

Sotatercept (pulmonale arterielle Hypertonie)

Addendum zum Projekt A24-96
(Dossierbewertung)

A horizontal bar composed of 18 rectangular segments of varying shades of blue and grey. The text 'ADDENDUM (DOSSIERBEWERTUNG)' is centered in white on a dark blue segment.

ADDENDUM (DOSSIERBEWERTUNG)

Projekt: A25-14

Version: 1.0

Stand: 14.02.2025

IQWiG-Berichte – Nr. 1937

DOI: 10.60584/A25-14

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Sotatercept (pulmonale arterielle Hypertonie) – Addendum zum Projekt A24-96

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

28.01.2025

Interne Projektnummer

A25-14

DOI-URL

<https://doi.org/10.60584/A25-14>

Anschrift des Herausgebers

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Siegburger Str. 237
50679 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Zitiervorschlag

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Sotatercept (pulmonale arterielle Hypertonie); Addendum zum Projekt A24-96 (Dossierbewertung) [online]. 2025 [Zugriff: TT.MM.JJJJ]. URL: <https://doi.org/10.60584/A25-14>.

Schlagwörter

Sotatercept, Pulmonale Hypertonie, Nutzenbewertung, NCT04576988

Keywords

Sotatercept, Pulmonary Arterial Hypertension, Benefit Assessment, NCT04576988

An dem Addendum beteiligte Mitarbeiterinnen und Mitarbeiter des IQWiG

- Christian Siebel
- Moritz Felsch
- Prateek Mishra
- Daniela Preukschat
- Claudia Selbach

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	v
Abkürzungsverzeichnis.....	vi
1 Hintergrund.....	1
2 Bewertung.....	2
2.1 Bewertung der Studie STELLAR	2
2.1.1 Studiencharakteristika.....	2
2.1.2 Ergebnisse.....	6
2.1.2.1 Dargestellte Endpunkte	6
2.1.2.2 Verzerrungspotenzial.....	10
2.1.2.3 Ergebnisse	12
2.1.2.4 Subgruppen und andere Effektmodifikatoren.....	16
2.1.3 Zusammenfassung der Ergebnisse	19
3 Literatur	20
Anhang A Gründe, warum bei Studienteilnehmenden mit Mono- oder Zweifachtherapie keine zusätzliche medikamentöse Therapie zu Studienbeginn eingesetzt wurde	22
Anhang B Ergebnisse zu Nebenwirkungen	23

Tabellenverzeichnis

	Seite
Tabelle 1: Endpunktübergreifendes Verzerrungspotenzial (Studienebene) – RCT, direkter Vergleich: Sotatercept vs. Placebo	5
Tabelle 2: Matrix der Endpunkte – RCT, direkter Vergleich: Sotatercept vs. Placebo.....	7
Tabelle 3: Endpunktübergreifendes und endpunktspezifisches Verzerrungspotenzial – RCT, direkter Vergleich: Sotatercept vs. Placebo	11
Tabelle 4: Ergebnisse (Mortalität, Morbidität, Gesundheitsbezogene Lebensqualität, Nebenwirkungen) – RCT, direkter Vergleich: Sotatercept vs. Placebo	13
Tabelle 5: Ergebnisse (Morbidität, stetig) – RCT, direkter Vergleich: Sotatercept vs. Placebo.....	14
Tabelle 6: Subgruppen (Morbidität, stetig) – RCT, direkter Vergleich: Sotatercept vs. Placebo.....	17
Tabelle 7: Subgruppen (Morbidität, dichotom) – RCT, direkter Vergleich: Sotatercept vs. Placebo.....	18
Tabelle 8: Gründe, warum bei Studienteilnehmenden mit Mono- oder Zweifachtherapie keine zusätzliche Arzneimitteltherapie eingesetzt wurde – RCT, direkter Vergleich: Sotatercept vs. Placebo	22
Tabelle 9: Häufige UEs – RCT, direkter Vergleich: Sotatercept vs. Placebo.....	24
Tabelle 10: Häufige SUEs – RCT, direkter Vergleich: Sotatercept vs. Placebo.....	25
Tabelle 11: Abbruch wegen UEs – RCT, direkter Vergleich: Sotatercept vs. Placebo.....	26

