Financial Difficulties - Proportions of Subjects with Clinically-Meaningful Improvement  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects at baseline	N	31	20
	Number of HRQoL evaluable subjects at Week 24	N	21	13
	Number of subjects with clinically meaningful improvement at Week 24	n	4	2
	Percent of baseline HRQoL evaluable subjects with clinically meaningful improvement at Week 24 [a]	%	12.9%	10.0%
		OR (95% CI) [b]	1.33 (0.22, 8.06)	
		p-value [b]	0.7540	
		RR (95% CI) [d]	1.29 (0.26, 6.40)	
		Risk Diff (95% CI) [e]	2.9% (-14.8, 20.6)	
	Percent of HRQoL evaluable subjects with clinically meaningful improvement at Week 24 (excluding subjects with missing scores at Week 24 from the denominator)	%	19.0%	15.4%
		OR (95% CI) [b]	1.29 (0.20, 8.31)	
	p-value [b]	0.7858		
	RR (95% CI) [d]	1.24 (0.26, 5.83)		
	Risk Diff (95% CI) [e]	3.7% (-22.2, 29.5)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. A clinically meaningful improvement is defined based on a pre-specified responder definition of 10. BSC = Best Supportive Care; CI = Confidence Interval; Coeff = Coefficient; Diff = Difference; HRQoL = Health-related Quality of Life; OR = Odds Ratio; RR = Relative Risk; SD = Standard Deviation.

[a] Subjects with missing change score at Week 24 are assumed to be non-responders.

[b] Odds ratios (OR; 95% CI and p-value) are calculated based on logistic regression models with treatment group (luspatercept+BSC vs. placebo+BSC), subgroup, and subgroup by treatment interaction as covariates.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

[d] Confidence intervals for relative risk (RR) comparing luspatercept+BSC vs. placebo+BSC are calculated assuming normal distribution. RR is not estimated if no events occur in both treatment groups.

[e] Risk differences between treatment (luspatercept+BSC vs. placebo+BSC) groups are estimated assuming a normal distribution.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQLQC30

Run Date: 20AUG2019

Data Extraction Date: 21JUN2018

Luspatercept (Reblozyl®)

Seite 1384 von 1722

Table 14.2.18.6.1.2.1.4 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects	N	31	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.66 (17.58)	56.58 (19.26)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (93.5%)	15 (78.9%)
	Observed mean domain score at Week 24	Mean (SD)	64.22 (21.84)	53.33 (17.34)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.53 (2.79)	-0.16 (3.72)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	3.69 (4.65)	
		95% CI	-5.46, 12.84	
		p-value [b]	0.4280	
		Effect size by Hedges' g (95% CI)	0.20 (-0.36, 0.76)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0146	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.4 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	115 (97.5%)	57 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.61 (21.36)	49.78 (23.92)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	81 (68.6%)	38 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	50.77 (22.04)	54.61 (26.22)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.42 (1.93)	2.21 (2.70)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.64 (3.32)	
		95% CI	-14.17, -1.10	
		p-value [b]	0.0222	
	Effect size by Hedges' g (95% CI)	-0.34 (-0.66, -0.02)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.1 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	28 (100.0%)	16 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.14 (21.09)	56.25 (20.92)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (92.9%)	11 (68.8%)
	Observed mean domain score at Week 24	Mean (SD)	55.77 (21.57)	52.27 (26.70)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.07 (2.96)	-2.72 (3.99)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.79 (4.97)	
		95% CI	-6.99, 12.57	
		p-value [b]	0.5753	
		Effect size by Hedges' g (95% CI)	0.13 (-0.47, 0.73)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0034	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.1 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	69 (98.6%)	29 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.25 (20.94)	51.29 (19.00)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (61.4%)	22 (75.9%)
	Observed mean domain score at Week 24	Mean (SD)	56.10 (21.71)	57.95 (19.88)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.66 (2.24)	6.11 (3.28)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-12.76 (3.97)	
		95% CI	-20.58, -4.94	
		p-value [b]	0.0015	
	Effect size by Hedges' g (95% CI)	-0.62 (-1.06, -0.18)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.1 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	49 (96.1%)	31 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	49.49 (19.76)	49.19 (27.18)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	41 (80.4%)	20 (64.5%)
	Observed mean domain score at Week 24	Mean (SD)	51.52 (24.56)	51.25 (26.87)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.79 (2.49)	-0.39 (3.22)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.40 (4.07)	
		95% CI	-8.41, 7.61	
		p-value [b]	0.9223	
	Effect size by Hedges' g (95% CI)	-0.02 (-0.46, 0.43)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.1 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	28 (100.0%)	16 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.33 (31.66)	50.00 (31.69)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (92.9%)	11 (68.8%)
	Observed mean domain score at Week 24	Mean (SD)	62.82 (31.10)	54.55 (35.99)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.89 (4.77)	-8.31 (6.47)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	11.20 (8.04)	
		95% CI	-4.64, 27.04	
		p-value [b]	0.1650	
		Effect size by Hedges' g (95% CI)	0.35 (-0.26, 0.95)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0223	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.1 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	69 (98.6%)	29 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.72 (33.25)	47.89 (28.96)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (61.4%)	22 (75.9%)
	Observed mean domain score at Week 24	Mean (SD)	50.65 (32.92)	59.60 (32.08)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.99 (3.46)	8.26 (5.10)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-14.26 (6.17)	
		95% CI	-26.41, -2.10	
		p-value [b]	0.0217	
	Effect size by Hedges' g (95% CI)	-0.44 (-0.88, -0.01)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.1 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	49 (96.1%)	31 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.61 (35.84)	54.48 (29.59)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	39 (76.5%)	20 (64.5%)
	Observed mean domain score at Week 24	Mean (SD)	52.42 (33.33)	66.11 (31.73)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.67 (3.91)	1.74 (4.98)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.41 (6.34)	
		95% CI	-19.88, 5.07	
		p-value [b]	0.2433	
Effect size by Hedges' g (95% CI)		-0.22 (-0.67, 0.23)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.1 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	24 (85.7%)	14 (87.5%)
	Observed mean domain score at baseline	Mean (SD)	52.08 (37.53)	46.43 (41.44)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (85.7%)	10 (62.5%)
	Observed mean domain score at Week 24	Mean (SD)	56.25 (35.55)	50.00 (37.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.80 (4.54)	-3.15 (6.12)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.35 (7.62)	
		95% CI	-12.67, 17.36	
		p-value [b]	0.7584	
		Effect size by Hedges' g (95% CI)	0.06 (-0.59, 0.70)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1146	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.1 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	64 (91.4%)	26 (89.7%)
	Observed mean domain score at baseline	Mean (SD)	50.00 (37.27)	48.08 (36.69)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	40 (57.1%)	20 (69.0%)
	Observed mean domain score at Week 24	Mean (SD)	45.00 (34.99)	48.75 (39.30)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-11.08 (3.19)	3.66 (4.82)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-14.74 (5.77)	
		95% CI	-26.12, -3.37	
		p-value [b]	0.0113	
	Effect size by Hedges' g (95% CI)	-0.39 (-0.85, 0.06)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.1 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	49 (96.1%)	28 (90.3%)
	Observed mean domain score at baseline	Mean (SD)	45.92 (38.98)	47.32 (38.09)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	38 (74.5%)	18 (58.1%)
	Observed mean domain score at Week 24	Mean (SD)	46.05 (34.15)	61.11 (36.60)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.83 (3.52)	3.57 (4.71)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.40 (5.88)	
		95% CI	-17.00, 6.19	
		p-value [b]	0.3594	
	Effect size by Hedges' g (95% CI)	-0.14 (-0.60, 0.32)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.1 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	26 (92.9%)	15 (93.8%)
	Observed mean domain score at baseline	Mean (SD)	57.69 (33.91)	64.44 (36.39)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (92.9%)	11 (68.8%)
	Observed mean domain score at Week 24	Mean (SD)	66.35 (32.95)	57.58 (32.59)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.87 (4.61)	-4.15 (6.15)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	11.02 (7.69)	
		95% CI	-4.14, 26.18	
		p-value [b]	0.1534	
		Effect size by Hedges' g (95% CI)	0.31 (-0.32, 0.94)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0727	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.1 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	68 (97.1%)	29 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.72 (35.10)	49.14 (37.36)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	42 (60.0%)	21 (72.4%)
	Observed mean domain score at Week 24	Mean (SD)	58.53 (37.34)	58.73 (38.86)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.18 (3.28)	5.72 (4.84)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.91 (5.83)	
		95% CI	-19.39, 3.57	
		p-value [b]	0.1763	
Effect size by Hedges' g (95% CI)		-0.22 (-0.65, 0.21)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.1 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (92.2%)	27 (87.1%)
	Observed mean domain score at baseline	Mean (SD)	70.92 (34.49)	76.23 (39.00)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	37 (72.5%)	19 (61.3%)
	Observed mean domain score at Week 24	Mean (SD)	73.87 (32.76)	67.98 (37.60)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.60 (3.77)	-0.42 (4.94)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	3.02 (6.19)	
		95% CI	-9.16, 15.21	
		p-value [b]	0.6259	
	Effect size by Hedges' g (95% CI)	0.08 (-0.39, 0.55)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.1 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	28 (100.0%)	16 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.51 (14.76)	69.05 (14.34)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (92.9%)	11 (68.8%)
	Observed mean domain score at Week 24	Mean (SD)	74.18 (12.72)	72.29 (15.33)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.42 (1.92)	-1.72 (2.62)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.31 (3.24)	
		95% CI	-6.08, 6.69	
		p-value [b]	0.9248	
		Effect size by Hedges' g (95% CI)	0.02 (-0.58, 0.62)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1193	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.1 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	70 (100.0%)	29 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	76.67 (13.62)	75.04 (14.32)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (61.4%)	22 (75.9%)
	Observed mean domain score at Week 24	Mean (SD)	72.87 (15.33)	74.46 (15.86)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.61 (1.43)	2.37 (2.10)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.98 (2.54)	
		95% CI	-11.98, -1.97	
		p-value [b]	0.0065	
Effect size by Hedges' g (95% CI)		-0.50 (-0.94, -0.07)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.1 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	51 (100.0%)	31 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	74.04 (15.40)	74.81 (13.98)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	41 (80.4%)	20 (64.5%)
	Observed mean domain score at Week 24	Mean (SD)	73.29 (15.61)	83.81 (10.18)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.68 (1.57)	3.44 (2.07)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.12 (2.60)	
		95% CI	-10.24, 0.00	
		p-value [b]	0.0502	
Effect size by Hedges' g (95% CI)		-0.34 (-0.79, 0.10)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.1 (AMNOG)  
 QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	27 (96.4%)	16 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.29 (26.65)	56.99 (24.21)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (92.9%)	11 (68.8%)
	Observed mean domain score at Week 24	Mean (SD)	62.27 (24.75)	58.44 (27.77)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.09 (3.08)	-1.89 (4.09)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.98 (5.12)	
		95% CI	-7.10, 13.07	
		p-value [b]	0.5607	
		Effect size by Hedges' g (95% CI)	0.11 (-0.49, 0.72)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2070	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.1 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	70 (100.0%)	29 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.50 (22.39)	56.98 (22.73)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (61.4%)	22 (75.9%)
	Observed mean domain score at Week 24	Mean (SD)	56.98 (26.19)	55.95 (25.22)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.81 (2.18)	2.30 (3.24)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.11 (3.91)	
		95% CI	-13.81, 1.58	
		p-value [b]	0.1189	
Effect size by Hedges' g (95% CI)		-0.27 (-0.70, 0.16)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.1 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	51 (100.0%)	30 (96.8%)	
	Observed mean domain score at baseline	Mean (SD)	53.41 (25.36)	58.10 (23.52)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	40 (78.4%)	20 (64.5%)	
	Observed mean domain score at Week 24	Mean (SD)	62.38 (24.55)	62.38 (22.64)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	3.99 (2.44)	3.37 (3.22)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.63 (4.04)		
		95% CI	-7.33, 8.58		
		p-value [b]	0.8767		
Effect size by Hedges' g (95% CI)		0.03 (-0.42, 0.47)			

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.1 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	24 (85.7%)	14 (87.5%)
	Observed mean domain score at baseline	Mean (SD)	61.94 (21.97)	57.50 (22.86)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (85.7%)	10 (62.5%)
	Observed mean domain score at Week 24	Mean (SD)	63.27 (20.74)	58.49 (22.23)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.27 (2.64)	-3.23 (3.53)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.96 (4.41)	
		95% CI	-5.73, 11.64	
		p-value [b]	0.5026	
		Effect size by Hedges' g (95% CI)	0.13 (-0.52, 0.78)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0082	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.1 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	64 (91.4%)	26 (89.7%)
	Observed mean domain score at baseline	Mean (SD)	58.82 (19.26)	55.23 (19.14)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	40 (57.1%)	20 (69.0%)
	Observed mean domain score at Week 24	Mean (SD)	56.82 (18.94)	59.20 (21.03)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.40 (1.84)	5.79 (2.76)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-12.18 (3.32)	
		95% CI	-18.73, -5.64	
		p-value [b]	0.0003	
Effect size by Hedges' g (95% CI)		-0.63 (-1.09, -0.17)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.1 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (92.2%)	28 (90.3%)	
	Observed mean domain score at baseline	Mean (SD)	58.70 (23.68)	59.32 (21.74)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	36 (70.6%)	18 (58.1%)	
	Observed mean domain score at Week 24	Mean (SD)	59.22 (21.68)	68.00 (20.84)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	-1.76 (2.07)	2.37 (2.70)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.13 (3.40)		
		95% CI	-10.83, 2.57		
		p-value [b]	0.2255		
Effect size by Hedges' g (95% CI)		-0.18 (-0.64, 0.29)			