Abkürzungsverzeichnis

Abkürzung	Bedeutung
6MWT	6-Minutes Walking Test (6-Minuten-Gehtest)
G-BA	Gemeinsamer Bundesausschuss
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
MedDRA	Medical Dictionary for regulatory Activities (Medizinisches Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung)
PAH	pulmonale arterielle Hypertonie
PT	Preferred Term (bevorzugter Begriff)
pU	pharmazeutischer Unternehmer
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
SGB	Sozialgesetzbuch
SUE	schwerwiegendes unerwünschtes Ereignis
UE	unerwünschtes Ereignis
VAS	visuelle Analogskala
WHO	Weltgesundheitsorganisation

1 Hintergrund

Der Gemeinsame Bundesausschuss (G-BA) hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) am 28.01.2025 mit ergänzenden Bewertungen zum Projekt A24-96 (Sotatercept – Nutzenbewertung gemäß § 35a SGB V [1]) beauftragt.

Der Auftrag umfasst die Bewertung der Studie STELLAR unter Berücksichtigung der im Dossier [2] und durch den pharmazeutischen Unternehmer (pU) im Stellungnahmeverfahren [3,4] vorgelegten Daten.

Die Verantwortung für die vorliegende Bewertung und das Bewertungsergebnis liegt ausschließlich beim IQWiG. Die Bewertung wird dem G-BA übermittelt. Über den Zusatznutzen beschließt der G-BA.

2 Bewertung

Die Fragestellung der Nutzenbewertung in A24-96 [1] war die Bewertung des Zusatznutzens von Sotatercept in Kombination mit anderen Therapien gegen pulmonale arterielle Hypertonie (PAH) im Vergleich mit der zweckmäßigen Vergleichstherapie zur Verbesserung der körperlichen Leistungsfähigkeit bei erwachsenen Patientinnen und Patienten mit PAH mit der Weltgesundheitsorganisation(WHO)-Funktionsklasse II bis III.

In seinem Dossier legte der pU Ergebnisse zur Gesamtpopulation der randomisierten kontrollierten Studie (RCT) STELLAR vor [2]. In der Studie STELLAR erhielten etwa 40 % der Patientinnen und Patienten in beiden Studienarmen eine Therapie mit den parenteral zu verabreichenden Prostazyklin-Analoga Treprostinil und Epoprostenol. Die parenteralen Prostazyklin-Analoga Treprostinil und Epoprostenol waren jedoch nicht von der vom G-BA festgelegten zweckmäßigen Vergleichstherapie umfasst. Somit war für einen relevanten Teil der Patientinnen und Patienten in der Studie STELLAR die vom G-BA festgelegte zweckmäßige Vergleichstherapie nicht umgesetzt. Darüber hinaus bestanden unabhängig von einer Therapie mit parenteralen Prostazyklin-Analoga Unsicherheiten, ob die patientenindividuelle Therapie in der Studie STELLAR adäquat umgesetzt wurde. Zudem lagen keine Informationen zur jeweiligen Dosierung der im Rahmen der Hintergrundtherapie angewendeten Wirkstoffe vor, sodass nicht geprüft werden konnte, ob die Wirkstoffe in der Studie zulassungskonform dosiert wurden.

Aufgrund der fehlenden Umsetzung der zweckmäßigen Vergleichstherapie für einen relevanten Teil der eingeschlossenen Patientinnen und Patienten war die Studie STELLAR nicht für die Nutzenbewertung geeignet. Im Folgenden wird auftragsgemäß die Gesamtpopulation der Studie STELLAR bewertet.

2.1 Bewertung der Studie STELLAR

2.1.1 Studiencharakteristika

Die Charakterisierung der Studie STELLAR und Charakterisierung der Studienpopulation findet sich in der Dossierbewertung A24-96 [1].

Optimierung der Therapie im Vergleichsarm

Innerhalb des Stellungnahmeverfahrens sowie im Nachgang zur mündlichen Anhörung hat der pU weitere Daten zur patientenindividuellen Therapie in der Studie STELLAR nachgereicht [3,4]. Der pU legte mit seiner Stellungnahme unter anderem Angaben zur PAH-Hintergrundtherapie in Abhängigkeit von der Risikostratifizierung vor, aus denen hervorgeht, dass etwa ein Viertel der Patientinnen und Patienten im Interventions- vs. Vergleichsarm einem intermediär-niedrigem bis hohem Risiko zuzuordnen sind und trotz Nicht-Erreichen des Therapieziels (Niedrig-Risiko-Status) lediglich eine Zweifachtherapie erhielten. Für diese

Patienten blieb zunächst weiterhin unklar, ob für diese nicht, z. B. durch Hinzunahme einer 3. Wirkstoffkomponente (auch abseits der parenteralen Prostazykline), eine Therapieoptimierung zu Studienbeginn möglich gewesen wäre. Im Nachgang zur mündlichen Anhörung reichte der pU Daten nach, aus denen hervorgeht, dass für einen Großteil der Patientinnen und Patienten, die zu Studienbeginn eine Mono- oder Zweifachtherapie erhielten, die Hinzunahme einer weiteren medikamentösen Therapie entweder nicht empfohlen war oder diese aufgrund von Sicherheits- / Verträglichkeitsbedenken nicht verabreicht wurden (siehe Tabelle 8 in Anhang A). Demgegenüber war eine zusätzliche medikamentöse Therapie bei 8 Patientinnen und Patienten im Interventionsarm (4,9 %) und 15 Patientinnen und Patienten im Vergleichsarm (9,4 %) nicht erstattungsfähig oder nicht verfügbar, weswegen in der Gesamtbetrachtung weiterhin eine Unsicherheit verbleibt, die bei der Einschätzung der Aussagesicherheit berücksichtigt wird (siehe unten).

Zur Dosierung der im Rahmen der Hintergrundtherapie angewendeten Wirkstoffe legt der pU weiterhin keine Angaben vor. Jedoch wurde im Stellungnahmeverfahren betont, dass dies im vorliegenden Anwendungsgebiet von untergeordneter Bedeutung sei, da die Dosierung der eingesetzten Endothelin-Rezeptor-Antagonisten und Phosphodiesterase-Typ-5-Inhibitoren standardisiert sei [5,6]. Prostazyklin-Analoga, selektive Prostazyklin-Rezeptor-Agonisten und Riociguat werden dagegen individuell dosiert [7-10]. Es liegen keine Daten vor, die darauf hinweisen, dass nicht gemäß Fachinformation dosiert wurde. Die nachgereichten Angaben zu 19 Patientinnen und Patienten mit Notwendigkeit einer Therapieeskalation mit einer zugelassenen PAH-Hintergrundtherapie oder Erhöhung der Prostazyklindosis um mindestens 10 % deuten hingegen darauf hin, dass die Wirkstoffe weitgehend gemäß Fachinformation dosiert wurden.

Die verbleibenden Unsicherheiten bezüglich der Umsetzung der patientenindividuellen Therapie werden bei der Einschätzung der Aussagesicherheit berücksichtigt (siehe Abschnitt 2.1.2.2).

Auswertungszeiträume

In der Studie war die primäre Analyse der Endpunkte zu Wirksamkeit und Nebenwirkungen zu dem Zeitpunkt geplant, zu dem alle Patientinnen und Patienten die 24-wöchige primäre Behandlungsphase beendet hatten. Der Datenschnitt für diese primäre Analyse erfolgte am 26.08.2022. Der finale Datenschnitt erfolgte zu Studienende am 06.12.2022 (Datum der letzten Visite der letzten Patientin oder des letzten Patienten). Beide Datenschnitte waren präspezifiziert. Nach Abschluss der primären 24-wöchigen Behandlungsphase wurden die Patientinnen und Patienten in den jeweiligen Studienarmen der Studie STELLAR bis zu 72 Wochen weiter behandelt und bis zu 80 Wochen weiter beobachtet, bis die letzte eingeschlossene Patientin bzw. der letzte eingeschlossene Patient die primäre Behandlungsphase abgeschlossen hatte (Studienende). Aufgrund des Studiendesigns war die

Behandlungs- und Beobachtungsdauer in der Studie STELLAR patientenindividuell unterschiedlich. Mit Ausnahme der patientenberichteten Endpunkte zur Symptomatik, Gesundheitszustand und gesundheitsbezogenen Lebensqualität wurden die Endpunkte in der Studie auch nach Woche 24 weiter erhoben.