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.1 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	24 (85.7%)	14 (87.5%)
	Observed mean domain score at baseline	Mean (SD)	59.74 (23.73)	56.98 (23.42)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (85.7%)	10 (62.5%)
	Observed mean domain score at Week 24	Mean (SD)	63.23 (21.75)	59.48 (24.07)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.37 (2.71)	-1.87 (3.62)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	1.50 (4.52)	
		95% CI	-7.42, 10.41	
		p-value [b]	0.7409	
		Effect size by Hedges' g (95% CI)	0.06 (-0.58, 0.71)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0301	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.1 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	64 (91.4%)	26 (89.7%)
	Observed mean domain score at baseline	Mean (SD)	58.89 (19.04)	56.83 (19.01)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	40 (57.1%)	20 (69.0%)
	Observed mean domain score at Week 24	Mean (SD)	56.15 (21.50)	58.23 (22.39)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.54 (1.92)	4.89 (2.87)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.44 (3.45)	
		95% CI	-17.24, -3.64	
		p-value [b]	0.0028	
Effect size by Hedges' g (95% CI)		-0.54 (-1.00, -0.08)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.1 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (92.2%)	27 (87.1%)	
	Observed mean domain score at baseline	Mean (SD)	56.35 (23.00)	58.96 (22.50)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	36 (70.6%)	18 (58.1%)	
	Observed mean domain score at Week 24	Mean (SD)	60.93 (21.86)	65.81 (20.15)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	1.71 (2.15)	3.50 (2.85)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-1.79 (3.57)	
			95% CI	-8.83, 5.25	
			p-value [b]	0.6162	
Effect size by Hedges' g (95% CI)			-0.08 (-0.55, 0.39)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.1 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≤ 64	Number of HRQoL evaluable subjects	N	28	16
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	27 (96.4%)	16 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.94 (21.19)	54.41 (20.38)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (92.9%)	11 (68.8%)
	Observed mean domain score at Week 24	Mean (SD)	60.29 (22.81)	55.09 (22.47)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.42 (2.87)	-4.13 (3.80)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	4.56 (4.76)	
		95% CI	-4.82, 13.94	
		p-value [b]	0.3394	
		Effect size by Hedges' g (95% CI)	0.21 (-0.39, 0.82)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0062	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.1 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age 65-74	Number of HRQoL evaluable subjects	N	70	29
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	68 (97.1%)	29 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.47 (19.79)	52.05 (18.10)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (61.4%)	22 (75.9%)
	Observed mean domain score at Week 24	Mean (SD)	54.58 (21.99)	57.83 (22.02)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.63 (2.06)	6.09 (3.02)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-11.72 (3.66)	
		95% CI	-18.94, -4.51	
		p-value [b]	0.0016	
Effect size by Hedges' g (95% CI)		-0.60 (-1.04, -0.16)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.1 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Age ≥ 75	Number of HRQoL evaluable subjects	N	51	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	49 (96.1%)	30 (96.8%)
	Observed mean domain score at baseline	Mean (SD)	51.98 (22.14)	54.15 (22.39)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	38 (74.5%)	20 (64.5%)
	Observed mean domain score at Week 24	Mean (SD)	56.57 (22.89)	59.91 (22.20)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.59 (2.33)	1.10 (2.99)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.69 (3.79)	
		95% CI	-9.15, 5.77	
		p-value [b]	0.6560	
Effect size by Hedges' g (95% CI)		-0.08 (-0.53, 0.37)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.2 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	90 (96.8%)	50 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.17 (20.53)	51.75 (24.62)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	67 (72.0%)	34 (68.0%)
	Observed mean domain score at Week 24	Mean (SD)	58.40 (22.02)	54.41 (25.91)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.84 (2.08)	1.08 (2.85)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.91 (3.53)	
		95% CI	-9.86, 4.03	
		p-value [b]	0.4094	
		Effect size by Hedges' g (95% CI)	-0.13 (-0.48, 0.21)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1978	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.2 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	56 (100.0%)	26 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.90 (21.05)	50.96 (19.66)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (76.8%)	19 (73.1%)
	Observed mean domain score at Week 24	Mean (SD)	47.97 (22.48)	53.95 (20.43)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.65 (2.37)	2.62 (3.40)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.27 (4.14)	
		95% CI	-16.42, -0.12	
		p-value [b]	0.0468	
	Effect size by Hedges' g (95% CI)	-0.40 (-0.86, 0.07)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.2 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	91 (97.8%)	50 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.58 (33.65)	54.44 (28.68)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	66 (71.0%)	34 (68.0%)
	Observed mean domain score at Week 24	Mean (SD)	53.87 (32.45)	63.07 (32.45)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.66 (3.12)	2.08 (4.26)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.75 (5.28)	
		95% CI	-15.15, 5.66	
		p-value [b]	0.3697	
		Effect size by Hedges' g (95% CI)	-0.15 (-0.49, 0.20)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6132	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.2 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	55 (98.2%)	26 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.16 (34.04)	44.44 (30.63)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	42 (75.0%)	19 (73.1%)
	Observed mean domain score at Week 24	Mean (SD)	54.76 (33.50)	57.31 (32.97)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.11 (3.73)	2.27 (5.34)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.38 (6.52)	
		95% CI	-21.23, 4.46	
		p-value [b]	0.1999	
Effect size by Hedges' g (95% CI)		-0.25 (-0.72, 0.21)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.2 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	90 (96.8%)	43 (86.0%)
	Observed mean domain score at baseline	Mean (SD)	50.83 (37.75)	49.42 (39.52)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	63 (67.7%)	31 (62.0%)
	Observed mean domain score at Week 24	Mean (SD)	50.79 (35.06)	62.10 (38.12)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.13 (2.84)	3.95 (4.18)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.09 (5.05)	
		95% CI	-20.04, -0.13	
		p-value [b]	0.0471	
		Effect size by Hedges' g (95% CI)	-0.26 (-0.62, 0.10)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4375	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.2 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (83.9%)	25 (96.2%)
	Observed mean domain score at baseline	Mean (SD)	45.21 (37.82)	44.00 (34.82)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	39 (69.6%)	17 (65.4%)
	Observed mean domain score at Week 24	Mean (SD)	43.59 (34.28)	38.24 (32.01)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.44 (3.55)	-0.43 (4.88)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.02 (6.03)	
		95% CI	-16.90, 6.87	
		p-value [b]	0.4065	
	Effect size by Hedges' g (95% CI)	-0.13 (-0.62, 0.35)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.2 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	89 (95.7%)	46 (92.0%)
	Observed mean domain score at baseline	Mean (SD)	54.31 (34.98)	55.25 (42.43)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	66 (71.0%)	33 (66.0%)
	Observed mean domain score at Week 24	Mean (SD)	57.32 (37.57)	59.85 (37.81)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.92 (3.00)	-0.68 (4.12)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.23 (5.08)	
		95% CI	-12.23, 7.77	
		p-value [b]	0.6607	
		Effect size by Hedges' g (95% CI)	-0.06 (-0.41, 0.30)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3891	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.2 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	52 (92.9%)	25 (96.2%)
	Observed mean domain score at baseline	Mean (SD)	75.48 (31.02)	76.33 (28.33)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	39 (69.6%)	18 (69.2%)
	Observed mean domain score at Week 24	Mean (SD)	80.34 (24.52)	65.74 (35.34)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	8.38 (3.59)	4.93 (5.04)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	3.45 (6.14)	
		95% CI	-8.63, 15.54	
		p-value [b]	0.5743	
	Effect size by Hedges' g (95% CI)	0.11 (-0.36, 0.59)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.2 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	93 (100.0%)	50 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	76.86 (14.42)	74.67 (15.50)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	67 (72.0%)	34 (68.0%)
	Observed mean domain score at Week 24	Mean (SD)	75.91 (14.26)	80.11 (14.04)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.86 (1.31)	2.44 (1.82)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.30 (2.24)	
		95% CI	-8.71, 0.12	
		p-value [b]	0.0563	
		Effect size by Hedges' g (95% CI)	-0.29 (-0.63, 0.05)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5790	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.2 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	56 (100.0%)	26 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	73.38 (14.26)	71.79 (11.34)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (76.8%)	19 (73.1%)
	Observed mean domain score at Week 24	Mean (SD)	69.32 (14.74)	72.93 (14.47)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.78 (1.52)	1.04 (2.19)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.82 (2.67)	
		95% CI	-11.06, -0.57	
		p-value [b]	0.0300	
Effect size by Hedges' g (95% CI)		-0.43 (-0.89, 0.04)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.2 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	92 (98.9%)	50 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.41 (22.86)	59.10 (25.15)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	67 (72.0%)	34 (68.0%)
	Observed mean domain score at Week 24	Mean (SD)	61.34 (25.93)	62.18 (25.57)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.50 (1.99)	1.72 (2.75)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.22 (3.39)	
		95% CI	-8.90, 4.46	
		p-value [b]	0.5135	
		Effect size by Hedges' g (95% CI)	-0.09 (-0.44, 0.25)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9311	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.2 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	56 (100.0%)	25 (96.2%)
	Observed mean domain score at baseline	Mean (SD)	53.91 (26.21)	54.10 (18.17)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	42 (75.0%)	19 (73.1%)
	Observed mean domain score at Week 24	Mean (SD)	58.45 (24.08)	53.01 (21.89)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.26 (2.37)	2.08 (3.47)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.83 (4.20)	
		95% CI	-10.09, 6.44	
		p-value [b]	0.6637	
Effect size by Hedges' g (95% CI)		-0.08 (-0.54, 0.39)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.2 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	88 (94.6%)	43 (86.0%)
	Observed mean domain score at baseline	Mean (SD)	58.14 (21.00)	57.23 (22.68)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	62 (66.7%)	31 (62.0%)
	Observed mean domain score at Week 24	Mean (SD)	59.36 (20.98)	65.44 (22.76)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.30 (1.65)	2.95 (2.37)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.25 (2.89)	
		95% CI	-11.94, -0.55	
		p-value [b]	0.0317	
		Effect size by Hedges' g (95% CI)	-0.29 (-0.65, 0.08)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9666	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.2 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (83.9%)	25 (96.2%)
	Observed mean domain score at baseline	Mean (SD)	61.56 (21.74)	57.65 (17.48)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	38 (67.9%)	17 (65.4%)
	Observed mean domain score at Week 24	Mean (SD)	59.03 (19.52)	56.71 (17.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.21 (2.06)	1.88 (2.81)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.09 (3.49)	
		95% CI	-12.96, 0.79	
		p-value [b]	0.0823	
Effect size by Hedges' g (95% CI)		-0.30 (-0.78, 0.19)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.2 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	88 (94.6%)	43 (86.0%)
	Observed mean domain score at baseline	Mean (SD)	58.35 (20.25)	58.60 (23.28)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	62 (66.7%)	31 (62.0%)
	Observed mean domain score at Week 24	Mean (SD)	60.30 (22.32)	64.56 (22.72)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.03 (1.75)	2.92 (2.51)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.94 (3.05)	
		95% CI	-10.97, 1.08	
		p-value [b]	0.1070	
		Effect size by Hedges' g (95% CI)	-0.23 (-0.59, 0.13)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9438	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.2 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (83.9%)	24 (92.3%)
	Observed mean domain score at baseline	Mean (SD)	57.81 (23.16)	56.14 (16.86)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	38 (67.9%)	17 (65.4%)
	Observed mean domain score at Week 24	Mean (SD)	58.37 (20.80)	55.46 (19.08)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.28 (2.15)	2.95 (2.99)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.22 (3.68)	
		95% CI	-12.48, 2.03	
		p-value [b]	0.1572	
Effect size by Hedges' g (95% CI)		-0.24 (-0.73, 0.25)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.2 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Male	Number of HRQoL evaluable subjects	N	93	50
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	89 (95.7%)	50 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.97 (20.04)	55.10 (20.96)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	66 (71.0%)	34 (68.0%)
	Observed mean domain score at Week 24	Mean (SD)	58.10 (22.48)	59.89 (23.69)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.83 (1.89)	1.35 (2.56)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.18 (3.18)	
		95% CI	-9.44, 3.08	
		p-value [b]	0.3176	
		Effect size by Hedges' g (95% CI)	-0.16 (-0.50, 0.19)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3462	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.2 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Female	Number of HRQoL evaluable subjects	N	56	26
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	55 (98.2%)	25 (96.2%)
	Observed mean domain score at baseline	Mean (SD)	54.49 (22.37)	49.99 (18.33)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	41 (73.2%)	19 (73.1%)
	Observed mean domain score at Week 24	Mean (SD)	54.37 (22.37)	54.75 (18.14)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.14 (2.23)	3.05 (3.23)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.18 (3.92)	
		95% CI	-14.91, 0.54	
		p-value [b]	0.0683	
Effect size by Hedges' g (95% CI)		-0.34 (-0.81, 0.14)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.4 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects	N	31	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.66 (17.58)	56.58 (19.26)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (93.5%)	15 (78.9%)
	Observed mean domain score at Week 24	Mean (SD)	64.22 (21.84)	53.33 (17.34)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.53 (2.79)	-0.16 (3.72)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	3.69 (4.65)	
		95% CI	-5.46, 12.84	
		p-value [b]	0.4280	
		Effect size by Hedges' g (95% CI)	0.20 (-0.36, 0.76)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0146	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.4 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	115 (97.5%)	57 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.61 (21.36)	49.78 (23.92)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	81 (68.6%)	38 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	50.77 (22.04)	54.61 (26.22)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.42 (1.93)	2.21 (2.70)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.64 (3.32)	
		95% CI	-14.17, -1.10	
		p-value [b]	0.0222	
	Effect size by Hedges' g (95% CI)	-0.34 (-0.66, -0.02)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.4 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects	N	31	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	71.68 (32.94)	51.46 (31.91)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (93.5%)	15 (78.9%)
	Observed mean domain score at Week 24	Mean (SD)	57.47 (30.58)	56.30 (34.49)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.37 (4.63)	5.79 (6.04)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.42 (7.63)	
		95% CI	-20.45, 9.61	
		p-value [b]	0.4782	
		Effect size by Hedges' g (95% CI)	-0.16 (-0.73, 0.40)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9275	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.4 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	115 (97.5%)	57 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.72 (32.62)	50.88 (29.02)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	79 (66.9%)	38 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	53.02 (33.57)	62.87 (31.88)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.31 (2.92)	0.85 (4.09)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.15 (5.02)	
		95% CI	-16.04, 3.74	
		p-value [b]	0.2216	
	Effect size by Hedges' g (95% CI)	-0.19 (-0.51, 0.12)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.4 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects	N	31	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.87 (37.94)	43.42 (38.95)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (87.1%)	14 (73.7%)
	Observed mean domain score at Week 24	Mean (SD)	58.33 (32.52)	51.79 (39.79)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.72 (4.08)	5.84 (5.41)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.56 (6.78)	
		95% CI	-21.93, 4.80	
		p-value [b]	0.2079	
		Effect size by Hedges' g (95% CI)	-0.22 (-0.78, 0.34)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9072	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.4 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	106 (89.8%)	49 (86.0%)
	Observed mean domain score at baseline	Mean (SD)	45.99 (37.34)	48.98 (37.49)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	75 (63.6%)	34 (59.6%)
	Observed mean domain score at Week 24	Mean (SD)	44.33 (35.01)	54.41 (37.17)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.99 (2.78)	0.74 (4.05)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.73 (4.92)	
		95% CI	-17.42, 1.96	
		p-value [b]	0.1173	
	Effect size by Hedges' g (95% CI)	-0.21 (-0.54, 0.13)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.4 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects	N	31	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	64.25 (36.84)	58.77 (43.87)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	28 (90.3%)	15 (78.9%)
	Observed mean domain score at Week 24	Mean (SD)	76.19 (31.97)	62.78 (41.77)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.46 (4.24)	4.30 (5.58)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.15 (7.01)	
		95% CI	-11.65, 15.96	
		p-value [b]	0.7588	
		Effect size by Hedges' g (95% CI)	0.05 (-0.51, 0.62)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7032	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.4 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	110 (93.2%)	52 (91.2%)
	Observed mean domain score at baseline	Mean (SD)	61.52 (34.61)	64.10 (37.69)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	77 (65.3%)	36 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	62.12 (35.52)	61.57 (35.03)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.48 (2.79)	0.17 (3.98)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.65 (4.86)	
		95% CI	-10.23, 8.93	
		p-value [b]	0.8936	
	Effect size by Hedges' g (95% CI)	-0.02 (-0.35, 0.31)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.4 (AMNOG)  
QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects	N	31	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	81.41 (11.29)	76.44 (11.50)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (93.5%)	15 (78.9%)
	Observed mean domain score at Week 24	Mean (SD)	76.52 (15.58)	80.32 (15.36)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.96 (1.86)	3.55 (2.47)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.51 (3.08)	
		95% CI	-11.58, 0.57	
		p-value [b]	0.0754	
		Effect size by Hedges' g (95% CI)	-0.48 (-1.05, 0.09)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7890	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.4 (AMNOG)  
QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	118 (100.0%)	57 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	74.01 (14.78)	72.77 (14.97)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	81 (68.6%)	38 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	72.19 (14.35)	76.44 (14.18)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.27 (1.24)	1.40 (1.76)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.68 (2.15)	
		95% CI	-8.91, -0.44	
		p-value [b]	0.0307	
		Effect size by Hedges' g (95% CI)	-0.31 (-0.63, 0.00)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.4 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
North America	Number of HRQoL evaluable subjects	N	31	19	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)	
	Observed mean domain score at baseline	Mean (SD)	61.60 (22.78)	51.88 (21.23)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (93.5%)	15 (78.9%)	
	Observed mean domain score at Week 24	Mean (SD)	66.26 (22.02)	56.51 (25.60)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	2.47 (2.91)	3.15 (3.86)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.69 (4.84)		
		95% CI	-10.22, 8.85		
		p-value [b]	0.8876		
		Effect size by Hedges' g (95% CI)	-0.03 (-0.59, 0.53)		
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7365		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.4 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	117 (99.2%)	56 (98.2%)
	Observed mean domain score at baseline	Mean (SD)	55.41 (24.49)	59.31 (23.52)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	80 (67.8%)	38 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	58.04 (25.99)	59.84 (24.35)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.02 (1.86)	1.38 (2.64)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.40 (3.23)	
		95% CI	-8.76, 3.96	
		p-value [b]	0.4578	
	Effect size by Hedges' g (95% CI)	-0.10 (-0.42, 0.22)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.4 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
North America	Number of HRQoL evaluable subjects	N	31	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	66.78 (18.30)	57.34 (21.08)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (87.1%)	14 (73.7%)
	Observed mean domain score at Week 24	Mean (SD)	66.99 (17.74)	62.39 (21.85)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.28 (2.37)	4.07 (3.10)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.35 (3.91)	
		95% CI	-12.05, 3.36	
		p-value [b]	0.2670	
		Effect size by Hedges' g (95% CI)	-0.22 (-0.78, 0.34)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5808	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.4 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	104 (88.1%)	49 (86.0%)
	Observed mean domain score at baseline	Mean (SD)	57.11 (21.63)	57.40 (20.89)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	73 (61.9%)	34 (59.6%)
	Observed mean domain score at Week 24	Mean (SD)	56.37 (20.60)	62.34 (21.28)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.78 (1.58)	1.88 (2.26)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.67 (2.76)	
		95% CI	-12.10, -1.23	
		p-value [b]	0.0165	
	Effect size by Hedges' g (95% CI)	-0.31 (-0.65, 0.03)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.4 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
North America	Number of HRQoL evaluable subjects	N	31	19	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)	
	Observed mean domain score at baseline	Mean (SD)	64.19 (18.54)	54.61 (20.44)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (87.1%)	14 (73.7%)	
	Observed mean domain score at Week 24	Mean (SD)	67.18 (18.28)	59.42 (22.94)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	1.18 (2.45)	4.42 (3.23)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.24 (4.06)		
		95% CI	-11.25, 4.77		
		p-value [b]	0.4256		
		Effect size by Hedges' g (95% CI)	-0.17 (-0.73, 0.40)		
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5900		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.4 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	104 (88.1%)	48 (84.2%)
	Observed mean domain score at baseline	Mean (SD)	56.36 (21.72)	58.95 (21.45)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	73 (61.9%)	34 (59.6%)
	Observed mean domain score at Week 24	Mean (SD)	56.76 (22.25)	62.12 (21.54)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.25 (1.66)	2.30 (2.40)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.54 (2.92)	
		95% CI	-11.31, 0.22	
		p-value [b]	0.0592	
	Effect size by Hedges' g (95% CI)	-0.25 (-0.60, 0.09)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.4 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
North America	Number of HRQoL evaluable subjects	N	31	19	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	19 (100.0%)	
	Observed mean domain score at baseline	Mean (SD)	63.65 (18.74)	53.31 (18.50)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (93.5%)	15 (78.9%)	
	Observed mean domain score at Week 24	Mean (SD)	62.65 (19.72)	55.38 (20.51)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	0.80 (2.73)	3.05 (3.58)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.25 (4.51)		
		95% CI	-11.13, 6.64		
		p-value [b]	0.6188		
		Effect size by Hedges' g (95% CI)	-0.12 (-0.68, 0.44)		
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5166		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.4 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Europe, the Middle East and Africa	Number of HRQoL evaluable subjects	N	118	57
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	113 (95.8%)	56 (98.2%)
	Observed mean domain score at baseline	Mean (SD)	53.14 (20.96)	53.43 (20.83)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	78 (66.1%)	38 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	54.45 (23.05)	59.10 (22.49)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.78 (1.76)	1.53 (2.45)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.31 (3.02)	
		95% CI	-11.26, 0.64	
		p-value [b]	0.0799	
	Effect size by Hedges' g (95% CI)	-0.25 (-0.57, 0.07)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.5 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	138 (97.9%)	74 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.08 (20.64)	51.01 (22.45)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	104 (73.8%)	51 (68.9%)
	Observed mean domain score at Week 24	Mean (SD)	53.73 (22.16)	54.17 (24.07)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.55 (1.90)	1.62 (2.63)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.17 (3.25)	
		95% CI	-11.58, 1.23	
		p-value [b]	0.1125	
		Effect size by Hedges' g (95% CI)	-0.24 (-0.52, 0.04)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5393	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.5 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	64.29 (20.95)	68.75 (44.19)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	70.00 (31.37)	56.25 (26.52)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.42 (5.45)	1.72 (9.91)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	1.70 (11.28)	
		95% CI	-20.54, 23.94	
		p-value [b]	0.8805	
	Effect size by Hedges' g (95% CI)	0.06 (-1.34, 1.46)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.5 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	138 (97.9%)	74 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.99 (33.49)	50.75 (29.20)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	103 (73.0%)	51 (68.9%)
	Observed mean domain score at Week 24	Mean (SD)	52.86 (32.65)	61.66 (32.61)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.27 (2.72)	1.81 (3.77)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.08 (4.65)	
		95% CI	-15.26, 3.09	
		p-value [b]	0.1925	
Effect size by Hedges' g (95% CI)		-0.19 (-0.47, 0.09)		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9359		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.5 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	65.08 (37.09)	61.11 (55.00)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	82.22 (20.18)	44.44 (31.43)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.32 (9.24)	8.96 (16.98)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.65 (19.33)	
		95% CI	-45.74, 30.45	
		p-value [b]	0.6928	
	Effect size by Hedges' g (95% CI)	-0.17 (-1.57, 1.23)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.5 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	129 (91.5%)	66 (89.2%)
	Observed mean domain score at baseline	Mean (SD)	47.87 (37.63)	47.35 (37.24)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	96 (68.1%)	46 (62.2%)
	Observed mean domain score at Week 24	Mean (SD)	46.88 (34.46)	54.89 (37.50)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.64 (2.58)	2.56 (3.70)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.19 (4.52)	
		95% CI	-17.10, 0.72	
		p-value [b]	0.0712	
		Effect size by Hedges' g (95% CI)	-0.22 (-0.51, 0.08)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6138	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.5 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	67.86 (40.09)	50.00 (70.71)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	75.00 (35.36)	25.00 (35.36)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.21 (8.02)	-6.48 (14.68)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.27 (16.72)	
		95% CI	-32.73, 33.27	
		p-value [b]	0.9872	
	Effect size by Hedges' g (95% CI)	0.01 (-1.39, 1.40)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.5 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	133 (94.3%)	69 (93.2%)	
	Observed mean domain score at baseline	Mean (SD)	60.59 (35.13)	62.08 (39.49)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	99 (70.2%)	49 (66.2%)	
	Observed mean domain score at Week 24	Mean (SD)	64.98 (35.51)	62.07 (37.31)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	1.33 (2.66)	1.84 (3.72)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.51 (4.57)		
		95% CI	-9.52, 8.49		
		p-value [b]	0.9108		
		Effect size by Hedges' g (95% CI)	-0.01 (-0.30, 0.28)		
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4480		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.5 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	85.71 (21.90)	83.33 (23.57)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	76.67 (23.12)	58.33 (23.57)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.42 (8.97)	-15.43 (15.39)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	13.01 (17.77)	
		95% CI	-22.04, 48.06	
		p-value [b]	0.4651	
	Effect size by Hedges' g (95% CI)	0.52 (-0.90, 1.94)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.5 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	141 (100.0%)	74 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.18 (14.54)	73.68 (14.26)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	104 (73.8%)	51 (68.9%)
	Observed mean domain score at Week 24	Mean (SD)	72.89 (14.84)	78.06 (13.87)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.09 (1.18)	2.06 (1.65)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.15 (2.03)	
		95% CI	-9.16, -1.15	
		p-value [b]	0.0119	
		Effect size by Hedges' g (95% CI)	-0.36 (-0.64, -0.07)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4452	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.5 (AMNOG)  
QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	83.67 (10.23)	73.81 (16.84)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	83.81 (9.87)	64.29 (30.30)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.63 (3.61)	-2.17 (6.55)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.54 (7.48)	
		95% CI	-14.21, 15.29	
		p-value [b]	0.9425	
	Effect size by Hedges' g (95% CI)	0.04 (-1.35, 1.44)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.5 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	140 (99.3%)	73 (98.6%)
	Observed mean domain score at baseline	Mean (SD)	56.00 (24.27)	57.47 (22.83)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	103 (73.0%)	51 (68.9%)
	Observed mean domain score at Week 24	Mean (SD)	59.52 (25.10)	59.43 (23.48)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.13 (1.76)	1.85 (2.47)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.98 (3.03)	
		95% CI	-7.97, 4.00	
		p-value [b]	0.5145	
		Effect size by Hedges' g (95% CI)	-0.08 (-0.37, 0.20)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8570	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.5 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	70.41 (21.60)	55.95 (42.09)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	75.71 (26.75)	45.24 (57.24)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.56 (5.85)	1.63 (10.73)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.19 (12.23)	
		95% CI	-28.30, 19.92	
		p-value [b]	0.7321	
	Effect size by Hedges' g (95% CI)	-0.15 (-1.54, 1.25)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.5 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	127 (90.1%)	66 (89.2%)
	Observed mean domain score at baseline	Mean (SD)	58.56 (21.11)	57.08 (20.37)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	95 (67.4%)	46 (62.2%)
	Observed mean domain score at Week 24	Mean (SD)	58.27 (20.29)	62.90 (21.06)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.64 (1.46)	2.79 (2.06)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.43 (2.53)	
		95% CI	-11.42, -1.44	
		p-value [b]	0.0118	
		Effect size by Hedges' g (95% CI)	-0.31 (-0.60, -0.01)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4504	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.5 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	73.32 (21.90)	67.40 (42.06)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	77.54 (11.05)	49.66 (29.43)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.82 (4.75)	-4.88 (8.66)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	1.06 (9.85)	
		95% CI	-18.38, 20.50	
		p-value [b]	0.9144	
	Effect size by Hedges' g (95% CI)	0.04 (-1.36, 1.43)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.5 (AMNOG)  
QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	127 (90.1%)	65 (87.8%)
	Observed mean domain score at baseline	Mean (SD)	57.39 (21.13)	57.60 (20.76)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	95 (67.4%)	46 (62.2%)
	Observed mean domain score at Week 24	Mean (SD)	58.67 (21.53)	61.94 (21.05)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.09 (1.57)	3.05 (2.22)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.14 (2.72)	
		95% CI	-10.51, 0.24	
		p-value [b]	0.0608	
		Effect size by Hedges' g (95% CI)	-0.24 (-0.54, 0.06)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8702	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.5 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	71.87 (21.32)	61.68 (42.08)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	76.63 (18.70)	47.45 (43.34)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.21 (4.83)	-0.72 (8.79)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.49 (10.02)	
		95% CI	-23.27, 16.29	
		p-value [b]	0.7278	
	Effect size by Hedges' g (95% CI)	-0.12 (-1.52, 1.28)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.5 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RCMD	Number of HRQoL evaluable subjects	N	141	74
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	136 (96.5%)	73 (98.6%)
	Observed mean domain score at baseline	Mean (SD)	54.91 (20.77)	53.16 (19.60)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	102 (72.3%)	51 (68.9%)
	Observed mean domain score at Week 24	Mean (SD)	55.73 (22.27)	58.42 (21.49)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.88 (1.67)	1.86 (2.29)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.74 (2.84)	
		95% CI	-10.34, 0.86	
		p-value [b]	0.0964	
		Effect size by Hedges' g (95% CI)	-0.23 (-0.52, 0.05)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9668	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.5 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
MDS WHO classification at baseline MDS RARS	Number of HRQoL evaluable subjects	N	7	2
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	2 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	66.59 (22.92)	61.94 (47.09)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	2 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	75.98 (16.99)	48.64 (38.40)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.60 (5.43)	3.67 (9.95)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.27 (11.32)	
		95% CI	-26.58, 18.05	
		p-value [b]	0.7065	
	Effect size by Hedges' g (95% CI)	-0.14 (-1.54, 1.26)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.6 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	38 (97.4%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.99 (23.86)	49.34 (18.85)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	31 (79.5%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	50.40 (22.47)	52.08 (21.87)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.62 (2.75)	1.28 (3.91)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.91 (4.77)	
		95% CI	-14.30, 4.49	
		p-value [b]	0.3049	
		Effect size by Hedges' g (95% CI)	-0.22 (-0.76, 0.33)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9450	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.6 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects	N	60	34
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	60 (100.0%)	34 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.33 (19.24)	49.26 (23.42)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	45 (75.0%)	23 (67.6%)
	Observed mean domain score at Week 24	Mean (SD)	55.28 (21.23)	54.35 (22.80)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.11 (2.33)	2.16 (3.16)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.27 (3.93)	
		95% CI	-12.00, 3.47	
		p-value [b]	0.2781	
		Effect size by Hedges' g (95% CI)	-0.20 (-0.62, 0.22)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.6 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects	N	50	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	48 (96.0%)	23 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.25 (19.81)	56.52 (25.25)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	34 (68.0%)	18 (78.3%)
	Observed mean domain score at Week 24	Mean (SD)	56.62 (24.85)	55.56 (27.53)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.77 (2.58)	1.09 (3.57)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.86 (4.40)	
		95% CI	-14.51, 2.79	
		p-value [b]	0.1836	
		Effect size by Hedges' g (95% CI)	-0.27 (-0.76, 0.23)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.6 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	38 (97.4%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.88 (33.74)	51.46 (27.27)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (74.4%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	47.89 (32.95)	51.85 (35.56)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.40 (4.29)	6.12 (6.17)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-11.52 (7.52)	
		95% CI	-26.32, 3.28	
		p-value [b]	0.1266	
		Effect size by Hedges' g (95% CI)	-0.36 (-0.90, 0.19)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3588	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.6 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects	N	60	34
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	60 (100.0%)	34 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.22 (33.60)	42.81 (27.78)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	45 (75.0%)	23 (67.6%)
	Observed mean domain score at Week 24	Mean (SD)	59.75 (32.68)	63.77 (32.00)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.51 (3.57)	1.07 (4.83)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.56 (6.02)	
		95% CI	-12.41, 11.30	
		p-value [b]	0.9264	
		Effect size by Hedges' g (95% CI)	-0.02 (-0.43, 0.40)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.6 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluatable subjects	N	50	23
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	48 (96.0%)	23 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.02 (34.15)	62.80 (31.00)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	34 (68.0%)	18 (78.3%)
	Observed mean domain score at Week 24	Mean (SD)	52.29 (32.31)	63.58 (31.61)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-8.73 (4.00)	0.83 (5.56)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-9.57 (6.83)	
		95% CI	-23.02, 3.89	
		p-value [b]	0.1627	
		Effect size by Hedges' g (95% CI)	-0.29 (-0.78, 0.21)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).  
 [b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.  
 [c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Table 14.2.18.6.1.2.3.6 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	36 (92.3%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.78 (40.87)	52.63 (38.09)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	30 (76.9%)	11 (57.9%)
	Observed mean domain score at Week 24	Mean (SD)	47.50 (34.96)	52.27 (42.51)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.47 (3.94)	4.45 (5.59)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.92 (6.84)	
		95% CI	-21.40, 5.55	
		p-value [b]	0.2480	
		Effect size by Hedges' g (95% CI)	-0.20 (-0.74, 0.35)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9375	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.6 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects	N	60	34
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	55 (91.7%)	30 (88.2%)
	Observed mean domain score at baseline	Mean (SD)	45.45 (36.03)	44.17 (34.54)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	40 (66.7%)	22 (64.7%)
	Observed mean domain score at Week 24	Mean (SD)	49.38 (33.73)	52.27 (31.73)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.96 (3.39)	3.03 (4.66)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.99 (5.75)	
		95% CI	-18.33, 4.34	
		p-value [b]	0.2255	
		Effect size by Hedges' g (95% CI)	-0.20 (-0.64, 0.25)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.6 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects	N	50	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	46 (92.0%)	19 (82.6%)
	Observed mean domain score at baseline	Mean (SD)	50.00 (37.64)	47.37 (43.22)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	32 (64.0%)	15 (65.2%)
	Observed mean domain score at Week 24	Mean (SD)	46.88 (36.89)	56.67 (43.78)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-10.49 (3.66)	-0.81 (5.41)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-9.67 (6.53)	
		95% CI	-22.55, 3.21	
		p-value [b]	0.1403	
		Effect size by Hedges' g (95% CI)	-0.24 (-0.77, 0.29)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.6 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	36 (92.3%)	18 (94.7%)
	Observed mean domain score at baseline	Mean (SD)	65.05 (40.91)	80.56 (34.54)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (74.4%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	67.82 (38.62)	73.61 (34.97)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.66 (4.18)	0.59 (5.92)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	6.06 (7.23)	
		95% CI	-8.18, 20.30	
		p-value [b]	0.4025	
		Effect size by Hedges' g (95% CI)	0.15 (-0.40, 0.71)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4489	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.6 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluatable subjects	N	60	34
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	58 (96.7%)	31 (91.2%)
	Observed mean domain score at baseline	Mean (SD)	64.66 (32.08)	57.26 (40.41)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	44 (73.3%)	23 (67.6%)
	Observed mean domain score at Week 24	Mean (SD)	67.23 (35.56)	59.78 (37.32)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.99 (3.40)	1.32 (4.68)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.33 (5.78)	
		95% CI	-11.72, 11.05	
		p-value [b]	0.9543	
		Effect size by Hedges' g (95% CI)	-0.01 (-0.44, 0.42)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.6 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects	N	50	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (94.0%)	22 (95.7%)
	Observed mean domain score at baseline	Mean (SD)	56.74 (33.68)	55.68 (38.09)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	32 (64.0%)	16 (69.6%)
	Observed mean domain score at Week 24	Mean (SD)	62.24 (31.54)	56.25 (37.45)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.11 (3.80)	1.80 (5.32)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.92 (6.53)	
		95% CI	-17.77, 7.93	
		p-value [b]	0.4518	
		Effect size by Hedges' g (95% CI)	-0.14 (-0.64, 0.36)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