Der pU legt in Modul 4 A sowie mit den nachgereichten Daten Auswertungen des finalen Datenschnitts vom 06.12.2022 vor, die sich je nach Endpunkt auf 2 verschiedene Zeiträume beziehen:

- Auswertungszeitraum: zu Woche 24
- Auswertungszeitraum: gesamter Beobachtungszeitraum bis zum Studienende

Für das vorliegende Addendum wird der gesamte Beobachtungszeitraum basierend auf dem finalen Datenschnitt der Studie betrachtet. Für die Endpunkte Gehfähigkeit und Dyspnoe sowie die patientenberichteten Endpunkte zur Symptomatik, Gesundheitszustand sowie gesundheitsbezogenen Lebensqualität legt der pU Auswertungen zum Auswertungszeitraum zu Woche 24 vor. Die patientenberichteten Endpunkte zur Symptomatik, Gesundheitszustand sowie gesundheitsbezogenen Lebensqualität wurden zu Behandlungsbeginn und zu Woche 24 erhoben, sodass hier der längste vorhandene Beobachtungszeitraum vorliegt. Die Endpunkte Gehfähigkeit und Dyspnoe wurden zu Woche 24 bei 97 % bzw. 92 % der Patientinnen und Patienten im Interventions- und Vergleichsarm, zu Woche 36 noch bei 90 % bzw. 76 % und zu Woche 48 nur noch bei 51 % bzw. 33 % der Patientinnen und Patienten erhoben. Ungeachtet der deutlich geringeren Rücklaufquoten deuten die vom pU in Anhang 4 G dargestellten deskriptiven Analysen darauf hin, dass sich die Ergebnisse nach Woche 24 nicht relevant ändern. Deshalb werden für die Endpunkte Gehfähigkeit und Dyspnoe genauso wie für die patientenberichteten Endpunkte zur Symptomatik, Gesundheitszustand sowie gesundheitsbezogenen Lebensqualität die Auswertungen zu Woche 24 betrachtet.

Endpunktübergreifendes Verzerrungspotenzial (Studienebene)

Tabelle 1 zeigt das endpunktübergreifende Verzerrungspotenzial (Verzerrungspotenzial auf Studienebene).

Tabelle 1: Endpunktübergreifendes Verzerrungspotenzial (Studienebene) – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Verzerrungspotenzial auf Studienebene
			Patientinnen und Patienten	Behandelnde Personen			
STELLAR	ja	ja	ja	ja	ja	ja	niedrig
RCT: randomisierte kontrollierte Studie							

Das endpunktübergreifende Verzerrungspotenzial wird für die Studie STELLAR als niedrig eingestuft.

Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext

Der pU gibt an, dass es sich bei der Studie STELLAR mit 91 Studienzentren in 21 Ländern um eine internationale und multizentrische Studie handelt. In Europa wurden Patientinnen und Patienten in 40 Zentren eingeschlossen. Neun Zentren in Deutschland hätten 71 Patientinnen und Patienten eingeschlossen, dies entspreche 22 % der Patientinnen und Patienten der Studie STELLAR.

Der pU gibt weiter an, dass die deutliche Mehrheit der Patientinnen und Patienten weiß sei (90,2 % im Sotatercept-Arm und 88,1 % im Placebo-Arm). PAH betreffe dem pU zufolge weitaus häufiger Frauen, dementsprechend seien in der Studie mehrheitlich Frauen eingeschlossen. Die Patientinnen und Patienten seien seit bis zu 40 Jahren an PAH erkrankt, die mediane Erkrankungsdauer betrage 7,26 Jahre. Dies entspreche dem pU zufolge der großen Bandbreite, die sich auch bei deutschen PAH-Patientinnen und -Patienten zeige, die für eine Behandlung mit Sotatercept gemäß Zulassung infrage kommen. Insgesamt entsprächen laut pU die krankheitsspezifischen Charakteristika der Studienpopulation der deutschen Zielpopulation.

Laut Einschätzung des pU erfolgte die Erhebung der Endpunkte standardisiert und mit niedrigem Verzerrungspotential. Es sei somit davon auszugehen, dass die in der Studie STELLAR beobachteten Behandlungseffekte auf den deutschen Versorgungskontext übertragbar sind.

Der pU legt keine weiteren Informationen zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext vor.

2.1.2 Ergebnisse

2.1.2.1 Dargestellte Endpunkte

Im vorliegenden Addendum werden folgende patientenrelevante Endpunkte dargestellt:

- Mortalität
 - Gesamtmortalität
- Morbidität
 - Gehfähigkeit, erhoben mit dem 6-Minuten-Gehtest (6MWT)
 - Symptomatik, erhoben mit dem PAH-SYMPACT
 - Dyspnoe, erhoben mit der Borg CR10 Skala
 - Gesundheitszustand, erhoben mit der visuellen Analogskala (VAS) des EQ-5D
- Gesundheitsbezogene Lebensqualität
 - erhoben mit dem PAH-SYMPACT
- Nebenwirkungen
 - schwerwiegende UEs (SUEs)
 - Abbruch wegen UEs
 - gegebenenfalls weitere spezifische UEs

Die Auswahl der patientenrelevanten Endpunkte weicht von der Auswahl des pU ab, der im Dossier (Modul 4 A) weitere Endpunkte heranzieht.