- [a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).
- [b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.
- [c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.6 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	39 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.70 (18.01)	73.43 (15.33)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	31 (79.5%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	72.04 (14.89)	77.78 (17.27)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		-2.22 (1.72)	4.57 (2.50)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-6.79 (3.04)	
	95% CI		-12.77, -0.81	
	p-value [b]		0.0261	
	Effect size by Hedges' g (95% CI)		-0.39 (-0.94, 0.16)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.7199	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.6 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects	N	60	34
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	60 (100.0%)	34 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.95 (12.69)	70.45 (14.58)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	45 (75.0%)	23 (67.6%)
	Observed mean domain score at Week 24	Mean (SD)	73.76 (15.38)	76.19 (13.39)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.60 (1.48)	2.71 (2.01)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.31 (2.50)	
		95% CI	-9.23, 0.61	
		p-value [b]	0.0857	
		Effect size by Hedges' g (95% CI)	-0.32 (-0.74, 0.10)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.6 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects	N	50	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	50 (100.0%)	23 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	74.95 (13.46)	78.67 (11.56)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	34 (68.0%)	18 (78.3%)
	Observed mean domain score at Week 24	Mean (SD)	73.95 (14.08)	79.10 (14.53)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.44 (1.61)	-1.05 (2.28)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.39 (2.79)	
		95% CI	-9.88, 1.11	
		p-value [b]	0.1172	
		Effect size by Hedges' g (95% CI)	-0.34 (-0.83, 0.16)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.6 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	38 (97.4%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.15 (24.47)	57.89 (22.14)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	30 (76.9%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	57.30 (21.69)	58.93 (26.12)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.39 (2.77)	2.65 (3.95)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.26 (4.82)	
		95% CI	-9.75, 9.23	
		p-value [b]	0.9572	
		Effect size by Hedges' g (95% CI)	-0.01 (-0.55, 0.53)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8878	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.6 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects	N	60	34
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	60 (100.0%)	33 (97.1%)
	Observed mean domain score at baseline	Mean (SD)	56.11 (23.63)	58.51 (23.39)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	45 (75.0%)	23 (67.6%)
	Observed mean domain score at Week 24	Mean (SD)	62.86 (24.65)	59.01 (22.14)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.20 (2.30)	2.84 (3.14)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.64 (3.89)	
		95% CI	-10.30, 5.02	
		p-value [b]	0.4980	
		Effect size by Hedges' g (95% CI)	-0.11 (-0.53, 0.31)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.6 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects	N	50	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	50 (100.0%)	23 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.33 (25.09)	55.49 (24.18)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	34 (68.0%)	18 (78.3%)
	Observed mean domain score at Week 24	Mean (SD)	59.31 (28.83)	58.73 (27.60)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.95 (2.52)	-0.17 (3.57)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.78 (4.37)	
		95% CI	-11.40, 5.83	
		p-value [b]	0.5252	
		Effect size by Hedges' g (95% CI)	-0.11 (-0.60, 0.38)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.6 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	35 (89.7%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	60.05 (23.07)	61.12 (19.73)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	28 (71.8%)	11 (57.9%)
	Observed mean domain score at Week 24	Mean (SD)	57.57 (20.60)	62.91 (25.09)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		-2.22 (2.34)	3.86 (3.19)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-6.08 (3.95)	
	95% CI		-13.88, 1.71	
	p-value [b]		0.1253	
	Effect size by Hedges' g (95% CI)		-0.27 (-0.83, 0.28)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.7459	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.6 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects	N	60	34
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	55 (91.7%)	30 (88.2%)
	Observed mean domain score at baseline	Mean (SD)	60.03 (20.06)	53.20 (20.82)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	40 (66.7%)	22 (64.7%)
	Observed mean domain score at Week 24	Mean (SD)	62.00 (19.66)	60.46 (19.67)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.55 (1.93)	3.00 (2.64)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.55 (3.27)	
		95% CI	-11.00, 1.90	
		p-value [b]	0.1658	
		Effect size by Hedges' g (95% CI)	-0.22 (-0.66, 0.22)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.6 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects	N	50	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	45 (90.0%)	19 (82.6%)
	Observed mean domain score at baseline	Mean (SD)	57.91 (21.59)	60.24 (21.64)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	32 (64.0%)	15 (65.2%)
	Observed mean domain score at Week 24	Mean (SD)	57.24 (21.18)	64.71 (21.66)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-7.31 (2.11)	0.63 (3.11)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.94 (3.76)	
		95% CI	-15.35, -0.53	
		p-value [b]	0.0358	
		Effect size by Hedges' g (95% CI)	-0.36 (-0.90, 0.17)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.6 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	35 (89.7%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	59.45 (21.83)	59.51 (19.98)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	28 (71.8%)	11 (57.9%)
	Observed mean domain score at Week 24	Mean (SD)	57.86 (19.66)	61.87 (24.41)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		-0.25 (2.44)	4.28 (3.30)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-4.53 (4.10)	
	95% CI		-12.61, 3.56	
	p-value [b]		0.2708	
	Effect size by Hedges' g (95% CI)		-0.21 (-0.76, 0.34)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.9402	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.6 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects	N	60	34
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	55 (91.7%)	29 (85.3%)
	Observed mean domain score at baseline	Mean (SD)	58.07 (20.26)	55.66 (21.65)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	40 (66.7%)	22 (64.7%)
	Observed mean domain score at Week 24	Mean (SD)	62.58 (21.28)	59.89 (19.57)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.72 (2.01)	3.71 (2.79)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.43 (3.44)	
		95% CI	-11.20, 2.35	
		p-value [b]	0.1995	
		Effect size by Hedges' g (95% CI)	-0.21 (-0.66, 0.24)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.6 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects	N	50	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	45 (90.0%)	19 (82.6%)
	Observed mean domain score at baseline	Mean (SD)	57.26 (22.31)	59.07 (22.19)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	32 (64.0%)	15 (65.2%)
	Observed mean domain score at Week 24	Mean (SD)	57.30 (23.95)	63.07 (24.10)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.38 (2.20)	0.52 (3.23)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.91 (3.91)	
		95% CI	-13.62, 1.80	
		p-value [b]	0.1325	
		Effect size by Hedges' g (95% CI)	-0.26 (-0.79, 0.27)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.6 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline ≤ 2 years	Number of HRQoL evaluable subjects	N	39	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	37 (94.9%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.35 (21.50)	52.90 (18.09)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	28 (71.8%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	53.18 (22.39)	54.29 (22.85)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.26 (2.60)	3.69 (3.64)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.96 (4.48)	
		95% CI	-14.77, 2.86	
		p-value [b]	0.1845	
		Effect size by Hedges' g (95% CI)	-0.29 (-0.84, 0.26)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8921	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.6 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline >2 to 5 years	Number of HRQoL evaluable subjects	N	60	34
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	60 (100.0%)	33 (97.1%)
	Observed mean domain score at baseline	Mean (SD)	55.56 (20.03)	50.28 (20.24)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	45 (75.0%)	23 (67.6%)
	Observed mean domain score at Week 24	Mean (SD)	59.30 (20.31)	59.04 (21.23)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.80 (2.12)	2.83 (2.89)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.64 (3.59)	
		95% CI	-10.71, 3.43	
		p-value [b]	0.3118	
		Effect size by Hedges' g (95% CI)	-0.18 (-0.60, 0.24)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.6 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Time since initial diagnosis at baseline > 5 years	Number of HRQoL evaluable subjects	N	50	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (94.0%)	23 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.04 (21.90)	58.27 (21.47)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	34 (68.0%)	18 (78.3%)
	Observed mean domain score at Week 24	Mean (SD)	56.07 (25.16)	59.29 (22.85)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.70 (2.40)	-0.61 (3.30)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.09 (4.07)	
		95% CI	-13.11, 2.93	
		p-value [b]	0.2124	
		Effect size by Hedges' g (95% CI)	-0.23 (-0.73, 0.26)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.7 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	138 (97.9%)	65 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.35 (20.80)	53.46 (21.25)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	106 (75.2%)	45 (69.2%)
	Observed mean domain score at Week 24	Mean (SD)	53.89 (22.52)	55.56 (23.16)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.37 (1.89)	1.91 (2.69)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.28 (3.29)	
		95% CI	-11.77, 1.20	
		p-value [b]	0.1096	
		Effect size by Hedges' g (95% CI)	-0.25 (-0.55, 0.04)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6688	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.7 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	42.19 (14.85)	39.77 (29.48)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	65.63 (27.72)	46.88 (28.15)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.47 (5.45)	-0.18 (4.70)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.28 (7.16)	
		95% CI	-16.37, 11.80	
		p-value [b]	0.7499	
Effect size by Hedges' g (95% CI)		-0.09 (-0.96, 0.78)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.7 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	138 (97.9%)	65 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.64 (33.82)	52.48 (28.63)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	104 (73.8%)	45 (69.2%)
	Observed mean domain score at Week 24	Mean (SD)	54.27 (33.00)	61.23 (31.75)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.36 (2.72)	3.50 (3.90)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.86 (4.76)	
		95% CI	-17.25, 1.52	
		p-value [b]	0.1001	
		Effect size by Hedges' g (95% CI)	-0.24 (-0.54, 0.05)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1449	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.7 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	47.22 (32.39)	42.42 (34.72)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	52.78 (27.78)	59.72 (38.46)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.51 (9.09)	-6.16 (7.66)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	9.67 (11.88)	
		95% CI	-13.72, 33.06	
		p-value [b]	0.4163	
		Effect size by Hedges' g (95% CI)	0.27 (-0.60, 1.15)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.7 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	129 (91.5%)	57 (87.7%)
	Observed mean domain score at baseline	Mean (SD)	49.03 (37.55)	50.44 (37.35)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	98 (69.5%)	41 (63.1%)
	Observed mean domain score at Week 24	Mean (SD)	47.19 (34.97)	56.10 (37.41)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.62 (2.56)	4.53 (3.82)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.15 (4.60)	
		95% CI	-19.23, -1.08	
		p-value [b]	0.0286	
		Effect size by Hedges' g (95% CI)	-0.27 (-0.58, 0.04)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3711	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.7 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	46.88 (43.17)	31.82 (37.23)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	7 (63.6%)
	Observed mean domain score at Week 24	Mean (SD)	68.75 (23.94)	39.29 (37.80)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-10.47 (8.07)	-9.76 (6.77)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.71 (10.53)	
		95% CI	-21.46, 20.05	
		p-value [b]	0.9467	
Effect size by Hedges' g (95% CI)		-0.02 (-0.89, 0.85)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.7 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	133 (94.3%)	60 (92.3%)
	Observed mean domain score at baseline	Mean (SD)	61.84 (35.25)	62.08 (39.19)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	101 (71.6%)	43 (66.2%)
	Observed mean domain score at Week 24	Mean (SD)	64.52 (35.03)	62.60 (35.88)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.94 (2.65)	0.83 (3.84)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.11 (4.66)	
		95% CI	-9.08, 9.30	
		p-value [b]	0.9815	
		Effect size by Hedges' g (95% CI)	0.00 (-0.30, 0.31)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9180	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.7 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	66.67 (32.12)	65.91 (40.90)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	100.00 (0.00)	58.33 (43.42)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	5.08 (8.53)	3.83 (7.01)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	1.25 (11.04)	
		95% CI	-20.49, 22.99	
		p-value [b]	0.9097	
Effect size by Hedges' g (95% CI)		0.03 (-0.84, 0.90)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.7 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	141 (100.0%)	65 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.99 (14.31)	75.75 (12.70)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	106 (75.2%)	45 (69.2%)
	Observed mean domain score at Week 24	Mean (SD)	72.82 (14.71)	77.14 (14.16)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.19 (1.17)	1.62 (1.69)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.81 (2.05)	
		95% CI	-8.87, -0.76	
		p-value [b]	0.0202	
		Effect size by Hedges' g (95% CI)	-0.35 (-0.64, -0.05)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6423	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.7 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	67.86 (15.00)	61.47 (16.96)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	86.90 (8.13)	79.76 (17.03)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.36 (3.56)	4.02 (3.09)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.66 (4.68)	
		95% CI	-11.87, 6.55	
		p-value [b]	0.5698	
	Effect size by Hedges' g (95% CI)	-0.16 (-1.03, 0.71)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.7 (AMNOG)  
 QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	140 (99.3%)	64 (98.5%)
	Observed mean domain score at baseline	Mean (SD)	56.82 (23.94)	59.45 (22.37)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	105 (74.5%)	45 (69.2%)
	Observed mean domain score at Week 24	Mean (SD)	59.14 (24.92)	59.37 (23.98)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.56 (1.74)	1.41 (2.53)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.97 (3.07)	
		95% CI	-8.03, 4.08	
		p-value [b]	0.5212	
		Effect size by Hedges' g (95% CI)	-0.08 (-0.38, 0.21)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5842		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.7 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.76 (30.12)	45.67 (24.51)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	88.69 (12.20)	56.25 (28.96)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.52 (5.70)	4.35 (4.86)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.16 (7.49)	
		95% CI	-12.59, 16.91	
		p-value [b]	0.7731	
		Effect size by Hedges' g (95% CI)	0.08 (-0.79, 0.95)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.8 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	84 (100.0%)	42 (97.7%)
	Observed mean domain score at baseline	Mean (SD)	59.81 (23.38)	58.90 (22.73)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	66 (78.6%)	31 (72.1%)
	Observed mean domain score at Week 24	Mean (SD)	65.62 (22.46)	64.29 (20.42)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.13 (1.99)	2.78 (2.84)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.65 (3.46)	
		95% CI	-7.47, 6.18	
		p-value [b]	0.8521	
		Effect size by Hedges' g (95% CI)	-0.03 (-0.40, 0.34)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3892	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.8 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	64 (98.5%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.64 (24.82)	55.56 (23.67)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (66.2%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	51.94 (27.05)	51.30 (28.08)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.79 (2.31)	0.60 (3.13)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.39 (3.89)	
		95% CI	-12.04, 3.26	
		p-value [b]	0.2599	
	Effect size by Hedges' g (95% CI)	-0.18 (-0.60, 0.24)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.7 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	127 (90.1%)	57 (87.7%)
	Observed mean domain score at baseline	Mean (SD)	59.66 (21.26)	59.14 (20.17)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	96 (68.1%)	41 (63.1%)
	Observed mean domain score at Week 24	Mean (SD)	58.59 (20.46)	62.83 (20.87)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.75 (1.46)	3.05 (2.15)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.80 (2.60)	
		95% CI	-11.92, -1.67	
		p-value [b]	0.0096	
		Effect size by Hedges' g (95% CI)	-0.32 (-0.64, -0.01)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3743	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.7 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.16 (21.77)	48.28 (22.53)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	7 (63.6%)
	Observed mean domain score at Week 24	Mean (SD)	74.81 (7.44)	59.54 (24.72)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.41 (4.72)	-0.18 (3.98)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.24 (6.17)	
		95% CI	-13.40, 10.92	
		p-value [b]	0.8412	
Effect size by Hedges' g (95% CI)		-0.05 (-0.92, 0.82)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.7 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	127 (90.1%)	56 (86.2%)
	Observed mean domain score at baseline	Mean (SD)	58.39 (21.08)	59.83 (20.29)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	96 (68.1%)	41 (63.1%)
	Observed mean domain score at Week 24	Mean (SD)	58.65 (21.56)	61.93 (20.94)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.31 (1.56)	2.84 (2.30)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.15 (2.78)	
		95% CI	-10.63, 0.33	
		p-value [b]	0.0654	
Effect size by Hedges' g (95% CI)		-0.25 (-0.56, 0.07)		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6301		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.7 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)	
	Observed mean domain score at baseline	Mean (SD)	54.46 (24.71)	46.98 (22.88)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	7 (63.6%)	
	Observed mean domain score at Week 24	Mean (SD)	81.75 (9.70)	57.83 (27.64)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	1.28 (4.78)	3.39 (4.08)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.11 (6.27)		
		95% CI	-14.48, 10.25		
		p-value [b]	0.7368		
Effect size by Hedges' g (95% CI)		-0.09 (-0.96, 0.79)			

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.7 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 0-1	Number of HRQoL evaluable subjects	N	141	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	136 (96.5%)	64 (98.5%)
	Observed mean domain score at baseline	Mean (SD)	55.84 (20.85)	55.25 (18.92)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	103 (73.0%)	45 (69.2%)
	Observed mean domain score at Week 24	Mean (SD)	56.19 (22.66)	58.72 (20.77)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.02 (1.66)	2.25 (2.36)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.27 (2.88)	
		95% CI	-10.96, 0.42	
		p-value [b]	0.0691	
		Effect size by Hedges' g (95% CI)	-0.26 (-0.56, 0.04)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2354	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.7 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
ECOG at baseline 2	Number of HRQoL evaluable subjects	N	8	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	48.06 (21.66)	42.62 (24.44)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	4 (50.0%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	69.03 (8.66)	54.28 (28.41)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.11 (5.30)	0.07 (4.53)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	3.04 (6.95)	
		95% CI	-10.64, 16.72	
		p-value [b]	0.6620	
	Effect size by Hedges' g (95% CI)	0.12 (-0.75, 1.00)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.8 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	84 (100.0%)	43 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.06 (20.18)	53.20 (24.71)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	67 (79.8%)	31 (72.1%)
	Observed mean domain score at Week 24	Mean (SD)	56.16 (22.86)	60.08 (22.92)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.14 (2.08)	3.57 (2.94)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.70 (3.60)	
		95% CI	-13.79, 0.39	
		p-value [b]	0.0637	
		Effect size by Hedges' g (95% CI)	-0.31 (-0.67, 0.06)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3063	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.8 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	62 (95.4%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	51.81 (21.33)	49.24 (20.48)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (66.2%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	51.45 (22.36)	46.02 (23.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.62 (2.40)	-1.06 (3.21)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.56 (4.01)	
		95% CI	-10.44, 5.33	
		p-value [b]	0.5238	
	Effect size by Hedges' g (95% CI)	-0.12 (-0.54, 0.30)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.8 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	83 (98.8%)	43 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	63.86 (33.69)	49.61 (28.83)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	66 (78.6%)	31 (72.1%)
	Observed mean domain score at Week 24	Mean (SD)	58.92 (31.79)	69.18 (30.32)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.16 (3.12)	6.93 (4.36)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-11.09 (5.38)	
		95% CI	-21.70, -0.49	
		p-value [b]	0.0404	
		Effect size by Hedges' g (95% CI)	-0.34 (-0.71, 0.03)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0910	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.8 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	63 (96.9%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	43.74 (30.32)	52.86 (30.81)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	42 (64.6%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	46.83 (33.16)	49.49 (32.48)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.56 (3.68)	-4.36 (4.90)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.79 (6.12)	
		95% CI	-11.25, 12.84	
		p-value [b]	0.8968	
	Effect size by Hedges' g (95% CI)	0.03 (-0.39, 0.44)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.8 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	77 (91.7%)	39 (90.7%)
	Observed mean domain score at baseline	Mean (SD)	53.90 (36.51)	52.56 (37.52)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	62 (73.8%)	30 (69.8%)
	Observed mean domain score at Week 24	Mean (SD)	53.63 (32.59)	60.00 (35.72)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.52 (2.91)	4.80 (4.19)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-9.32 (5.10)	
		95% CI	-19.37, 0.73	
		p-value [b]	0.0689	
		Effect size by Hedges' g (95% CI)	-0.25 (-0.64, 0.13)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6535	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.8 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	60 (92.3%)	29 (87.9%)
	Observed mean domain score at baseline	Mean (SD)	42.50 (38.59)	40.52 (37.45)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	40 (61.5%)	18 (54.5%)
	Observed mean domain score at Week 24	Mean (SD)	39.38 (36.64)	43.06 (39.11)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-8.04 (3.41)	-1.57 (4.79)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.47 (5.87)	
		95% CI	-18.02, 5.08	
		p-value [b]	0.2708	
	Effect size by Hedges' g (95% CI)	-0.17 (-0.61, 0.27)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.8 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	80 (95.2%)	40 (93.0%)
	Observed mean domain score at baseline	Mean (SD)	66.98 (34.05)	65.83 (38.16)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	63 (75.0%)	31 (72.1%)
	Observed mean domain score at Week 24	Mean (SD)	74.74 (31.50)	68.55 (34.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.26 (2.96)	5.71 (4.22)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.45 (5.15)	
		95% CI	-12.60, 7.69	
		p-value [b]	0.6342	
		Effect size by Hedges' g (95% CI)	-0.07 (-0.45, 0.31)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4130	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.8 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluatable subjects	N	65	33
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	61 (93.8%)	31 (93.9%)
	Observed mean domain score at baseline	Mean (SD)	55.74 (35.47)	58.60 (40.74)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	42 (64.6%)	20 (60.6%)
	Observed mean domain score at Week 24	Mean (SD)	52.58 (36.18)	51.67 (38.86)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.87 (3.46)	-4.71 (4.70)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.83 (5.83)	
		95% CI	-8.63, 14.30	
		p-value [b]	0.6269	
		Effect size by Hedges' g (95% CI)	0.08 (-0.35, 0.50)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.8 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	84 (100.0%)	43 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	77.04 (13.77)	73.86 (14.49)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	67 (79.8%)	31 (72.1%)
	Observed mean domain score at Week 24	Mean (SD)	74.56 (14.79)	82.33 (11.15)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.32 (1.30)	3.72 (1.85)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.05 (2.27)	
		95% CI	-10.51, -1.58	
		p-value [b]	0.0082	
		Effect size by Hedges' g (95% CI)	-0.43 (-0.80, -0.06)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3226	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.8 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	65 (100.0%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	73.63 (15.08)	73.45 (14.04)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (66.2%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	71.43 (14.62)	70.78 (16.12)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.93 (1.49)	-0.51 (2.04)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.42 (2.53)	
		95% CI	-8.40, 1.55	
		p-value [b]	0.1761	
	Effect size by Hedges' g (95% CI)	-0.23 (-0.65, 0.19)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.8 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	77 (91.7%)	39 (90.7%)
	Observed mean domain score at baseline	Mean (SD)	63.91 (20.72)	59.75 (20.71)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	61 (72.6%)	30 (69.8%)
	Observed mean domain score at Week 24	Mean (SD)	63.99 (18.74)	67.61 (17.80)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.55 (1.67)	5.29 (2.37)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.85 (2.89)	
		95% CI	-14.55, -3.14	
		p-value [b]	0.0026	
		Effect size by Hedges' g (95% CI)	-0.42 (-0.81, -0.04)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0728	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.8 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	58 (89.2%)	29 (87.9%)
	Observed mean domain score at baseline	Mean (SD)	53.25 (20.55)	54.20 (20.82)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	39 (60.0%)	18 (54.5%)
	Observed mean domain score at Week 24	Mean (SD)	51.80 (20.74)	53.59 (23.96)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.65 (2.00)	-1.57 (2.72)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.08 (3.36)	
		95% CI	-8.70, 4.55	
		p-value [b]	0.5371	
	Effect size by Hedges' g (95% CI)	-0.10 (-0.54, 0.34)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.8 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	77 (91.7%)	38 (88.4%)
	Observed mean domain score at baseline	Mean (SD)	62.21 (20.53)	59.28 (21.35)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	61 (72.6%)	30 (69.8%)
	Observed mean domain score at Week 24	Mean (SD)	64.80 (19.33)	66.15 (17.51)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.04 (1.76)	4.86 (2.53)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.90 (3.08)	
		95% CI	-11.97, 0.17	
		p-value [b]	0.0567	
		Effect size by Hedges' g (95% CI)	-0.28 (-0.67, 0.11)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6139	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.8 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	58 (89.2%)	29 (87.9%)
	Observed mean domain score at baseline	Mean (SD)	52.78 (21.11)	55.67 (20.98)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	39 (60.0%)	18 (54.5%)
	Observed mean domain score at Week 24	Mean (SD)	51.39 (22.83)	53.32 (25.99)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.78 (2.08)	0.19 (2.83)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.97 (3.51)	
		95% CI	-10.87, 2.94	
		p-value [b]	0.2587	
	Effect size by Hedges' g (95% CI)	-0.19 (-0.63, 0.26)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.8 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	84	43
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	83 (98.8%)	42 (97.7%)
	Observed mean domain score at baseline	Mean (SD)	59.71 (21.20)	54.06 (20.23)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	65 (77.4%)	31 (72.1%)
	Observed mean domain score at Week 24	Mean (SD)	60.78 (21.31)	64.51 (18.05)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.51 (1.88)	4.50 (2.65)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.01 (3.25)	
		95% CI	-13.42, -0.60	
		p-value [b]	0.0322	
		Effect size by Hedges' g (95% CI)	-0.33 (-0.70, 0.04)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1788	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.8 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluatable subjects	N	65	33
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	61 (93.8%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	49.54 (19.13)	52.55 (20.31)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	42 (64.6%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	50.31 (22.83)	48.94 (23.76)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.97 (2.22)	-1.47 (2.92)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.50 (3.66)	
		95% CI	-8.70, 5.71	
		p-value [b]	0.6825	
	Effect size by Hedges' g (95% CI)	-0.08 (-0.50, 0.34)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.9 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	44 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.52 (18.24)	56.88 (23.11)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	37 (84.1%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	58.78 (24.45)	54.17 (19.29)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.86 (2.55)	3.37 (3.75)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.22 (4.52)	
		95% CI	-14.13, 3.68	
		p-value [b]	0.2492	
		Effect size by Hedges' g (95% CI)	-0.26 (-0.78, 0.26)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4960	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.9 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	40 (100.0%)	23 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	51.25 (21.71)	50.00 (26.11)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	30 (75.0%)	16 (69.6%)
	Observed mean domain score at Week 24	Mean (SD)	52.92 (20.68)	65.63 (25.21)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.53 (2.63)	3.73 (3.57)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.27 (4.43)	
		95% CI	-16.99, 0.45	
		p-value [b]	0.0630	
	Effect size by Hedges' g (95% CI)	-0.35 (-0.86, 0.16)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.9 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	62 (95.4%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	51.81 (21.33)	49.24 (20.48)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (66.2%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	51.45 (22.36)	46.02 (23.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.65 (2.41)	-1.08 (3.22)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.57 (4.01)	
		95% CI	-10.47, 5.33	
		p-value [b]	0.5221	
	Effect size by Hedges' g (95% CI)	-0.12 (-0.54, 0.30)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.9 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	44 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	69.44 (31.44)	49.44 (29.83)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	37 (84.1%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	65.77 (31.79)	62.96 (33.51)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.84 (3.99)	5.59 (5.86)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.43 (7.11)	
		95% CI	-21.44, 6.58	
		p-value [b]	0.2969	
		Effect size by Hedges' g (95% CI)	-0.24 (-0.76, 0.29)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1691	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.9 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	39 (97.5%)	23 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.55 (35.40)	49.76 (28.60)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (72.5%)	16 (69.6%)
	Observed mean domain score at Week 24	Mean (SD)	50.19 (30.08)	75.00 (26.76)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.72 (4.18)	8.11 (5.55)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-14.83 (6.95)	
		95% CI	-28.52, -1.14	
		p-value [b]	0.0338	
	Effect size by Hedges' g (95% CI)	-0.44 (-0.96, 0.07)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.9 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	63 (96.9%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	43.74 (30.32)	52.86 (30.81)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	42 (64.6%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	46.83 (33.16)	49.49 (32.48)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.69 (3.68)	-4.39 (4.90)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.71 (6.12)	
		95% CI	-11.35, 12.76	
		p-value [b]	0.9084	
	Effect size by Hedges' g (95% CI)	0.02 (-0.39, 0.44)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.9 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (93.2%)	18 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	60.37 (35.78)	52.78 (36.27)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	35 (79.5%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	57.14 (34.60)	48.33 (34.68)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.66 (3.68)	5.19 (5.52)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.84 (6.61)	
		95% CI	-23.88, 2.20	
		p-value [b]	0.1027	
		Effect size by Hedges' g (95% CI)	-0.30 (-0.85, 0.25)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8442	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.9 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	36 (90.0%)	21 (91.3%)
	Observed mean domain score at baseline	Mean (SD)	46.53 (36.43)	52.38 (39.45)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (67.5%)	15 (65.2%)
	Observed mean domain score at Week 24	Mean (SD)	49.07 (29.81)	71.67 (33.89)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.22 (3.92)	4.44 (5.29)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.66 (6.59)	
		95% CI	-20.64, 5.32	
		p-value [b]	0.2460	
	Effect size by Hedges' g (95% CI)	-0.20 (-0.73, 0.33)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.9 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluatable subjects	N	65	33
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	60 (92.3%)	29 (87.9%)
	Observed mean domain score at baseline	Mean (SD)	42.50 (38.59)	40.52 (37.45)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	40 (61.5%)	18 (54.5%)
	Observed mean domain score at Week 24	Mean (SD)	39.38 (36.64)	43.06 (39.11)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-7.98 (3.42)	-1.55 (4.81)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.43 (5.89)	
		95% CI	-18.03, 5.16	
		p-value [b]	0.2758	
	Effect size by Hedges' g (95% CI)	-0.17 (-0.61, 0.27)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.9 (AMNOG)  
 QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	44 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	67.05 (32.40)	71.67 (40.14)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	36 (81.8%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	72.69 (30.32)	70.56 (34.63)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.65 (3.68)	4.72 (5.45)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.07 (6.57)	
		95% CI	-14.02, 11.88	
		p-value [b]	0.8709	
		Effect size by Hedges' g (95% CI)	-0.03 (-0.55, 0.49)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6819	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.9 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	36 (90.0%)	20 (87.0%)
	Observed mean domain score at baseline	Mean (SD)	66.90 (36.43)	60.00 (36.13)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (67.5%)	16 (69.6%)
	Observed mean domain score at Week 24	Mean (SD)	77.47 (33.40)	66.67 (34.96)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.79 (4.07)	6.70 (5.46)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.91 (6.82)	
		95% CI	-17.34, 9.51	
		p-value [b]	0.5665	
	Effect size by Hedges' g (95% CI)	-0.11 (-0.65, 0.43)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.9 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluatable subjects	N	65	33
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	61 (93.8%)	31 (93.9%)
	Observed mean domain score at baseline	Mean (SD)	55.74 (35.47)	58.60 (40.74)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	42 (64.6%)	20 (60.6%)
	Observed mean domain score at Week 24	Mean (SD)	52.58 (36.18)	51.67 (38.86)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.89 (3.47)	-4.70 (4.71)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.81 (5.84)	
		95% CI	-8.69, 14.31	
		p-value [b]	0.6306	
	Effect size by Hedges' g (95% CI)	0.07 (-0.35, 0.50)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.9 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	44 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	80.30 (10.99)	72.86 (15.53)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	37 (84.1%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	76.58 (14.56)	80.95 (11.66)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.78 (1.63)	3.96 (2.40)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.73 (2.91)	
		95% CI	-11.45, -0.01	
		p-value [b]	0.0496	
		Effect size by Hedges' g (95% CI)	-0.45 (-0.98, 0.08)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5957	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.9 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	40 (100.0%)	23 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	73.45 (15.66)	74.74 (13.80)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	30 (75.0%)	16 (69.6%)
	Observed mean domain score at Week 24	Mean (SD)	72.06 (14.94)	83.63 (10.85)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.92 (1.68)	3.52 (2.28)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.44 (2.83)	
		95% CI	-12.02, -0.86	
		p-value [b]	0.0238	
	Effect size by Hedges' g (95% CI)	-0.42 (-0.94, 0.09)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.9 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	65 (100.0%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	73.63 (15.08)	73.45 (14.04)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (66.2%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	71.43 (14.62)	70.78 (16.12)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.95 (1.50)	-0.51 (2.05)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.45 (2.53)	
		95% CI	-8.43, 1.54	
		p-value [b]	0.1746	
	Effect size by Hedges' g (95% CI)	-0.23 (-0.65, 0.18)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.9 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	44 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	63.69 (19.89)	59.76 (24.46)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	37 (84.1%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	69.82 (20.73)	64.13 (20.41)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	4.64 (2.52)	4.85 (3.73)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.21 (4.49)	
		95% CI	-9.05, 8.64	
		p-value [b]	0.9635	
		Effect size by Hedges' g (95% CI)	-0.01 (-0.53, 0.51)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6908	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.9 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	40 (100.0%)	22 (95.7%)
	Observed mean domain score at baseline	Mean (SD)	55.54 (26.30)	58.12 (21.58)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	29 (72.5%)	16 (69.6%)
	Observed mean domain score at Week 24	Mean (SD)	60.26 (23.77)	64.43 (21.09)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.61 (2.62)	0.91 (3.60)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.53 (4.45)	
		95% CI	-10.29, 7.24	
		p-value [b]	0.7321	
	Effect size by Hedges' g (95% CI)	-0.06 (-0.57, 0.45)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.9 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	64 (98.5%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.64 (24.82)	55.56 (23.67)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	43 (66.2%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	51.94 (27.05)	51.30 (28.08)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.87 (2.30)	0.60 (3.12)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.46 (3.87)	
		95% CI	-12.09, 3.16	
		p-value [b]	0.2504	
	Effect size by Hedges' g (95% CI)	-0.18 (-0.60, 0.24)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.9 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (93.2%)	18 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	68.51 (17.98)	61.97 (21.26)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	35 (79.5%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	66.76 (18.93)	63.39 (19.01)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.22 (2.14)	4.13 (3.17)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.34 (3.80)	
		95% CI	-14.84, 0.16	
		p-value [b]	0.0550	
		Effect size by Hedges' g (95% CI)	-0.38 (-0.93, 0.17)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1717	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.9 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	36 (90.0%)	21 (91.3%)
	Observed mean domain score at baseline	Mean (SD)	58.67 (22.58)	57.84 (20.55)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (65.0%)	15 (65.2%)
	Observed mean domain score at Week 24	Mean (SD)	60.26 (18.18)	71.82 (16.02)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.95 (2.27)	6.33 (3.01)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.28 (3.77)	
		95% CI	-17.71, -2.85	
		p-value [b]	0.0069	
	Effect size by Hedges' g (95% CI)	-0.46 (-1.00, 0.07)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.9 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	58 (89.2%)	29 (87.9%)
	Observed mean domain score at baseline	Mean (SD)	53.25 (20.55)	54.20 (20.82)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	39 (60.0%)	18 (54.5%)
	Observed mean domain score at Week 24	Mean (SD)	51.80 (20.74)	53.59 (23.96)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.67 (2.00)	-1.56 (2.73)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.11 (3.37)	
		95% CI	-8.75, 4.54	
		p-value [b]	0.5330	
	Effect size by Hedges' g (95% CI)	-0.10 (-0.54, 0.34)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.9 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (93.2%)	18 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	66.08 (16.91)	60.75 (22.57)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	35 (79.5%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	68.51 (18.22)	63.76 (18.93)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.14 (2.21)	5.22 (3.30)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.08 (3.96)	
		95% CI	-12.88, 2.72	
		p-value [b]	0.2007	
		Effect size by Hedges' g (95% CI)	-0.27 (-0.82, 0.28)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8155	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.9 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	36 (90.0%)	20 (87.0%)
	Observed mean domain score at baseline	Mean (SD)	57.81 (23.47)	57.96 (20.68)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (65.0%)	15 (65.2%)
	Observed mean domain score at Week 24	Mean (SD)	59.80 (19.99)	68.53 (16.28)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.39 (2.35)	4.54 (3.20)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.93 (3.97)	
		95% CI	-14.75, 0.88	
		p-value [b]	0.0818	
	Effect size by Hedges' g (95% CI)	-0.30 (-0.85, 0.24)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.9 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	58 (89.2%)	29 (87.9%)
	Observed mean domain score at baseline	Mean (SD)	52.78 (21.11)	55.67 (20.98)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	39 (60.0%)	18 (54.5%)
	Observed mean domain score at Week 24	Mean (SD)	51.39 (22.83)	53.32 (25.99)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.84 (2.08)	0.18 (2.83)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.02 (3.51)	
		95% CI	-10.93, 2.89	
		p-value [b]	0.2532	
	Effect size by Hedges' g (95% CI)	-0.19 (-0.63, 0.25)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.9 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden < 4 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	44	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	44 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	63.89 (17.62)	55.36 (21.11)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	37 (84.1%)	15 (75.0%)
	Observed mean domain score at Week 24	Mean (SD)	64.79 (22.18)	60.42 (18.58)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.12 (2.39)	4.61 (3.49)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.73 (4.23)	
		95% CI	-14.07, 2.60	
		p-value [b]	0.1768	
		Effect size by Hedges' g (95% CI)	-0.30 (-0.83, 0.22)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3510	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.9 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden 4-5 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	40	23
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	39 (97.5%)	22 (95.7%)
	Observed mean domain score at baseline	Mean (SD)	55.00 (23.99)	52.87 (19.82)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	28 (70.0%)	16 (69.6%)
	Observed mean domain score at Week 24	Mean (SD)	55.49 (19.21)	68.35 (17.24)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.04 (2.50)	4.40 (3.37)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.44 (4.20)	
		95% CI	-16.71, -0.18	
		p-value [b]	0.0452	
	Effect size by Hedges' g (95% CI)	-0.37 (-0.89, 0.15)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.9 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline transfusion burden ≥ 6 RBC units/8 weeks	Number of HRQoL evaluable subjects	N	65	33
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	61 (93.8%)	33 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	49.54 (19.13)	52.55 (20.31)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	42 (64.6%)	22 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	50.31 (22.83)	48.94 (23.76)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.03 (2.22)	-1.49 (2.92)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.54 (3.66)	
		95% CI	-8.75, 5.68	
		p-value [b]	0.6754	
	Effect size by Hedges' g (95% CI)	-0.08 (-0.50, 0.34)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.10 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	120 (97.6%)	63 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.33 (20.86)	51.39 (22.91)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	92 (74.8%)	42 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	53.53 (22.19)	54.76 (24.69)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		-2.99 (1.95)	1.78 (2.73)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-4.77 (3.36)	
	95% CI		-11.38, 1.85	
	p-value [b]		0.1569	
	Effect size by Hedges' g (95% CI)		-0.22 (-0.52, 0.08)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.7212	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.10 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	25 (100.0%)	13 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.00 (20.13)	51.92 (23.85)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (72.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	58.33 (25.36)	52.27 (21.52)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.96 (3.34)	0.71 (4.35)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.67 (5.48)	
		95% CI	-17.46, 4.13	
		p-value [b]	0.2250	
		Effect size by Hedges' g (95% CI)	-0.30 (-0.96, 0.36)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.10 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	120 (97.6%)	63 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.06 (33.36)	51.85 (27.72)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	90 (73.2%)	42 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	53.21 (32.49)	64.55 (31.83)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.39 (2.81)	1.39 (3.95)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.78 (4.85)	
		95% CI	-14.34, 4.77	
		p-value [b]	0.3248	
		Effect size by Hedges' g (95% CI)	-0.15 (-0.46, 0.15)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5047	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.10 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	25 (100.0%)	13 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	63.56 (34.02)	47.01 (38.26)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (72.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	59.26 (34.30)	47.47 (32.62)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.36 (5.45)	6.54 (7.02)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.90 (8.92)	
		95% CI	-28.47, 6.66	
		p-value [b]	0.2227	
		Effect size by Hedges' g (95% CI)	-0.30 (-0.96, 0.36)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.10 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	114 (92.7%)	56 (88.9%)
	Observed mean domain score at baseline	Mean (SD)	47.81 (36.82)	46.88 (37.56)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	85 (69.1%)	37 (58.7%)
	Observed mean domain score at Week 24	Mean (SD)	47.65 (32.87)	58.11 (36.35)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		-5.60 (2.64)	3.96 (3.88)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-9.56 (4.70)	
	95% CI		-18.82, -0.30	
	p-value [b]		0.0431	
	Effect size by Hedges' g (95% CI)		-0.26 (-0.58, 0.06)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.2909	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.10 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	22 (88.0%)	12 (92.3%)
	Observed mean domain score at baseline	Mean (SD)	55.68 (42.91)	50.00 (39.89)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	17 (68.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	50.00 (44.19)	38.64 (39.31)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.10 (5.21)	-5.43 (6.49)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.67 (8.32)	
		95% CI	-17.07, 15.73	
		p-value [b]	0.9361	
		Effect size by Hedges' g (95% CI)	-0.02 (-0.70, 0.67)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.10 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	115 (93.5%)	58 (92.1%)
	Observed mean domain score at baseline	Mean (SD)	62.54 (34.29)	61.35 (40.52)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	87 (70.7%)	40 (63.5%)
	Observed mean domain score at Week 24	Mean (SD)	67.43 (34.82)	62.29 (36.88)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		1.72 (2.73)	2.92 (3.86)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-1.20 (4.73)	
	95% CI		-10.51, 8.11	
	p-value [b]		0.8004	
	Effect size by Hedges' g (95% CI)		-0.03 (-0.35, 0.28)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.3632	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.10 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	25 (100.0%)	13 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	60.67 (39.35)	68.59 (33.36)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (72.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	58.33 (36.04)	60.61 (37.84)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.84 (4.96)	-5.52 (6.42)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	6.35 (8.11)	
		95% CI	-9.62, 22.33	
		p-value [b]	0.4342	
		Effect size by Hedges' g (95% CI)	0.17 (-0.49, 0.82)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.10 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	123 (100.0%)	63 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.76 (14.30)	73.02 (14.00)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	92 (74.8%)	42 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	73.50 (13.82)	78.57 (13.15)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.92 (1.21)	2.10 (1.72)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.01 (2.10)	
		95% CI	-9.16, -0.86	
		p-value [b]	0.0182	
		Effect size by Hedges' g (95% CI)	-0.35 (-0.66, -0.05)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9871	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.10 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	25 (100.0%)	13 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	74.29 (15.43)	76.92 (15.29)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (72.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	72.49 (19.23)	73.59 (18.97)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.69 (2.17)	1.38 (2.83)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.07 (3.56)	
		95% CI	-12.08, 1.95	
		p-value [b]	0.1560	
		Effect size by Hedges' g (95% CI)	-0.32 (-0.98, 0.34)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.10 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	122 (99.2%)	62 (98.4%)
	Observed mean domain score at baseline	Mean (SD)	56.58 (24.68)	56.91 (23.03)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	91 (74.0%)	42 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	59.92 (24.50)	62.02 (23.12)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.56 (1.81)	2.65 (2.57)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.21 (3.14)	
		95% CI	-9.41, 2.99	
		p-value [b]	0.3084	
		Effect size by Hedges' g (95% CI)	-0.13 (-0.44, 0.17)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2955	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.10 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	25 (100.0%)	13 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.19 (22.22)	59.89 (23.91)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (72.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	61.77 (29.02)	46.97 (27.11)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		0.82 (3.44)	-2.03 (4.48)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		2.84 (5.64)	
	95% CI		-8.28, 13.96	
	p-value [b]		0.6153	
	Effect size by Hedges' g (95% CI)		0.12 (-0.53, 0.78)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.10 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	112 (91.1%)	56 (88.9%)
	Observed mean domain score at baseline	Mean (SD)	58.71 (20.46)	56.71 (20.20)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	83 (67.5%)	37 (58.7%)
	Observed mean domain score at Week 24	Mean (SD)	59.16 (19.97)	64.68 (20.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.22 (1.50)	2.96 (2.16)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.18 (2.63)	
		95% CI	-11.37, -0.99	
		p-value [b]	0.0199	
		Effect size by Hedges' g (95% CI)	-0.30 (-0.62, 0.02)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8624	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.10 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	22 (88.0%)	12 (92.3%)
	Observed mean domain score at baseline	Mean (SD)	62.51 (25.60)	60.51 (24.03)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	17 (68.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	59.62 (22.73)	54.52 (23.43)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.63 (3.03)	0.70 (3.74)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.32 (4.81)	
		95% CI	-14.80, 4.15	
		p-value [b]	0.2693	
		Effect size by Hedges' g (95% CI)	-0.21 (-0.90, 0.48)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.10 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	112 (91.1%)	55 (87.3%)
	Observed mean domain score at baseline	Mean (SD)	57.81 (21.11)	56.87 (20.82)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	83 (67.5%)	37 (58.7%)
	Observed mean domain score at Week 24	Mean (SD)	59.31 (21.21)	64.48 (20.24)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.11 (1.61)	3.76 (2.32)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.87 (2.82)	
		95% CI	-11.45, -0.30	
		p-value [b]	0.0390	
		Effect size by Hedges' g (95% CI)	-0.28 (-0.60, 0.04)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3697	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.10 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	22 (88.0%)	12 (92.3%)
	Observed mean domain score at baseline	Mean (SD)	60.42 (22.52)	61.60 (22.89)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	17 (68.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	60.83 (24.45)	50.74 (24.24)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.07 (3.12)	-0.74 (3.89)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.33 (4.98)	
		95% CI	-11.15, 8.49	
		p-value [b]	0.7895	
Effect size by Hedges' g (95% CI)		-0.06 (-0.74, 0.63)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.10 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Very Low, Low	Number of HRQoL evaluable subjects	N	123	63
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	118 (95.9%)	62 (98.4%)
	Observed mean domain score at baseline	Mean (SD)	54.57 (20.75)	53.49 (19.28)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	89 (72.4%)	42 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	56.04 (22.27)	60.44 (20.83)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		-2.33 (1.72)	1.76 (2.40)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-4.09 (2.95)	
	95% CI		-9.92, 1.73	
	p-value [b]		0.1673	
	Effect size by Hedges' g (95% CI)		-0.20 (-0.51, 0.11)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.5903	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.10 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline IRSS-R Intermediate	Number of HRQoL evaluable subjects	N	25	13
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	25 (100.0%)	13 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	59.25 (22.00)	52.94 (24.76)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (72.0%)	11 (84.6%)
	Observed mean domain score at Week 24	Mean (SD)	59.79 (23.50)	48.91 (24.11)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.26 (3.21)	2.72 (4.16)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.98 (5.26)	
		95% CI	-17.34, 3.38	
		p-value [b]	0.1855	
		Effect size by Hedges' g (95% CI)	-0.30 (-0.96, 0.36)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.11 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	6 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.00 (23.15)	52.08 (22.94)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	50.00 (26.73)	67.50 (18.96)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-10.67 (5.37)	5.22 (5.92)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-15.89 (7.99)	
		95% CI	-31.63, -0.16	
		p-value [b]	0.0477	
		Effect size by Hedges' g (95% CI)	-0.65 (-1.66, 0.37)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3324	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.11 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	125	61
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	122 (97.6%)	61 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.00 (20.47)	53.28 (22.81)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	89 (71.2%)	41 (67.2%)
	Observed mean domain score at Week 24	Mean (SD)	53.79 (22.33)	53.66 (24.88)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.96 (1.96)	1.35 (2.74)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.32 (3.37)	
		95% CI	-10.96, 2.33	
		p-value [b]	0.2016	
	Effect size by Hedges' g (95% CI)	-0.20 (-0.51, 0.10)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.11 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	16 (100.0%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.13 (22.13)	38.89 (22.05)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	13 (81.3%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	60.58 (23.30)	48.21 (19.67)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.93 (3.80)	0.85 (5.07)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.78 (6.33)	
		95% CI	-16.25, 8.69	
		p-value [b]	0.5514	
	Effect size by Hedges' g (95% CI)	-0.17 (-0.96, 0.63)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.11 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	6 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	65.28 (37.77)	55.56 (32.96)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	55.56 (42.00)	73.33 (20.18)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.31 (9.01)	0.69 (10.14)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.39 (13.56)	
		95% CI	-27.11, 26.33	
		p-value [b]	0.9773	
		Effect size by Hedges' g (95% CI)	-0.01 (-1.00, 0.98)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8909	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.11 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	125	61
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	122 (97.6%)	61 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.83 (33.59)	53.01 (29.98)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	88 (70.4%)	41 (67.2%)
	Observed mean domain score at Week 24	Mean (SD)	53.16 (31.98)	59.62 (33.77)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.33 (2.85)	2.39 (4.00)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.73 (4.92)	
		95% CI	-16.42, 2.96	
		p-value [b]	0.1725	
	Effect size by Hedges' g (95% CI)	-0.21 (-0.51, 0.10)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.11 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	16 (100.0%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	60.42 (33.20)	34.57 (20.37)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	12 (75.0%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	61.11 (33.67)	60.32 (33.25)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.30 (6.30)	1.44 (8.42)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.74 (10.54)	
		95% CI	-25.51, 16.03	
		p-value [b]	0.6533	
	Effect size by Hedges' g (95% CI)	-0.16 (-0.95, 0.63)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.11 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at baseline	Mean (SD)	62.50 (42.26)	45.00 (44.72)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	34.38 (37.65)	55.00 (32.60)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-13.42 (7.63)	7.15 (9.43)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-20.58 (12.13)	
		95% CI	-44.51, 3.36	
		p-value [b]	0.0916	
		Effect size by Hedges' g (95% CI)	-0.44 (-1.50, 0.61)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4111	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.11 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	125	61
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	114 (91.2%)	56 (91.8%)
	Observed mean domain score at baseline	Mean (SD)	48.90 (37.98)	51.34 (37.06)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	81 (64.8%)	37 (60.7%)
	Observed mean domain score at Week 24	Mean (SD)	48.77 (35.55)	56.76 (38.48)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.41 (2.69)	2.75 (3.87)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.16 (4.71)	
		95% CI	-17.45, 1.13	
		p-value [b]	0.0849	
	Effect size by Hedges' g (95% CI)	-0.22 (-0.53, 0.10)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.11 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	15 (93.8%)	7 (77.8%)
	Observed mean domain score at baseline	Mean (SD)	41.67 (33.63)	17.86 (27.82)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	13 (81.3%)	6 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	51.92 (27.88)	33.33 (34.16)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.86 (5.60)	-5.42 (8.43)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.44 (10.07)	
		95% CI	-20.30, 19.41	
		p-value [b]	0.9650	
	Effect size by Hedges' g (95% CI)	-0.01 (-0.88, 0.85)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.11 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	6 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	65.63 (45.08)	61.11 (44.62)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	53.13 (40.57)	68.33 (45.80)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.32 (7.95)	3.96 (9.06)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.28 (12.06)	
		95% CI	-34.07, 13.50	
		p-value [b]	0.3949	
		Effect size by Hedges' g (95% CI)	-0.21 (-1.21, 0.78)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3336	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.11 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluatable subjects	N	125	61
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	118 (94.4%)	56 (91.8%)
	Observed mean domain score at baseline	Mean (SD)	59.96 (34.68)	65.63 (38.83)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	84 (67.2%)	39 (63.9%)
	Observed mean domain score at Week 24	Mean (SD)	64.88 (35.52)	65.17 (34.67)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.45 (2.71)	2.59 (3.89)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.14 (4.75)	
		95% CI	-10.49, 8.21	
		p-value [b]	0.8104	
	Effect size by Hedges' g (95% CI)	-0.03 (-0.35, 0.29)		