Tabelle 2 zeigt, für welche Endpunkte in der Studie Daten zur Verfügung stehen.

Tabelle 2: Matrix der Endpunkte – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie	Endpunkte									
	Gesamtmortalität ^a	Gehfähigkeit (6MWT)	Symptomatik (PAH-SYMPACT)	Dyspnoe (Borg CR10 Skala)	Gesundheitszustand (EQ-5D VAS)	Gesundheitsbezogene Lebensqualität (PAH-SYMPACT)	SUEs	Abbruch wegen UEs	Augenerkrankungen (SOC, UEs)	Nasenbluten (PT, UEs)
STELLAR	ja	ja	ja	ja	ja	ja	ja	ja	ja	ja
a. Die Ergebnisse zur Gesamtmortalität beruhen auf den Angaben zu tödlich verlaufenen UEs. 6MWT: 6-Minuten-Gehtest; CR10: 10 Punkte Category Ratio Scale; PAH: pulmonale arterielle Hypertonie; PT: bevorzugter Begriff; RCT: randomisierte kontrollierte Studie; SOC: Systemorganklasse; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis; VAS: visuelle Analogskala										

Anmerkungen zu den Endpunkten

Gesamtmortalität

Der pU legt in Modul 4 A des Dossiers Auswertungen zur Gesamtmortalität zu Woche 24 vor. In der Studie STELLAR wurde der Überlebensstatus jedoch auch über Woche 24 hinaus bis zum Studienende erhoben. Im Nachgang zur mündlichen Anhörung hat der pU Auswertungen nachgereicht, die den gesamten Studienzeitraum umfassen. Im vorliegenden Addendum werden die Angaben zu Todesfällen über die gesamte Studiendauer dargestellt. Der Unterschied zwischen den Beobachtungszeiten in beiden Behandlungsarmen (45,3 Wochen vs. 39,0 Wochen), lässt die Betrachtung des relativen Risikos zu. Die sich aus den unterschiedlichen Beobachtungsdauern – insbesondere infolge des vorzeitigen Wechsels von Patientinnen und Patienten aus der Studie STELLAR in die Extensionsstudie SOTERIA – ergebende Unsicherheit wird bei der Bewertung des Verzerrungspotenzials berücksichtigt (siehe Abschnitt 2.1.2.2). Die vom pU im Nachgang zur Anhörung vorgelegten Ereigniszeitanalysen zum Gesamtüberleben in der Studie STELLAR zeigen ebenfalls keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen.

Zusätzlich zu der Auswertung bis Woche 24 legt der pU im Modul 4 A des Dossiers ergänzende Ereigniszeitanalysen zum Gesamtüberleben vor, in die neben den Beobachtungen über den kompletten Studienverlauf der Studie STELLAR auch die Erhebungen des Überlebensstatus in der 1-armigen Extensionsstudie SOTERIA bis zum Datenschnitt 08.11.2023 eingehen. Der pU legt dabei sowohl eine Analyse vor, in der für den Behandlungswechsel von Patientinnen und

Patienten aus dem Placeboarm der Studie STELLAR auf eine Behandlung mit Sotatercept in der Extensionsstudie SOTERIA adjustiert wird (Rank Preserving Structural Failure Time Modell), als auch eine Analyse ohne Adjustierung. Zu beiden Auswertungen legt der pU jedoch keine ausreichenden Informationen für die Bewertung vor. So bleibt unklar, wie viele Patientinnen und Patienten aus beiden Behandlungsarmen in die Extensionsstudie SOTERIA übergangen und wie lange die Patientinnen und Patienten in der Extensionsstudie beobachtet wurden. Zudem ist anzumerken, dass es sich bei dem vom pU dargestellten Datenschnitt vom 08.11.2023 nicht um einen präspezifizierten Datenschnitt handelt. Die ergänzenden Analysen werden daher nicht dargestellt.