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.11 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	15 (93.8%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	77.22 (29.79)	45.37 (37.76)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	13 (81.3%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	80.13 (24.66)	39.29 (39.00)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.74 (5.92)	-8.00 (7.65)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	10.74 (9.72)	
		95% CI	-8.43, 29.90	
		p-value [b]	0.2708	
		Effect size by Hedges' g (95% CI)	0.31 (-0.49, 1.12)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.11 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	6 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	82.14 (20.16)	77.78 (11.13)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	72.02 (21.21)	83.81 (4.26)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.47 (3.52)	3.50 (3.90)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.96 (5.25)	
		95% CI	-19.30, 1.38	
		p-value [b]	0.0891	
		Effect size by Hedges' g (95% CI)	-0.49 (-1.50, 0.51)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4377	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.11 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	125	61
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	125 (100.0%)	61 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.35 (13.74)	74.94 (13.38)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	89 (71.2%)	41 (67.2%)
	Observed mean domain score at Week 24	Mean (SD)	73.41 (14.21)	77.70 (15.07)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.83 (1.22)	2.35 (1.74)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.19 (2.12)	
		95% CI	-9.37, -1.00	
		p-value [b]	0.0154	
	Effect size by Hedges' g (95% CI)	-0.38 (-0.69, -0.07)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.11 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	16 (100.0%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	73.81 (16.50)	62.43 (17.41)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	13 (81.3%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	73.63 (15.14)	72.11 (14.91)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.88 (2.47)	-1.77 (3.34)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.11 (4.15)	
		95% CI	-9.28, 7.07	
		p-value [b]	0.7901	
		Effect size by Hedges' g (95% CI)	-0.06 (-0.85, 0.73)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.11 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	6 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	61.01 (30.22)	65.87 (20.90)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	57.14 (35.93)	73.33 (23.89)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.27 (5.61)	8.20 (6.25)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.93 (8.40)	
		95% CI	-21.48, 11.63	
		p-value [b]	0.5581	
		Effect size by Hedges' g (95% CI)	-0.17 (-1.17, 0.82)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3528	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.11 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	125	61
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	124 (99.2%)	60 (98.4%)
	Observed mean domain score at baseline	Mean (SD)	57.37 (24.07)	58.77 (22.74)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	88 (70.4%)	41 (67.2%)
	Observed mean domain score at Week 24	Mean (SD)	59.93 (25.04)	57.43 (25.43)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.45 (1.82)	1.39 (2.58)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.84 (3.16)	
		95% CI	-9.07, 3.39	
		p-value [b]	0.3695	
	Effect size by Hedges' g (95% CI)	-0.12 (-0.43, 0.19)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.11 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	16 (100.0%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	49.40 (22.09)	42.86 (22.84)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	13 (81.3%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	64.10 (19.38)	57.14 (17.98)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.93 (3.94)	0.47 (5.27)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	6.46 (6.54)	
		95% CI	-6.44, 19.35	
		p-value [b]	0.3248	
		Effect size by Hedges' g (95% CI)	0.28 (-0.51, 1.07)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.11 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at baseline	Mean (SD)	65.11 (30.57)	54.37 (20.78)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	53.02 (25.91)	69.60 (6.71)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-8.48 (4.56)	4.91 (5.65)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-13.39 (7.27)	
		95% CI	-27.73, 0.94	
		p-value [b]	0.0668	
		Effect size by Hedges' g (95% CI)	-0.45 (-1.51, 0.60)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5667	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.11 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	125	61
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	112 (89.6%)	56 (91.8%)
	Observed mean domain score at baseline	Mean (SD)	58.63 (21.08)	60.08 (19.95)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	80 (64.0%)	37 (60.7%)
	Observed mean domain score at Week 24	Mean (SD)	58.83 (20.12)	62.89 (22.57)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.38 (1.54)	2.38 (2.17)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.76 (2.66)	
		95% CI	-11.01, -0.51	
		p-value [b]	0.0317	
Effect size by Hedges' g (95% CI)		-0.28 (-0.60, 0.04)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.11 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	15 (93.8%)	7 (77.8%)
	Observed mean domain score at baseline	Mean (SD)	61.50 (17.39)	37.98 (19.29)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	12 (75.0%)	6 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	66.09 (17.53)	53.02 (19.15)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.84 (3.30)	1.94 (4.97)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.79 (5.98)	
		95% CI	-16.58, 7.01	
		p-value [b]	0.4244	
		Effect size by Hedges' g (95% CI)	-0.26 (-1.12, 0.61)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.11 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at baseline	Mean (SD)	63.06 (29.55)	60.76 (21.24)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	55.08 (30.58)	71.46 (12.91)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.89 (4.63)	9.58 (5.72)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-12.46 (7.35)	
		95% CI	-26.97, 2.04	
		p-value [b]	0.0917	
		Effect size by Hedges' g (95% CI)	-0.43 (-1.48, 0.62)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.4313	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.11 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	125	61
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	112 (89.6%)	55 (90.2%)
	Observed mean domain score at baseline	Mean (SD)	58.31 (21.04)	59.97 (20.23)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	80 (64.0%)	37 (60.7%)
	Observed mean domain score at Week 24	Mean (SD)	59.13 (21.41)	60.98 (23.13)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.74 (1.63)	2.28 (2.31)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.02 (2.82)	
		95% CI	-10.59, 0.56	
		p-value [b]	0.0773	
	Effect size by Hedges' g (95% CI)	-0.24 (-0.56, 0.08)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.11 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	15 (93.8%)	7 (77.8%)
	Observed mean domain score at baseline	Mean (SD)	54.40 (18.38)	37.87 (20.06)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	12 (75.0%)	6 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	65.49 (16.88)	55.08 (17.62)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.28 (3.37)	2.86 (5.03)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.59 (6.03)	
		95% CI	-12.48, 11.30	
		p-value [b]	0.9225	
		Effect size by Hedges' g (95% CI)	-0.03 (-0.89, 0.83)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.11 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count < 100 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	8	6
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	8 (100.0%)	6 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.76 (27.38)	57.84 (21.04)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	8 (100.0%)	5 (83.3%)
	Observed mean domain score at Week 24	Mean (SD)	54.23 (31.11)	71.39 (15.01)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.37 (5.27)	3.94 (5.94)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.31 (7.94)	
		95% CI	-21.96, 9.33	
		p-value [b]	0.4274	
		Effect size by Hedges' g (95% CI)	-0.24 (-1.23, 0.76)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9119	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.11 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count 100-400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	125	61
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	120 (96.0%)	60 (98.4%)
	Observed mean domain score at baseline	Mean (SD)	55.32 (20.68)	55.15 (19.71)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	87 (69.6%)	41 (67.2%)
	Observed mean domain score at Week 24	Mean (SD)	56.05 (22.02)	56.90 (22.80)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.14 (1.73)	1.63 (2.42)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.77 (2.98)	
		95% CI	-10.65, 1.10	
		p-value [b]	0.1107	
	Effect size by Hedges' g (95% CI)	-0.23 (-0.54, 0.08)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.11 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline platelet count > 400 x 10 <sup>9</sup> /L	Number of HRQoL evaluable subjects	N	16	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	16 (100.0%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.32 (20.30)	38.77 (18.23)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	12 (75.0%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	62.83 (19.57)	55.22 (18.52)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.04 (3.71)	2.52 (4.98)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.48 (6.20)	
		95% CI	-14.70, 9.74	
		p-value [b]	0.6897	
	Effect size by Hedges' g (95% CI)	-0.12 (-0.91, 0.67)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.12 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	30 (93.8%)	12 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	49.17 (20.74)	44.79 (23.51)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (81.3%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	51.92 (22.83)	55.36 (31.34)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.54 (3.02)	-0.48 (4.59)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.06 (5.48)	
		95% CI	-11.85, 9.73	
		p-value [b]	0.8470	
		Effect size by Hedges' g (95% CI)	-0.05 (-0.71, 0.61)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6840	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.12 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	109 (99.1%)	62 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.93 (20.70)	53.02 (23.05)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	78 (70.9%)	45 (72.6%)
	Observed mean domain score at Week 24	Mean (SD)	56.57 (22.13)	53.61 (23.02)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.69 (2.00)	2.04 (2.73)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.73 (3.38)	
		95% CI	-12.40, 0.93	
		p-value [b]	0.0915	
	Effect size by Hedges' g (95% CI)	-0.26 (-0.58, 0.05)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.12 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	1 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.57 (20.04)	50.00 (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	6 (85.7%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	35.42 (22.94)	75.00 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.18 (5.36)	-0.64 (13.73)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.54 (14.74)	
		95% CI	-33.62, 24.55	
		p-value [b]	0.7586	
	Effect size by Hedges' g (95% CI)	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.12 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (90.6%)	12 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.64 (37.10)	56.48 (25.71)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (75.0%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	50.93 (37.77)	73.02 (31.98)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-7.66 (4.90)	-5.69 (7.44)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.97 (8.91)	
		95% CI	-19.51, 15.58	
		p-value [b]	0.8256	
Effect size by Hedges' g (95% CI)		-0.06 (-0.72, 0.60)		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8482		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.12 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	110 (100.0%)	62 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.75 (32.65)	50.18 (30.69)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	78 (70.9%)	45 (72.6%)
	Observed mean domain score at Week 24	Mean (SD)	54.27 (30.96)	59.01 (32.81)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.75 (2.90)	3.44 (3.93)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.20 (4.89)	
		95% CI	-16.83, 2.44	
		p-value [b]	0.1423	
	Effect size by Hedges' g (95% CI)	-0.22 (-0.54, 0.09)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.12 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	1 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	68.25 (37.64)	33.33 (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	6 (85.7%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	66.67 (36.51)	66.67 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.71 (9.02)	15.07 (23.39)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.36 (25.10)	
		95% CI	-57.88, 41.16	
		p-value [b]	0.7395	
	Effect size by Hedges' g (95% CI)	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.12 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	30 (93.8%)	10 (83.3%)
	Observed mean domain score at baseline	Mean (SD)	48.33 (37.68)	42.50 (42.57)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (75.0%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	48.96 (32.54)	42.86 (31.34)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.75 (4.26)	-3.55 (7.08)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.20 (8.26)	
		95% CI	-16.48, 16.09	
		p-value [b]	0.9809	
Effect size by Hedges' g (95% CI)		0.00 (-0.71, 0.70)		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2234		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.12 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	100 (90.9%)	56 (90.3%)
	Observed mean domain score at baseline	Mean (SD)	50.00 (37.94)	49.11 (37.22)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	73 (66.4%)	40 (64.5%)
	Observed mean domain score at Week 24	Mean (SD)	49.32 (35.59)	55.63 (39.01)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-7.48 (2.76)	3.46 (3.83)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.94 (4.72)	
		95% CI	-20.25, -1.64	
		p-value [b]	0.0214	
	Effect size by Hedges' g (95% CI)	-0.29 (-0.62, 0.04)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.12 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	1 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	35.71 (37.80)	50.00 (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	25.00 (30.62)	50.00 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.03 (7.87)	-10.16 (19.97)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		16.19 (21.47)
			95% CI	-26.22, 58.60
			p-value [b]	0.4519
	Effect size by Hedges' g (95% CI)		NE (NE, NE)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.12 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (90.6%)	11 (91.7%)
	Observed mean domain score at baseline	Mean (SD)	74.14 (31.21)	60.61 (49.03)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (71.9%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	65.22 (29.69)	67.86 (38.62)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.85 (4.62)	-0.89 (7.00)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.95 (8.39)	
		95% CI	-18.48, 14.58	
		p-value [b]	0.8164	
Effect size by Hedges' g (95% CI)		-0.05 (-0.73, 0.63)		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8289		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.12 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	106 (96.4%)	59 (95.2%)
	Observed mean domain score at baseline	Mean (SD)	58.02 (35.58)	64.12 (36.99)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	77 (70.0%)	43 (69.4%)
	Observed mean domain score at Week 24	Mean (SD)	64.94 (37.05)	62.02 (36.56)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.66 (2.80)	1.75 (3.81)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.09 (4.73)	
		95% CI	-9.42, 9.23	
		p-value [b]	0.9842	
	Effect size by Hedges' g (95% CI)	0.00 (-0.32, 0.31)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.12 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	6 (85.7%)	0 (NE)
	Observed mean domain score at baseline	Mean (SD)	76.39 (28.59)	NE (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	83.33 (23.57)	16.67 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	10.00 (8.88)	NE (NE)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		NE (NE)
		95% CI		NE, NE
		p-value [b]		NE
	Effect size by Hedges' g (95% CI)		NE (NE, NE)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.12 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	12 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.74 (16.06)	72.22 (11.63)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (81.3%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	73.44 (13.14)	77.55 (12.51)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.86 (1.89)	0.65 (2.98)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.51 (3.53)	
		95% CI	-10.46, 3.44	
		p-value [b]	0.3208	
		Effect size by Hedges' g (95% CI)	-0.23 (-0.88, 0.42)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7483	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.12 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance ≥ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	110 (100.0%)	62 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.50 (14.23)	74.27 (14.77)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	78 (70.9%)	45 (72.6%)
	Observed mean domain score at Week 24	Mean (SD)	73.87 (14.72)	77.35 (15.00)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.23 (1.25)	2.03 (1.72)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.26 (2.13)	
		95% CI	-9.46, -1.06	
		p-value [b]	0.0143	
	Effect size by Hedges' g (95% CI)	-0.36 (-0.68, -0.05)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.12 (AMNOG)

QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	1 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.51 (10.44)	71.43 (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	6 (85.7%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	65.87 (21.60)	85.71 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.06 (3.51)	10.47 (9.05)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.53 (9.71)	
		95% CI	-29.68, 8.62	
		p-value [b]	0.2793	
	Effect size by Hedges' g (95% CI)	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.12 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	12 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.46 (27.08)	52.78 (21.41)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	25 (78.1%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	58.19 (25.84)	53.40 (26.61)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.71 (2.96)	-0.37 (4.68)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	4.08 (5.53)	
		95% CI	-6.81, 14.97	
		p-value [b]	0.4612	
		Effect size by Hedges' g (95% CI)	0.16 (-0.50, 0.81)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1402	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.12 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	109 (99.1%)	62 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.41 (23.51)	58.91 (23.11)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	78 (70.9%)	45 (72.6%)
	Observed mean domain score at Week 24	Mean (SD)	60.87 (25.56)	60.21 (24.43)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.05 (1.88)	2.25 (2.55)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.30 (3.17)	
		95% CI	-10.55, 1.95	
		p-value [b]	0.1763	
	Effect size by Hedges' g (95% CI)	-0.18 (-0.49, 0.13)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.12 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	0 (NE)
	Observed mean domain score at baseline	Mean (SD)	49.66 (19.99)	NE (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	6 (85.7%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	60.32 (19.61)	38.10 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	9.48 (5.65)	NE (NE)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		NE (NE)
		95% CI		NE, NE
		p-value [b]		NE
	Effect size by Hedges' g (95% CI)		NE (NE, NE)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.12 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	12 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.46 (27.08)	52.78 (21.41)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	25 (78.1%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	58.19 (25.84)	53.40 (26.61)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	3.71 (2.96)	-0.37 (4.68)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	4.08 (5.53)	
		95% CI	-6.81, 14.97	
		p-value [b]	0.4612	
Effect size by Hedges' g (95% CI)		0.16 (-0.50, 0.81)		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1402		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.12 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	109 (99.1%)	62 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	58.41 (23.51)	58.91 (23.11)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	78 (70.9%)	45 (72.6%)
	Observed mean domain score at Week 24	Mean (SD)	60.87 (25.56)	60.21 (24.43)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.05 (1.88)	2.25 (2.55)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.30 (3.17)	
		95% CI	-10.55, 1.95	
		p-value [b]	0.1763	
	Effect size by Hedges' g (95% CI)	-0.18 (-0.49, 0.13)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.12 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	0 (NE)
	Observed mean domain score at baseline	Mean (SD)	49.66 (19.99)	NE (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	6 (85.7%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	60.32 (19.61)	38.10 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	9.48 (5.65)	NE (NE)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		NE (NE)
		95% CI		NE, NE
		p-value [b]		NE
	Effect size by Hedges' g (95% CI)		NE (NE, NE)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.12 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	28 (87.5%)	10 (83.3%)
	Observed mean domain score at baseline	Mean (SD)	59.95 (23.27)	54.36 (20.95)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	22 (68.8%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	58.21 (19.88)	63.33 (25.02)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.33 (2.58)	-0.89 (4.15)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.44 (4.89)	
		95% CI	-13.08, 6.19	
		p-value [b]	0.4817	
Effect size by Hedges' g (95% CI)		-0.15 (-0.86, 0.56)		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7016		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.12 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	100 (90.9%)	56 (90.3%)
	Observed mean domain score at baseline	Mean (SD)	59.09 (20.78)	58.52 (20.91)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	73 (66.4%)	40 (64.5%)
	Observed mean domain score at Week 24	Mean (SD)	59.99 (20.27)	62.27 (21.11)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.95 (1.58)	3.11 (2.16)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.06 (2.68)	
		95% CI	-12.34, -1.78	
		p-value [b]	0.0091	
	Effect size by Hedges' g (95% CI)	-0.34 (-0.66, -0.01)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.12 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	1 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	60.28 (22.65)	51.19 (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	52.74 (26.45)	58.81 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.68 (4.66)	3.41 (11.97)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.72 (12.84)	
		95% CI	-26.09, 24.64	
		p-value [b]	0.9551	
	Effect size by Hedges' g (95% CI)	NE (NE, NE)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.12 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	28 (87.5%)	10 (83.3%)
	Observed mean domain score at baseline	Mean (SD)	56.16 (23.75)	53.37 (21.03)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	22 (68.8%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	58.00 (22.31)	58.36 (24.62)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.54 (2.65)	1.06 (4.21)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.60 (4.97)	
		95% CI	-11.39, 8.19	
		p-value [b]	0.7479	
Effect size by Hedges' g (95% CI)		-0.07 (-0.77, 0.64)		
	p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3308		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.12 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	100 (90.9%)	56 (90.3%)
	Observed mean domain score at baseline	Mean (SD)	58.94 (20.65)	59.07 (20.93)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	73 (66.4%)	40 (64.5%)
	Observed mean domain score at Week 24	Mean (SD)	60.23 (21.64)	62.18 (21.68)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.27 (1.67)	3.24 (2.29)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.51 (2.83)	
		95% CI	-12.11, -0.92	
		p-value [b]	0.0227	
	Effect size by Hedges' g (95% CI)	-0.31 (-0.64, 0.02)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.12 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	0 (NE)
	Observed mean domain score at baseline	Mean (SD)	54.97 (21.18)	NE (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	5 (71.4%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	56.85 (23.46)	48.45 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.71 (4.70)	NE (NE)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		NE (NE)
			95% CI	NE, NE
			p-value [b]	NE
Effect size by Hedges' g (95% CI)			NE (NE, NE)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.12 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance 40 - <60	Number of HRQoL evaluable subjects	N	32	12
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (90.6%)	12 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.22 (23.40)	51.35 (19.46)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (71.9%)	7 (58.3%)
	Observed mean domain score at Week 24	Mean (SD)	55.42 (25.20)	60.59 (28.40)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.65 (2.92)	-2.55 (4.38)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.90 (5.26)	
		95% CI	-9.46, 11.26	
		p-value [b]	0.8645	
		Effect size by Hedges' g (95% CI)	0.04 (-0.62, 0.70)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1930	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.12 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance $\geq$ 60	Number of HRQoL evaluable subjects	N	110	62
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	108 (98.2%)	62 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.14 (20.15)	54.04 (20.47)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	78 (70.9%)	45 (72.6%)
	Observed mean domain score at Week 24	Mean (SD)	57.24 (21.72)	57.61 (21.25)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.33 (1.78)	2.79 (2.39)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.12 (2.98)	
		95% CI	-12.00, -0.25	
		p-value [b]	0.0411	
	Effect size by Hedges' g (95% CI)	-0.30 (-0.61, 0.01)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.12 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline renal status creatinine clearance < 40	Number of HRQoL evaluable subjects	N	7	1
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	7 (100.0%)	0 (NE)
	Observed mean domain score at baseline	Mean (SD)	57.16 (23.41)	NE (NE)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	6 (85.7%)	1 (100.0%)
	Observed mean domain score at Week 24	Mean (SD)	54.13 (23.82)	59.92 (NE)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	2.07 (5.31)	NE (NE)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		NE (NE)
		95% CI		NE, NE
		p-value [b]		NE
	Effect size by Hedges' g (95% CI)		NE (NE, NE)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.13 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects	N	50	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	49 (98.0%)	31 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.38 (18.42)	52.42 (25.70)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	38 (76.0%)	23 (74.2%)
	Observed mean domain score at Week 24	Mean (SD)	57.89 (23.86)	57.07 (26.06)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.26 (2.51)	0.93 (3.23)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.19 (4.09)	
		95% CI	-13.25, 2.87	
		p-value [b]	0.2059	
		Effect size by Hedges' g (95% CI)	-0.24 (-0.69, 0.21)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8128	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.13 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	34 (94.4%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.74 (20.63)	49.34 (20.19)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (72.2%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	54.81 (26.48)	51.04 (19.55)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.82 (2.92)	3.71 (3.93)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.53 (4.89)	
		95% CI	-15.15, 4.09	
		p-value [b]	0.2588	
		Effect size by Hedges' g (95% CI)	-0.27 (-0.82, 0.29)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.13 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	15 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.18 (21.47)	56.67 (25.82)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (65.9%)	10 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	51.85 (19.21)	47.50 (26.87)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.08 (2.70)	-1.55 (4.19)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.54 (4.98)	
		95% CI	-11.34, 8.26	
		p-value [b]	0.7576	
	Effect size by Hedges' g (95% CI)	-0.07 (-0.65, 0.52)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.13 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects	N	21	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	21 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.60 (23.87)	45.45 (14.00)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (85.7%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	51.39 (19.12)	59.38 (20.86)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.50 (3.41)	4.31 (4.66)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.81 (5.76)	
		95% CI	-19.16, 3.54	
		p-value [b]	0.1764	
Effect size by Hedges' g (95% CI)		-0.36 (-1.08, 0.36)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.13 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects	N	50	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	49 (98.0%)	31 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	61.45 (33.88)	52.69 (28.40)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	37 (74.0%)	23 (74.2%)
	Observed mean domain score at Week 24	Mean (SD)	60.96 (35.87)	63.29 (32.73)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.08 (3.93)	1.91 (4.97)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.99 (6.34)	
		95% CI	-18.49, 6.50	
		p-value [b]	0.3458	
		Effect size by Hedges' g (95% CI)	-0.19 (-0.63, 0.26)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7972	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.13 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	34 (94.4%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	49.67 (34.52)	45.03 (30.42)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	25 (69.4%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	55.11 (29.48)	67.59 (33.32)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.67 (4.67)	5.34 (6.29)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-11.01 (7.82)	
		95% CI	-26.42, 4.39	
		p-value [b]	0.1604	
		Effect size by Hedges' g (95% CI)	-0.33 (-0.88, 0.23)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.13 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	15 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.56 (34.07)	60.74 (31.11)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (65.9%)	10 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	45.27 (30.50)	51.11 (27.81)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.20 (4.28)	-4.60 (6.80)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.60 (8.03)	
		95% CI	-16.41, 15.21	
		p-value [b]	0.9404	
	Effect size by Hedges' g (95% CI)	-0.02 (-0.60, 0.57)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.13 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects	N	21	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	21 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	51.32 (29.50)	43.43 (28.74)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (85.7%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	53.70 (33.28)	56.94 (38.23)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.25 (5.54)	5.90 (7.66)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.15 (9.44)	
		95% CI	-24.75, 12.45	
		p-value [b]	0.5152	
		Effect size by Hedges' g (95% CI)	-0.21 (-0.92, 0.51)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.13 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects	N	50	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	42 (84.0%)	27 (87.1%)
	Observed mean domain score at baseline	Mean (SD)	55.36 (38.07)	46.30 (35.15)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	35 (70.0%)	22 (71.0%)
	Observed mean domain score at Week 24	Mean (SD)	54.29 (34.56)	50.00 (39.34)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		-5.13 (3.73)	2.32 (4.70)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-7.45 (6.00)	
	95% CI		-19.28, 4.38	
	p-value [b]		0.2157	
	Effect size by Hedges' g (95% CI)		-0.20 (-0.68, 0.28)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.9826	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.13 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	34 (94.4%)	18 (94.7%)
	Observed mean domain score at baseline	Mean (SD)	43.38 (38.09)	38.89 (40.42)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (63.9%)	11 (57.9%)
	Observed mean domain score at Week 24	Mean (SD)	46.74 (33.97)	52.27 (39.46)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.71 (4.19)	4.21 (5.71)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.91 (7.07)	
		95% CI	-20.84, 7.02	
		p-value [b]	0.3292	
		Effect size by Hedges' g (95% CI)	-0.18 (-0.74, 0.39)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.13 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	12 (80.0%)
	Observed mean domain score at baseline	Mean (SD)	56.10 (36.13)	62.50 (40.59)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (63.4%)	8 (53.3%)
	Observed mean domain score at Week 24	Mean (SD)	48.08 (34.59)	56.25 (32.04)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-11.12 (3.82)	-0.79 (6.88)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.33 (7.84)	
		95% CI	-25.77, 5.11	
		p-value [b]	0.1887	
	Effect size by Hedges' g (95% CI)	-0.27 (-0.91, 0.36)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.13 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects	N	21	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	19 (90.5%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	31.58 (34.20)	47.73 (36.15)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	17 (81.0%)	7 (63.6%)
	Observed mean domain score at Week 24	Mean (SD)	36.76 (37.62)	64.29 (40.46)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.31 (5.12)	1.51 (6.72)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.83 (8.45)	
		95% CI	-23.49, 9.83	
		p-value [b]	0.4199	
		Effect size by Hedges' g (95% CI)	-0.19 (-0.91, 0.53)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.13 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC	
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects	N	50	31	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	46 (92.0%)	29 (93.5%)	
	Observed mean domain score at baseline	Mean (SD)	67.21 (32.89)	60.63 (39.21)	
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	35 (70.0%)	22 (71.0%)	
	Observed mean domain score at Week 24	Mean (SD)	70.48 (28.61)	56.44 (39.33)	
	LS mean change from baseline at Week 24				
	Within-group	LS mean change from baseline (SE)	2.08 (3.75)	2.36 (4.76)	
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.27 (6.06)		
		95% CI	-12.22, 11.67		
		p-value [b]	0.9640		
		Effect size by Hedges' g (95% CI)	-0.01 (-0.47, 0.45)		
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9233		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.13 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	34 (94.4%)	18 (94.7%)
	Observed mean domain score at baseline	Mean (SD)	62.75 (32.71)	57.87 (44.69)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (72.2%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	73.72 (33.89)	57.64 (45.84)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.35 (4.28)	2.24 (5.84)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.90 (7.24)	
		95% CI	-16.15, 12.35	
		p-value [b]	0.7936	
		Effect size by Hedges' g (95% CI)	-0.05 (-0.61, 0.51)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.13 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	14 (93.3%)
	Observed mean domain score at baseline	Mean (SD)	69.72 (34.90)	65.48 (40.94)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (63.4%)	9 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	77.56 (28.75)	77.78 (20.83)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	5.36 (4.01)	1.88 (6.44)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	3.49 (7.58)	
		95% CI	-11.45, 18.42	
		p-value [b]	0.6462	
	Effect size by Hedges' g (95% CI)	0.09 (-0.50, 0.69)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.13 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluatable subjects	N	21	11
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	19 (90.5%)	10 (90.9%)
	Observed mean domain score at baseline	Mean (SD)	35.53 (31.53)	73.33 (27.72)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	17 (81.0%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	28.92 (35.00)	65.63 (26.89)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-7.77 (5.40)	-4.48 (7.26)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.29 (9.10)	
		95% CI	-21.23, 14.64	
		p-value [b]	0.7179	
		Effect size by Hedges' g (95% CI)	-0.11 (-0.85, 0.64)	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.13 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects	N	50	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	50 (100.0%)	31 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	77.24 (14.88)	72.81 (12.61)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	38 (76.0%)	23 (74.2%)
	Observed mean domain score at Week 24	Mean (SD)	74.06 (13.18)	78.05 (12.16)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.19 (1.57)	1.30 (2.04)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.49 (2.58)	
		95% CI	-10.57, -0.41	
		p-value [b]	0.0341	
		Effect size by Hedges' g (95% CI)	-0.39 (-0.83, 0.06)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3955	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.13 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	36 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	73.68 (12.85)	70.43 (15.68)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (72.2%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	73.26 (16.66)	83.73 (15.52)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.36 (1.82)	5.18 (2.52)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.54 (3.11)	
		95% CI	-14.66, -2.43	
		p-value [b]	0.0063	
		Effect size by Hedges' g (95% CI)	-0.61 (-1.17, -0.05)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.13 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	15 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	77.12 (14.29)	80.32 (11.22)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (65.9%)	10 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	73.37 (13.36)	71.43 (13.28)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.44 (1.72)	-0.53 (2.71)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.91 (3.20)	
		95% CI	-8.21, 4.39	
		p-value [b]	0.5504	
	Effect size by Hedges' g (95% CI)	-0.14 (-0.72, 0.44)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.13 (AMNOG)

QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects	N	21	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	21 (100.0%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	73.47 (14.25)	72.73 (18.08)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (85.7%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	71.96 (18.18)	74.40 (18.87)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.21 (2.17)	1.77 (3.01)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.98 (3.70)	
		95% CI	-11.27, 3.31	
		p-value [b]	0.2835	
Effect size by Hedges' g (95% CI)		-0.25 (-0.96, 0.47)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.13 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluatable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluatable subjects	N	50	31
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at baseline	n(%)	50 (100.0%)	30 (96.8%)
	Observed mean domain score at baseline	Mean (SD)	58.57 (25.03)	53.25 (23.17)
	Number (percentage) of HRQoL evaluatable subjects with non-missing score at Week 24	n(%)	37 (74.0%)	23 (74.2%)
	Observed mean domain score at Week 24	Mean (SD)	66.09 (21.82)	60.04 (23.83)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.02 (2.45)	2.82 (3.20)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.80 (4.04)	
		95% CI	-9.75, 6.15	
		p-value [b]	0.6561	
		Effect size by Hedges' g (95% CI)	-0.07 (-0.52, 0.38)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1827	

Note: "HRQoL evaluatable population" are ITT subjects with an evaluatable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluatable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluatable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.13 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	36 (100.0%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.84 (25.65)	53.63 (23.51)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (72.2%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	64.29 (26.68)	59.72 (22.71)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	4.27 (2.86)	0.92 (3.93)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	3.36 (4.86)	
		95% CI	-6.21, 12.92	
		p-value [b]	0.4901	
		Effect size by Hedges' g (95% CI)	0.13 (-0.42, 0.68)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.13 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	15 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	60.57 (21.73)	71.90 (18.95)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (65.9%)	10 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	59.44 (23.02)	56.19 (26.92)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.20 (2.69)	-0.61 (4.29)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.59 (5.04)	
		95% CI	-10.53, 9.34	
		p-value [b]	0.9063	
	Effect size by Hedges' g (95% CI)	-0.03 (-0.61, 0.56)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.13 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects	N	21	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	20 (95.2%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.00 (24.35)	55.63 (21.91)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (85.7%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	44.18 (27.82)	57.74 (30.30)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-8.40 (3.54)	3.71 (4.77)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-12.10 (5.93)	
		95% CI	-23.80, -0.41	
		p-value [b]	0.0425	
		Effect size by Hedges' g (95% CI)	-0.50 (-1.23, 0.23)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.13 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects	N	50	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	42 (84.0%)	27 (87.1%)
	Observed mean domain score at baseline	Mean (SD)	64.02 (21.89)	57.05 (20.68)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	34 (68.0%)	22 (71.0%)
	Observed mean domain score at Week 24	Mean (SD)	63.84 (19.40)	61.23 (22.90)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.78 (2.19)	2.57 (2.69)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.36 (3.48)	
		95% CI	-13.21, 0.50	
		p-value [b]	0.0689	
		Effect size by Hedges' g (95% CI)	-0.29 (-0.77, 0.19)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6814	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.13 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (88.9%)	18 (94.7%)
	Observed mean domain score at baseline	Mean (SD)	56.46 (20.28)	52.91 (22.00)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	22 (61.1%)	11 (57.9%)
	Observed mean domain score at Week 24	Mean (SD)	59.90 (21.76)	64.48 (21.72)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.20 (2.49)	5.31 (3.29)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-8.51 (4.12)	
		95% CI	-16.63, -0.39	
		p-value [b]	0.0399	
		Effect size by Hedges' g (95% CI)	-0.40 (-0.97, 0.17)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.13 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	12 (80.0%)
	Observed mean domain score at baseline	Mean (SD)	62.73 (20.34)	66.33 (22.54)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (63.4%)	8 (53.3%)
	Observed mean domain score at Week 24	Mean (SD)	59.97 (17.44)	59.02 (15.21)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.84 (2.20)	-3.09 (3.97)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.75 (4.52)	
		95% CI	-10.66, 7.16	
		p-value [b]	0.6990	
	Effect size by Hedges' g (95% CI)	-0.08 (-0.72, 0.55)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.13 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects	N	21	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	19 (90.5%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	48.93 (17.68)	55.75 (16.32)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	17 (81.0%)	7 (63.6%)
	Observed mean domain score at Week 24	Mean (SD)	49.11 (22.60)	66.34 (24.25)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.05 (3.01)	3.29 (3.90)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.34 (4.92)	
		95% CI	-16.03, 3.35	
		p-value [b]	0.1987	
		Effect size by Hedges' g (95% CI)	-0.36 (-1.09, 0.37)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.13 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects	N	50	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	42 (84.0%)	26 (83.9%)
	Observed mean domain score at baseline	Mean (SD)	61.23 (22.73)	55.38 (21.53)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	34 (68.0%)	22 (71.0%)
	Observed mean domain score at Week 24	Mean (SD)	64.94 (19.97)	61.24 (21.74)
	LS mean change from baseline at Week 24			
Within-group	LS mean change from baseline (SE)		-1.45 (2.26)	4.14 (2.87)
Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)		-5.59 (3.66)	
	95% CI		-12.80, 1.62	
	p-value [b]		0.1281	
	Effect size by Hedges' g (95% CI)		-0.25 (-0.73, 0.24)	
	p-value (Subgroup*Treatment Interaction Coeff) [c]		0.5128	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.13 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (88.9%)	18 (94.7%)
	Observed mean domain score at baseline	Mean (SD)	55.43 (21.32)	53.37 (21.89)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	22 (61.1%)	11 (57.9%)
	Observed mean domain score at Week 24	Mean (SD)	61.61 (23.17)	62.22 (21.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	0.17 (2.58)	3.90 (3.39)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.73 (4.26)	
		95% CI	-12.12, 4.66	
		p-value [b]	0.3823	
		Effect size by Hedges' g (95% CI)	-0.17 (-0.74, 0.40)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.13 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	12 (80.0%)
	Observed mean domain score at baseline	Mean (SD)	61.65 (19.31)	71.16 (17.79)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	26 (63.4%)	8 (53.3%)
	Observed mean domain score at Week 24	Mean (SD)	60.16 (18.28)	58.53 (21.66)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.13 (2.28)	-2.96 (4.06)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.17 (4.64)	
		95% CI	-9.31, 8.96	
		p-value [b]	0.9702	
	Effect size by Hedges' g (95% CI)	-0.01 (-0.64, 0.62)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.13 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects	N	21	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	19 (90.5%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.03 (19.54)	55.69 (18.65)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	17 (81.0%)	7 (63.6%)
	Observed mean domain score at Week 24	Mean (SD)	46.19 (24.19)	63.44 (27.04)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.50 (3.07)	3.84 (4.00)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-9.34 (5.03)	
		95% CI	-19.27, 0.59	
		p-value [b]	0.0650	
Effect size by Hedges' g (95% CI)		-0.47 (-1.21, 0.26)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.13 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO < 100 U/L	Number of HRQoL evaluable subjects	N	50	31
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	48 (96.0%)	30 (96.8%)
	Observed mean domain score at baseline	Mean (SD)	58.82 (21.53)	53.03 (21.69)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	36 (72.0%)	23 (74.2%)
	Observed mean domain score at Week 24	Mean (SD)	62.69 (23.00)	60.13 (22.63)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.89 (2.35)	1.70 (2.99)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.59 (3.81)	
		95% CI	-12.09, 2.92	
		p-value [b]	0.2300	
		Effect size by Hedges' g (95% CI)	-0.21 (-0.66, 0.24)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5440	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.13 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 100- < 200 U/L	Number of HRQoL evaluable subjects	N	36	19
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	34 (94.4%)	19 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	51.63 (22.00)	49.34 (18.41)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	25 (69.4%)	12 (63.2%)
	Observed mean domain score at Week 24	Mean (SD)	58.68 (22.17)	59.45 (19.13)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.91 (2.76)	3.89 (3.68)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.80 (4.60)	
		95% CI	-13.85, 4.25	
		p-value [b]	0.2973	
		Effect size by Hedges' g (95% CI)	-0.23 (-0.78, 0.33)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.13 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO 200-500 U/L	Number of HRQoL evaluable subjects	N	41	15
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	41 (100.0%)	15 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.10 (19.70)	63.10 (19.76)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	27 (65.9%)	10 (66.7%)
	Observed mean domain score at Week 24	Mean (SD)	52.18 (20.52)	51.60 (20.65)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.30 (2.52)	-3.06 (3.99)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.24 (4.71)	
		95% CI	-9.52, 9.03	
		p-value [b]	0.9586	
	Effect size by Hedges' g (95% CI)	-0.01 (-0.60, 0.57)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.13 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Baseline serum EPO > 500 U/L	Number of HRQoL evaluable subjects	N	21	11
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	20 (95.2%)	11 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	51.78 (18.73)	48.17 (16.69)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	18 (85.7%)	8 (72.7%)
	Observed mean domain score at Week 24	Mean (SD)	49.76 (22.61)	58.02 (26.97)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.19 (3.34)	5.99 (4.49)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.18 (5.58)	
		95% CI	-21.18, 0.81	
		p-value [b]	0.0693	
		Effect size by Hedges' g (95% CI)	-0.55 (-1.28, 0.18)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.14 (AMNOG)  
 QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	135 (97.8%)	65 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.70 (20.03)	51.73 (22.95)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	101 (73.2%)	46 (70.8%)
	Observed mean domain score at Week 24	Mean (SD)	54.08 (21.87)	55.43 (23.96)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.12 (1.90)	2.58 (2.70)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.69 (3.30)	
		95% CI	-12.20, 0.82	
		p-value [b]	0.0864	
		Effect size by Hedges' g (95% CI)	-0.27 (-0.57, 0.03)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3640	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.14 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	11 (100.0%)	10 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.41 (28.55)	45.00 (17.87)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	56.94 (31.94)	39.58 (16.61)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.70 (4.53)	-4.93 (4.89)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.23 (6.66)	
		95% CI	-12.89, 13.35	
		p-value [b]	0.9725	
	Effect size by Hedges' g (95% CI)	0.01 (-0.81, 0.83)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.14 (AMNOG)  
 QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	135 (97.8%)	65 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.98 (33.20)	50.60 (30.24)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	99 (71.7%)	46 (70.8%)
	Observed mean domain score at Week 24	Mean (SD)	54.43 (32.28)	62.08 (32.25)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.67 (2.74)	1.92 (3.92)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.58 (4.79)	
		95% CI	-15.03, 3.86	
		p-value [b]	0.2450	
		Effect size by Hedges' g (95% CI)	-0.17 (-0.47, 0.12)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8834	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.14 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	11 (100.0%)	10 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.58 (40.92)	48.89 (22.35)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	51.85 (39.28)	46.30 (31.75)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.78 (7.58)	1.46 (8.08)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.23 (11.08)	
		95% CI	-29.06, 14.59	
		p-value [b]	0.5144	
	Effect size by Hedges' g (95% CI)	-0.21 (-1.03, 0.62)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.14 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	127 (92.0%)	58 (89.2%)
	Observed mean domain score at baseline	Mean (SD)	47.83 (37.14)	46.98 (38.35)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	93 (67.4%)	41 (63.1%)
	Observed mean domain score at Week 24	Mean (SD)	48.92 (34.95)	54.88 (38.00)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.75 (2.57)	1.24 (3.82)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.99 (4.61)	
		95% CI	-15.07, 3.10	
		p-value [b]	0.1952	
		Effect size by Hedges' g (95% CI)	-0.16 (-0.47, 0.15)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.0486	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.14 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	10 (90.9%)	9 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	62.50 (44.49)	44.44 (32.54)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	38.89 (33.33)	37.50 (30.62)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-18.77 (6.95)	7.58 (7.49)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-26.35 (10.23)	
		95% CI	-46.51, -6.20	
		p-value [b]	0.0106	
Effect size by Hedges' g (95% CI)		-0.64 (-1.52, 0.24)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.14 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	130 (94.2%)	61 (93.8%)
	Observed mean domain score at baseline	Mean (SD)	62.37 (34.78)	60.38 (39.49)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	96 (69.6%)	44 (67.7%)
	Observed mean domain score at Week 24	Mean (SD)	66.84 (34.16)	61.74 (36.89)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.88 (2.67)	1.43 (3.85)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.46 (4.68)	
		95% CI	-8.77, 9.68	
		p-value [b]	0.9226	
		Effect size by Hedges' g (95% CI)	0.01 (-0.29, 0.32)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.5236	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.14 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	11 (100.0%)	9 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	59.09 (39.17)	74.07 (37.14)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	55.56 (44.29)	56.94 (38.16)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.58 (6.98)	-0.34 (7.74)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.24 (10.43)	
		95% CI	-26.80, 14.31	
		p-value [b]	0.5503	
	Effect size by Hedges' g (95% CI)	-0.16 (-1.00, 0.69)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.14 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	138 (100.0%)	65 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	75.43 (14.15)	72.67 (14.37)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	101 (73.2%)	46 (70.8%)
	Observed mean domain score at Week 24	Mean (SD)	73.46 (14.35)	77.12 (14.53)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.86 (1.18)	1.73 (1.71)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.59 (2.08)	
		95% CI	-8.69, -0.48	
		p-value [b]	0.0286	
		Effect size by Hedges' g (95% CI)	-0.32 (-0.62, -0.03)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.7019	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.14 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	11 (100.0%)	10 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	77.06 (18.05)	78.10 (11.49)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	71.96 (19.55)	76.98 (13.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.82 (2.98)	2.42 (3.20)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.24 (4.37)	
		95% CI	-14.85, 2.36	
		p-value [b]	0.1542	
	Effect size by Hedges' g (95% CI)	-0.39 (-1.22, 0.44)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.14 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	137 (99.3%)	64 (98.5%)
	Observed mean domain score at baseline	Mean (SD)	56.73 (23.96)	56.21 (23.42)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	100 (72.5%)	46 (70.8%)
	Observed mean domain score at Week 24	Mean (SD)	61.24 (24.37)	56.83 (24.08)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.14 (1.77)	1.77 (2.56)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.91 (3.11)	
		95% CI	-8.04, 4.23	
		p-value [b]	0.5400	
		Effect size by Hedges' g (95% CI)	-0.08 (-0.38, 0.22)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9585	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.14 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	11 (100.0%)	10 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	56.49 (28.21)	61.67 (19.00)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	48.94 (32.28)	67.86 (23.56)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.54 (4.80)	1.00 (5.12)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-1.54 (7.01)	
		95% CI	-15.36, 12.28	
		p-value [b]	0.8264	
	Effect size by Hedges' g (95% CI)	-0.06 (-0.88, 0.76)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.14 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	125 (90.6%)	58 (89.2%)
	Observed mean domain score at baseline	Mean (SD)	59.00 (20.71)	56.58 (20.68)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	91 (65.9%)	41 (63.1%)
	Observed mean domain score at Week 24	Mean (SD)	59.65 (19.81)	63.09 (20.87)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.97 (1.46)	2.39 (2.14)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.36 (2.59)	
		95% CI	-10.48, -0.23	
		p-value [b]	0.0405	
		Effect size by Hedges' g (95% CI)	-0.26 (-0.57, 0.05)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.1747	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.14 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	10 (90.9%)	9 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	63.43 (28.08)	57.91 (18.71)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	55.04 (26.15)	51.46 (18.74)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-10.62 (4.07)	2.98 (4.37)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-13.60 (5.97)	
		95% CI	-25.37, -1.83	
		p-value [b]	0.0237	
	Effect size by Hedges' g (95% CI)	-0.54 (-1.42, 0.34)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.14 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	125 (90.6%)	57 (87.7%)
	Observed mean domain score at baseline	Mean (SD)	58.03 (20.80)	56.83 (21.30)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	91 (65.9%)	41 (63.1%)
	Observed mean domain score at Week 24	Mean (SD)	60.32 (20.91)	60.67 (21.59)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.65 (1.56)	2.68 (2.30)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.34 (2.78)	
		95% CI	-9.83, 1.16	
		p-value [b]	0.1212	
		Effect size by Hedges' g (95% CI)	-0.21 (-0.52, 0.11)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2950	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.14 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	10 (90.9%)	9 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	59.81 (27.28)	59.11 (17.57)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	51.99 (28.64)	59.66 (20.73)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-7.01 (4.16)	3.77 (4.47)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.79 (6.11)	
		95% CI	-22.82, 1.25	
		p-value [b]	0.0788	
	Effect size by Hedges' g (95% CI)	-0.44 (-1.32, 0.43)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.14 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Mutated	Number of HRQoL evaluable subjects	N	138	65
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	133 (96.4%)	64 (98.5%)
	Observed mean domain score at baseline	Mean (SD)	55.37 (20.49)	52.95 (20.30)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	98 (71.0%)	46 (70.8%)
	Observed mean domain score at Week 24	Mean (SD)	57.05 (21.66)	58.11 (21.84)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.54 (1.68)	2.14 (2.39)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.68 (2.92)	
		95% CI	-10.43, 1.08	
		p-value [b]	0.1105	
		Effect size by Hedges' g (95% CI)	-0.23 (-0.53, 0.07)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8680	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.14 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
SF3B1 mutation status Non-mutated	Number of HRQoL evaluable subjects	N	11	10
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	11 (100.0%)	10 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.83 (26.51)	51.85 (15.20)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	9 (81.8%)	6 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	52.58 (30.70)	51.25 (17.82)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.86 (4.45)	-0.28 (4.75)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.59 (6.51)	
		95% CI	-16.41, 9.24	
		p-value [b]	0.5822	
	Effect size by Hedges' g (95% CI)	-0.16 (-0.98, 0.67)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.15 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	46 (97.9%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	51.63 (24.10)	44.38 (20.47)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	38 (80.9%)	14 (70.0%)
	Observed mean domain score at Week 24	Mean (SD)	49.34 (24.31)	55.36 (24.37)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.76 (2.59)	3.28 (3.86)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.04 (4.64)	
		95% CI	-15.17, 3.09	
		p-value [b]	0.1941	
		Effect size by Hedges' g (95% CI)	-0.26 (-0.78, 0.26)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3917	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.15 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	33 (97.1%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	60.98 (18.42)	50.00 (23.39)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (70.6%)	8 (88.9%)
	Observed mean domain score at Week 24	Mean (SD)	63.02 (21.01)	67.19 (19.97)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.66 (2.93)	6.14 (5.11)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-10.80 (5.91)	
		95% CI	-22.44, 0.84	
		p-value [b]	0.0688	
	Effect size by Hedges' g (95% CI)	-0.54 (-1.28, 0.19)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.15 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	21 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.78 (20.06)	58.33 (20.67)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (71.9%)	15 (71.4%)
	Observed mean domain score at Week 24	Mean (SD)	51.09 (24.11)	42.50 (23.05)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.03 (2.92)	-3.36 (3.74)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	0.33 (4.75)	
		95% CI	-9.02, 9.68	
		p-value [b]	0.9445	
		Effect size by Hedges' g (95% CI)	0.02 (-0.53, 0.56)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.1.15 (AMNOG)