Gefähigkeit (6MWT)

Für den Endpunkt Gefähigkeit erhoben mittels des 6MWT führt der pU für das Dossier Responderanalysen zur Verbesserung zu Woche 24 im Vergleich zu Studienbeginn durch. Dabei verwendet der pU das post hoc spezifizierte Responsekriterium ≥ 40 m. Für den 6MWT ist in der vorliegenden Indikation der PAH keine Responseschwelle etabliert [11-14]. Zudem war das Responsekriterium ≥ 40 m nicht prädefiniert. Des Weiteren legt der pU für den Endpunkt Gefähigkeit erhoben mittels 6MWT eine Auswertung des medianen Unterschieds zwischen den Behandlungsarmen zu Woche 24 gegenüber Studienbeginn vor (Hodges-Lehman Location Shift). Die Erhebungszeitpunkte zwischen Studienbeginn und Woche 24 (Woche 3 und 12) gehen in diese Auswertung nicht ein. Hierbei handelt es sich um die primär geplante Analyse des primären Endpunkts in der Studie STELLAR. Im vorliegenden Addendum wird die präspezifizierte Auswertung des medianen Unterschieds zwischen den Behandlungsarmen zu Woche 24 dargestellt.

Kombinierter Endpunkt Zeit bis zur klinischen Verschlechterung oder Tod

Der vom pU vorgelegte kombinierte Endpunkt Zeit bis zur klinischen Verschlechterung oder Tod umfasst die folgenden Komponenten:

- Tod
- Listung für eine Lungen und / oder Herztransplantation aufgrund von Krankheitsprogression
- Notwendigkeit einer Therapieeskalation mit einer zugelassenen PAH-Hintergrundtherapie oder die Erhöhung der Dosis der Prostazyklin-Infusion um ≥ 10 %
- Notwendigkeit einer Atrioseptostomie
- Hospitalisierung aufgrund von Krankheitsprogression (≥ 24 Stunden)
- Verschlechterung der PAH definiert durch eine Verschlechterung der WHO-Funktionsklasse und eine Verschlechterung des 6MWT um ≥ 15 % im Vergleich zum Studienbeginn

Damit ein kombinierter Endpunkt berücksichtigt werden kann, ist es erforderlich, dass die einzelnen Komponenten des Endpunkts patientenrelevant sind. Eine alleinige Anpassung der Therapie der PAH ist nicht per se patientenrelevant. Symptome, die eine Therapieanpassung erfordern sowie mögliche Nachteile einer Therapieanpassung sollten sich dagegen in anderen patientenrelevanten Endpunkten widerspiegeln. Die Teilkomponente Notwendigkeit einer Therapieeskalation beeinflusst zudem maßgeblich das Ergebnis des kombinierten Endpunkts. Der kombinierte Endpunkt wird im vorliegenden Addendum daher nicht dargestellt.

Der pU legt in Modul 4 A des Dossiers zusätzlich eine Sensitivitätsanalyse vor, die die Komponenten Tod, Listung für eine Lungen und / oder Herztransplantation aufgrund von Krankheitsprogression, Notwendigkeit einer Atriostomie, Hospitalisierung aufgrund von Krankheitsprogression (≥ 24 Stunden) und eine Verschlechterung der PAH definiert durch eine Verschlechterung des 6MWT um ≥ 40 m umfasst. Hierbei handelt es sich um eine post hoc für das Dossier spezifizierte Operationalisierung des kombinierten Endpunkts. Das Ergebnis des kombinierten Endpunkts in dieser Operationalisierung wird maßgeblich durch Ereignisse in der Teilkomponente Verschlechterung des 6MWT um ≥ 40 m beeinflusst. Eine Verschlechterung des 6MWT wird durch die Betrachtung der Veränderung im Endpunkt Gehfähigkeit bereits abgebildet. Wie zuvor für den Endpunkt Gehfähigkeit beschrieben, ist anzumerken, dass das Responsekriterium ≥ 40 m keine etablierte Responsechwelle im Anwendungsgebiet darstellt und nicht präspezifiziert war. Die Einzelkomponenten Listung für eine Lungen und / oder Herztransplantation aufgrund von Krankheitsprogression, Notwendigkeit einer Atriostomie sowie Hospitalisierung aufgrund von Krankheitsprogression (≥ 24 Stunden) werden als patientenrelevant erachtet. Der pU legt für die Einzelkomponenten ausschließlich Ergebnisse zu den qualifizierenden Ereignissen vor, nicht jedoch zu allen aufgetretenen Ereignissen in den Einzelkomponenten. Für die Einzelkomponenten liegen somit keine geeigneten Daten vor.

Symptomatik und gesundheitsbezogene Lebensqualität (PAH-SYMPACT)

Beim PAH-SYMPACT handelt es sich um einen validierten Fragebogen zur Erhebung der Symptomatik und gesundheitsbezogenen Lebensqualität bei Patientinnen und Patienten mit PAH [15,16]. Der PAH-SYMPACT setzt sich zusammen aus 2 Domänen zur Symptomatik (Domänen kardiopulmonale Symptome und kardiovaskuläre Symptome) und 2 Domänen zu den Auswirkungen der Erkrankung (Domänen körperliche Beeinträchtigungen und kognitive / emotionale Beeinträchtigungen). Hinzu kommt 1 weitere Frage zur Verwendung von Sauerstoff. Die Fragen zur Symptomatik wurden täglich über einen Zeitraum von 7 Tagen vor der jeweiligen Visite von der Patientin bzw. dem Patienten beantwortet. Die Fragen zu Auswirkungen der Erkrankung wurden am letzten Tag des 7-tägigen Zeitraums beantwortet. Der Schweregrad der Symptome bzw. die Schwere der Auswirkungen der Erkrankung wurden auf einer 5-Punkt-Skala bewertet (0: keine; 4: sehr schwer). Der pU legt für die einzelnen Domänen Responderanalysen zur Verbesserung zu Woche 24 gegenüber Studienbeginn unter

Verwendung des Responsekriteriums 15 % der Skalenspannweite vor. Die in die Studie STELLAR eingeschlossenen Patientinnen und Patienten waren zu Beginn der Studie symptomatisch (WHO-Funktionsklasse II bis III) und zeigten körperliche sowie kognitive / emotionale Beeinträchtigungen. Da eine zusätzliche Behandlung mit Sotatercept somit prinzipiell eine Verbesserung der Symptomatik und der gesundheitsbezogenen Lebensqualität bewirken kann, werden jeweils die Auswertungen zur Verbesserung zu Woche 24 betrachtet.

Dyspnoe (Borg CR10 Skala)

In der Studie STELLAR wurde die von den Patientinnen bzw. den Patienten empfundene Dyspnoe mittels einer CR10 Skala nach Borg erhoben. Dabei sollten die Patientinnen und Patienten ihre empfundene Dyspnoe jeweils vor und nach Durchführung des 6MWT auf der Borg CR10 Skala angeben (0: „überhaupt keine Dyspnoe“, 10: „extreme Dyspnoe“). Der pU legt in Modul 4 A des Dossiers Responderanalysen zur Verbesserung der Dyspnoe zu Woche 24 im Vergleich zu Studienbeginn unter Verwendung des Responsekriteriums 15 % der Skalenspannweite vor. Der pU berücksichtigt dabei ausschließlich die Erhebung des Dyspnoe-Empfindens vor Durchführung des 6MWT. Wie bereits zuvor beschrieben, waren die in die Studie STELLAR eingeschlossenen Patientinnen und Patienten zu Beginn der Studie symptomatisch. Für die Borg CR10 Skala wird daher die Auswertung zur Verbesserung zu Woche 24 betrachtet.

Nebenwirkungen

In die Auswertungen der UEs gehen Ereignisse ein, die sowohl Nebenwirkungen als auch Symptome oder Folgekomplikationen der Grunderkrankung zugeordnet werden können. Da jedoch in den Gesamtraten SUEs und Abbrüche wegen UEs insgesamt nur wenige Ereignisse eingehen, die eindeutig der Grunderkrankung zuzuordnen sind, bleibt dies für das vorliegende Addendum ohne Konsequenz.

Schwere UEs

In der Studie STELLAR erfolgte die Einschätzung der Schwere der UEs nicht anhand einer etablierten Klassifikation, sondern nach vom pU definierten Kategorien. Dies ist keine adäquate Operationalisierung des Schweregrads. Deshalb wird der Endpunkt nicht dargestellt.

2.1.2.2 Verzerrungspotenzial

Tabelle 3 beschreibt das Verzerrungspotenzial für die Ergebnisse der relevanten Endpunkte.