QoL-E Physical Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	30 (96.8%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.58 (14.86)	53.75 (27.83)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	21 (67.7%)	13 (65.0%)
	Observed mean domain score at Week 24	Mean (SD)	60.12 (17.51)	60.58 (24.39)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.28 (3.06)	2.11 (3.77)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.39 (4.86)	
		95% CI	-12.95, 6.17	
		p-value [b]	0.4851	
		Effect size by Hedges' g (95% CI)	-0.16 (-0.72, 0.40)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.15 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	46 (97.9%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	51.21 (36.40)	39.44 (26.36)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	36 (76.6%)	14 (70.0%)
	Observed mean domain score at Week 24	Mean (SD)	47.53 (34.69)	45.24 (32.75)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.91 (4.06)	1.35 (6.14)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.26 (7.35)	
		95% CI	-19.74, 9.23	
		p-value [b]	0.4754	
		Effect size by Hedges' g (95% CI)	-0.15 (-0.67, 0.37)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9974	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.15 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (94.1%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	62.15 (35.68)	72.84 (31.48)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (70.6%)	8 (88.9%)
	Observed mean domain score at Week 24	Mean (SD)	67.59 (27.00)	72.22 (30.86)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.19 (4.76)	5.40 (8.61)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.58 (9.78)	
		95% CI	-25.86, 12.69	
		p-value [b]	0.5017	
		Effect size by Hedges' g (95% CI)	-0.19 (-0.91, 0.54)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.15 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	21 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.90 (30.20)	48.15 (25.17)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (71.9%)	15 (71.4%)
	Observed mean domain score at Week 24	Mean (SD)	51.21 (32.61)	57.04 (32.76)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.29 (4.71)	1.79 (5.91)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.08 (7.56)	
		95% CI	-21.99, 7.82	
		p-value [b]	0.3499	
		Effect size by Hedges' g (95% CI)	-0.25 (-0.79, 0.30)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.2.15 (AMNOG)

QoL-E Functional Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.48 (31.73)	55.56 (33.04)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	21 (67.7%)	13 (65.0%)
	Observed mean domain score at Week 24	Mean (SD)	57.14 (33.01)	82.91 (18.49)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.97 (4.90)	4.01 (5.99)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.97 (7.74)	
		95% CI	-22.21, 8.27	
		p-value [b]	0.3683	
		Effect size by Hedges' g (95% CI)	-0.21 (-0.77, 0.34)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.15 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	43 (91.5%)	18 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	51.16 (38.17)	44.44 (37.92)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	34 (72.3%)	13 (65.0%)
	Observed mean domain score at Week 24	Mean (SD)	41.18 (35.82)	51.92 (38.81)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.42 (3.74)	5.54 (5.68)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-11.96 (6.80)	
		95% CI	-25.37, 1.44	
		p-value [b]	0.0799	
		Effect size by Hedges' g (95% CI)	-0.31 (-0.86, 0.24)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.2514	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.15 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (91.2%)	8 (88.9%)
	Observed mean domain score at baseline	Mean (SD)	70.16 (30.56)	59.38 (42.13)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (67.6%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	64.13 (31.83)	71.43 (26.73)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-6.68 (4.34)	-1.92 (7.70)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.76 (8.79)	
		95% CI	-22.10, 12.58	
		p-value [b]	0.5886	
	Effect size by Hedges' g (95% CI)	-0.14 (-0.90, 0.62)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.15 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (90.6%)	18 (85.7%)
	Observed mean domain score at baseline	Mean (SD)	38.79 (35.09)	56.94 (32.99)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	22 (68.8%)	13 (61.9%)
	Observed mean domain score at Week 24	Mean (SD)	46.59 (36.43)	50.00 (36.80)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.21 (4.32)	-3.07 (5.69)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	2.86 (7.18)	
		95% CI	-11.29, 17.01	
		p-value [b]	0.6910	
		Effect size by Hedges' g (95% CI)	0.08 (-0.50, 0.66)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.3.15 (AMNOG)

QoL-E Social and Family Life - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (93.5%)	19 (95.0%)
	Observed mean domain score at baseline	Mean (SD)	35.34 (37.51)	40.79 (39.27)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	19 (61.3%)	12 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	48.68 (31.70)	56.25 (41.46)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-7.05 (4.47)	5.75 (5.43)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-12.80 (7.01)	
		95% CI	-26.61, 1.01	
		p-value [b]	0.0691	
		Effect size by Hedges' g (95% CI)	-0.33 (-0.90, 0.24)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.15 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	42 (89.4%)	18 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	54.76 (38.04)	68.52 (37.98)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	36 (76.6%)	14 (70.0%)
	Observed mean domain score at Week 24	Mean (SD)	62.73 (40.11)	60.12 (40.19)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	6.84 (3.89)	-1.35 (5.75)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	8.19 (6.95)	
		95% CI	-5.51, 21.88	
		p-value [b]	0.2401	
		Effect size by Hedges' g (95% CI)	0.21 (-0.33, 0.76)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.3201	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.15 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (94.1%)	7 (77.8%)
	Observed mean domain score at baseline	Mean (SD)	73.18 (34.31)	83.33 (30.05)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	21 (61.8%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	74.60 (32.11)	70.24 (43.53)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	5.55 (4.36)	1.51 (8.56)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	4.04 (9.57)	
		95% CI	-14.83, 22.91	
		p-value [b]	0.6735	
		Effect size by Hedges' g (95% CI)	0.12 (-0.68, 0.92)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.15 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	20 (95.2%)
	Observed mean domain score at baseline	Mean (SD)	62.50 (34.33)	44.17 (40.83)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (71.9%)	15 (71.4%)
	Observed mean domain score at Week 24	Mean (SD)	64.49 (32.68)	54.44 (38.56)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.16 (4.27)	3.31 (5.56)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.46 (6.99)	
		95% CI	-17.24, 10.32	
		p-value [b]	0.6211	
		Effect size by Hedges' g (95% CI)	-0.09 (-0.64, 0.46)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.4.15 (AMNOG)

QoL-E Sexual Well-being - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	30 (96.8%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	64.72 (28.08)	68.33 (37.52)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	21 (67.7%)	12 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	72.22 (27.17)	68.06 (30.74)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-5.50 (4.51)	1.70 (5.47)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.20 (7.09)	
		95% CI	-21.15, 6.75	
		p-value [b]	0.3106	
		Effect size by Hedges' g (95% CI)	-0.22 (-0.78, 0.34)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.15 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	47 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	74.97 (14.81)	69.29 (15.49)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	38 (80.9%)	14 (70.0%)
	Observed mean domain score at Week 24	Mean (SD)	70.93 (14.30)	74.49 (15.13)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.07 (1.62)	3.27 (2.47)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.34 (2.96)	
		95% CI	-13.16, -1.51	
		p-value [b]	0.0137	
		Effect size by Hedges' g (95% CI)	-0.48 (-1.01, 0.04)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9540	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.15 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	34 (100.0%)	9 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	76.75 (15.50)	78.31 (16.34)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (70.6%)	8 (88.9%)
	Observed mean domain score at Week 24	Mean (SD)	78.57 (11.75)	84.52 (10.72)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.53 (1.84)	5.44 (3.34)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.97 (3.81)	
		95% CI	-14.47, 0.53	
		p-value [b]	0.0685	
		Effect size by Hedges' g (95% CI)	-0.44 (-1.16, 0.29)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.15 (AMNOG)  
 QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	21 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	74.26 (14.15)	71.88 (12.95)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (71.9%)	15 (71.4%)
	Observed mean domain score at Week 24	Mean (SD)	71.22 (16.21)	74.60 (14.58)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.63 (1.88)	2.85 (2.39)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.48 (3.04)	
		95% CI	-11.46, 0.51	
		p-value [b]	0.0729	
	Effect size by Hedges' g (95% CI)	-0.39 (-0.94, 0.15)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.5.15 (AMNOG)

QoL-E Fatigue - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	76.65 (13.21)	76.90 (12.39)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	21 (67.7%)	13 (65.0%)
	Observed mean domain score at Week 24	Mean (SD)	73.24 (16.25)	84.98 (7.49)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.90 (1.95)	1.90 (2.42)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.80 (3.11)	
		95% CI	-11.92, 0.32	
		p-value [b]	0.0634	
		Effect size by Hedges' g (95% CI)	-0.44 (-1.00, 0.12)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.15 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	46 (97.9%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	57.09 (23.26)	57.50 (21.41)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	37 (78.7%)	14 (70.0%)
	Observed mean domain score at Week 24	Mean (SD)	56.37 (24.98)	57.82 (22.30)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.69 (2.61)	2.41 (3.90)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-0.72 (4.69)	
		95% CI	-9.96, 8.53	
		p-value [b]	0.8789	
		Effect size by Hedges' g (95% CI)	-0.03 (-0.55, 0.49)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.6220	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.15 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	34 (100.0%)	8 (88.9%)
	Observed mean domain score at baseline	Mean (SD)	64.22 (22.91)	68.75 (18.33)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (70.6%)	8 (88.9%)
	Observed mean domain score at Week 24	Mean (SD)	70.44 (20.55)	64.58 (23.69)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.83 (2.96)	0.18 (5.71)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	1.65 (6.40)	
		95% CI	-10.96, 14.26	
		p-value [b]	0.7969	
		Effect size by Hedges' g (95% CI)	0.07 (-0.68, 0.83)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.15 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	21 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	55.06 (23.81)	56.80 (23.44)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (71.9%)	15 (71.4%)
	Observed mean domain score at Week 24	Mean (SD)	55.90 (25.75)	55.56 (21.63)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.98 (2.99)	2.12 (3.79)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.10 (4.83)	
		95% CI	-16.61, 2.41	
		p-value [b]	0.1427	
	Effect size by Hedges' g (95% CI)	-0.30 (-0.84, 0.25)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.6.15 (AMNOG)

QoL-E MDS-specific symptoms - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	31 (100.0%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	50.92 (24.98)	55.24 (23.73)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	21 (67.7%)	13 (65.0%)
	Observed mean domain score at Week 24	Mean (SD)	65.65 (25.53)	65.57 (24.76)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	1.02 (3.12)	3.52 (3.84)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-2.50 (4.94)	
		95% CI	-12.23, 7.23	
		p-value [b]	0.6130	
		Effect size by Hedges' g (95% CI)	-0.10 (-0.65, 0.45)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.15 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	42 (89.4%)	18 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	57.46 (22.27)	53.20 (21.72)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	32 (68.1%)	13 (65.0%)
	Observed mean domain score at Week 24	Mean (SD)	53.66 (20.73)	58.75 (22.27)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.66 (2.24)	4.42 (3.32)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.08 (4.00)	
		95% CI	-14.96, 0.81	
		p-value [b]	0.0783	
		Effect size by Hedges' g (95% CI)	-0.32 (-0.86, 0.23)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.8802	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.15 (AMNOG)

QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	30 (88.2%)	8 (88.9%)
	Observed mean domain score at baseline	Mean (SD)	68.73 (19.60)	66.67 (19.75)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (67.6%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	69.17 (17.00)	75.33 (13.96)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.32 (2.59)	2.65 (4.63)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.96 (5.27)	
		95% CI	-16.36, 4.43	
		p-value [b]	0.2593	
		Effect size by Hedges' g (95% CI)	-0.30 (-1.06, 0.47)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.15 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (90.6%)	18 (85.7%)
	Observed mean domain score at baseline	Mean (SD)	57.06 (20.61)	57.37 (19.42)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	22 (68.8%)	13 (61.9%)
	Observed mean domain score at Week 24	Mean (SD)	57.47 (22.07)	54.72 (21.34)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.51 (2.55)	0.97 (3.32)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.48 (4.19)	
		95% CI	-11.73, 4.77	
		p-value [b]	0.4062	
		Effect size by Hedges' g (95% CI)	-0.17 (-0.75, 0.41)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.7.15 (AMNOG)  
 QoL-E General - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (93.5%)	19 (95.0%)
	Observed mean domain score at baseline	Mean (SD)	56.95 (19.26)	58.12 (21.62)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	19 (61.3%)	12 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	63.00 (16.69)	71.54 (16.80)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.68 (2.60)	2.74 (3.17)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-7.42 (4.10)	
		95% CI	-15.50, 0.67	
		p-value [b]	0.0720	
		Effect size by Hedges' g (95% CI)	-0.36 (-0.93, 0.21)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.15 (AMNOG)  
 QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
 (HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	42 (89.4%)	18 (90.0%)
	Observed mean domain score at baseline	Mean (SD)	57.50 (20.83)	55.11 (20.99)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	32 (68.1%)	13 (65.0%)
	Observed mean domain score at Week 24	Mean (SD)	54.77 (21.86)	58.32 (21.41)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.51 (2.35)	4.14 (3.44)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.65 (4.17)	
		95% CI	-12.87, 3.56	
		p-value [b]	0.2653	
		Effect size by Hedges' g (95% CI)	-0.22 (-0.77, 0.33)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9562	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.15 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	30 (88.2%)	7 (77.8%)
	Observed mean domain score at baseline	Mean (SD)	67.50 (19.39)	69.48 (19.42)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (67.6%)	7 (77.8%)
	Observed mean domain score at Week 24	Mean (SD)	69.32 (17.72)	71.34 (17.64)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.36 (2.68)	0.69 (5.10)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.05 (5.72)	
		95% CI	-14.33, 8.23	
		p-value [b]	0.5942	
		Effect size by Hedges' g (95% CI)	-0.15 (-0.96, 0.65)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.15 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (90.6%)	18 (85.7%)
	Observed mean domain score at baseline	Mean (SD)	55.99 (20.56)	57.65 (19.63)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	22 (68.8%)	13 (61.9%)
	Observed mean domain score at Week 24	Mean (SD)	56.60 (22.25)	55.93 (20.51)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-3.87 (2.65)	2.65 (3.45)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-6.52 (4.35)	
		95% CI	-15.09, 2.05	
		p-value [b]	0.1350	
		Effect size by Hedges' g (95% CI)	-0.32 (-0.90, 0.26)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.8.15 (AMNOG)

QoL-E All - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	29 (93.5%)	19 (95.0%)
	Observed mean domain score at baseline	Mean (SD)	53.52 (21.29)	56.56 (21.80)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	19 (61.3%)	12 (60.0%)
	Observed mean domain score at Week 24	Mean (SD)	64.52 (20.40)	69.90 (18.47)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.36 (2.69)	3.95 (3.30)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-4.31 (4.26)	
		95% CI	-12.70, 4.08	
		p-value [b]	0.3125	
		Effect size by Hedges' g (95% CI)	-0.20 (-0.77, 0.37)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.15 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment < 6 months	Number of HRQoL evaluable subjects	N	47	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	45 (95.7%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	54.14 (21.57)	47.11 (18.97)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	35 (74.5%)	14 (70.0%)
	Observed mean domain score at Week 24	Mean (SD)	52.08 (22.66)	52.81 (22.72)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-1.50 (2.48)	4.03 (3.66)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.53 (4.42)	
		95% CI	-14.23, 3.18	
		p-value [b]	0.2123	
		Effect size by Hedges' g (95% CI)	-0.26 (-0.78, 0.26)	
		p-value (Subgroup*Treatment Interaction Coeff) [c]	0.9769	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.15 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment 6-12 months	Number of HRQoL evaluable subjects	N	34	9
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (94.1%)	8 (88.9%)
	Observed mean domain score at baseline	Mean (SD)	62.84 (19.44)	65.51 (21.27)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	24 (70.6%)	8 (88.9%)
	Observed mean domain score at Week 24	Mean (SD)	67.02 (19.18)	68.00 (22.09)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-2.57 (2.86)	1.05 (5.37)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.63 (6.05)	
		95% CI	-15.55, 8.30	
		p-value [b]	0.5496	
		Effect size by Hedges' g (95% CI)	-0.18 (-0.94, 0.58)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.15 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment >12-24 months	Number of HRQoL evaluable subjects	N	32	21
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	32 (100.0%)	21 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	53.91 (20.80)	54.43 (17.77)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	23 (71.9%)	15 (71.4%)
	Observed mean domain score at Week 24	Mean (SD)	52.73 (22.44)	51.70 (18.79)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-4.70 (2.80)	0.47 (3.53)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-5.17 (4.51)	
		95% CI	-14.05, 3.72	
		p-value [b]	0.2530	
	Effect size by Hedges' g (95% CI)	-0.26 (-0.80, 0.29)		

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018

Table 14.2.18.6.1.2.9.15 (AMNOG)

QoL-E Treatment Outcome Index - Least Square Mean Change from Baseline from Mixed Model Repeated Measures Analysis (Model 1[a])  
(HRQoL Evaluable Population by Subgroup)

Subgroup	Description	Statistics	Luspatercept+BSC	Placebo+BSC
Longest duration of prior ESA treatment > 24 months	Number of HRQoL evaluable subjects	N	31	20
	Number (percentage) of HRQoL evaluable subjects with non-missing score at baseline	n(%)	30 (96.8%)	20 (100.0%)
	Observed mean domain score at baseline	Mean (SD)	52.81 (19.78)	54.85 (22.37)
	Number (percentage) of HRQoL evaluable subjects with non-missing score at Week 24	n(%)	21 (67.7%)	13 (65.0%)
	Observed mean domain score at Week 24	Mean (SD)	60.97 (21.36)	69.68 (16.60)
	LS mean change from baseline at Week 24			
	Within-group	LS mean change from baseline (SE)	-0.53 (2.94)	2.74 (3.58)
	Difference (Luspatercept+BSC vs. Placebo+BSC)	LS mean difference (SE)	-3.27 (4.63)	
		95% CI	-12.39, 5.86	
		p-value [b]	0.4812	
		Effect size by Hedges' g (95% CI)	-0.15 (-0.71, 0.40)	

Note: "HRQoL evaluable population" are ITT subjects with an evaluable QLQ-C30 questionnaire at C1D1 (baseline) visit and at least one post-baseline visit. If assessment at the C1D1 visit is missing, the assessment at the screening visit is used for the C1D1 visit. A QLQ-C30 is considered evaluable for a given visit if at least half of the QLQ-C30 items are answered. Only assessment data through Week 25 are considered in the analyses (End of Treatment visit is excluded). Within-group least square (LS) mean change and between-group difference in LS mean change are estimated from a mixed effects model. Changes in scores from baseline (screening) are used as the dependent variable. The model assumes an unstructured covariance unless convergence was not met and in which case compound-symmetry covariance assumption was used. The model includes the following covariates: baseline domain score, treatment group (luspatercept+BSC vs. placebo+BSC), time (continuous; coded as number of weeks since C1D1) and a treatment by time interaction as fixed effects, with intercept and time as random effects. Effect size (95% CI) for between-group difference is estimated using Hedges' g. BSC = Best Supportive Care; CI = Confidence Interval; EORTC = European Organisation for Research and Treatment of Cancer; HRQoL = Health-related Quality of Life; LS = Least Square; NE = Not Evaluable; QLQ-C30 = Quality of Life Questionnaire - Core 30; SD = Standard Deviation; SE = Standard Error.

[a] Model 1 includes assessment data through Week 24 (excluding End of Treatment visit).

[b] The between-group p-value is calculated based on a random intercept/slope model with covariates listed above.

[c] The interaction p-value is calculated for the subgroup by treatment interaction coefficient in the model.

Program Path: P:\EVA-25150-00\B - Team\_Folder\13 - Statistics\MEDALIST\WL\SASPrograms

Data Source: ADQOLE

Run Date: 17AUG2019

Data Extraction Date: 21JUN2018