Tabelle 3: Endpunktübergreifendes und endpunktspezifisches Verzerrungspotenzial – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie	Studienebene	Endpunkte									
		Gesamtmortalität ^a	Gehfähigkeit (6MWT)	Symptomatik (PAH-SYMPACT)	Dyspnoe (Borg CR10 Skala)	Gesundheitszustand (EQ-5D VAS)	Gesundheitsbezogene Lebensqualität (PAH-SYMPACT)	SUEs	Abbruch wegen UEs	Augenerkrankungen (SOC, UEs)	Nasenbluten (PT, UEs)
STELLAR	N	H ^b	N	H ^c	N	H ^c	H ^c	H ^b	N ^d	H ^b	H ^b

a. Die Ergebnisse zur Gesamtmortalität beruhen auf den Angaben zu tödlich verlaufenen UEs.
 b. unvollständige Beobachtungen aus potenziell informativen Gründen
 c. hoher Anteil an Patientinnen und Patienten (> 10 %), die nicht in der Auswertung berücksichtigt wurden
 d. Trotz niedrigen Verzerrungspotenzials wird für den Endpunkt Abbruch wegen UEs von einer eingeschränkten Ergebnissicherheit ausgegangen.

6MWT: 6-Minuten-Gehtest; CR10: 10 Punkte Category Ratio Scale; H: hoch; N: niedrig; PAH: pulmonale arterielle Hypertonie; PT: bevorzugter Begriff; RCT: randomisierte kontrollierte Studie;
 SOC: Systemorganklasse; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis;
 VAS: visuelle Analogskala

Das endpunktspezifische Verzerrungspotenzial wird für die Ergebnisse zu den Endpunkten Gesamtmortalität, SUEs, Augenerkrankungen und Nasenbluten aufgrund unvollständiger Beobachtungen aus potenziell informativen Gründen als hoch bewertet. Es ist beispielsweise unklar, wie viele Patientinnen und Patienten aufgrund einer klinischen Verschlechterung vorzeitig aus der Studie STELLAR in die Extensionsstudie SOTERIA wechselten und zu welchem Zeitpunkt dies erfolgte. Für die Ergebnisse der patientenberichteten Endpunkte Symptomatik, Gesundheitszustand und gesundheitsbezogene Lebensqualität ergibt sich das hohe Verzerrungspotenzial aus dem hohen Anteil an Patientinnen und Patienten, die nicht in der Auswertung berücksichtigt wurden.

Für die Endpunkte Gehfähigkeit und Dyspnoe zu Woche 24 wird das Verzerrungspotenzial der Ergebnisse als niedrig eingeschätzt.

Für den Endpunkt Abbruch wegen UE ist trotz eines niedrigen Verzerrungspotenzials die Ergebnissicherheit eingeschränkt. Ein vorzeitiger Abbruch der Therapie aus anderen Gründen als UEs stellt ein konkurrierendes Ereignis für den zu erfassenden Endpunkt Abbruch wegen UEs dar. Dies bedeutet, dass nach Abbruch der Therapie aus anderen Gründen zwar UEs, die

zum Abbruch geführt hätten, auftreten können, das Kriterium Abbruch für diese jedoch nicht mehr erfassbar ist. Wie viele UEs das betrifft, ist nicht abschätzbar.

Zusammenfassende Einschätzung der Aussagesicherheit

Wie bereits in Abschnitt 2.1.1 beschrieben bestehen weiterhin Unsicherheiten bezüglich der optimalen Einstellung der PAH-Hintergrundtherapie in der Studie STELLAR. Aus den im Rahmen des Stellungnahmeverfahrens gemachten Angaben des pU geht hervor, dass bei 8 Patientinnen und Patienten im Interventionsarm (4,9 %) und 15 Patientinnen und Patienten im Vergleichsarm (9,4 %), die eine Mono- oder Zweifachtherapie zu Studienbeginn erhielten, eine zusätzliche medikamentöse Therapie infrage gekommen wäre. Zudem liegen kaum Angaben zur Dosierung der im Rahmen der Hintergrundtherapie angewendeten Wirkstoffe vor, so dass nur näherungsweise und in einer kleinen Stichprobe geprüft werden konnte, ob die Wirkstoffe in der Studie zulassungskonform dosiert wurden. Die Aussagesicherheit der Studienergebnisse ist daher insgesamt reduziert.

2.1.2.3 Ergebnisse

Tabelle 4 und Tabelle 5 fassen die Ergebnisse zum Vergleich von Sotatercept mit Placebo bei Patientinnen und Patienten mit PAH mit der WHO-Funktionsklasse II bis III zusammen. Die Daten aus dem Dossier des pU wurden, wo notwendig, durch eigene Berechnungen ergänzt.

Die Ergebnisse zu häufigen UEs, SUEs und Abbrüchen wegen UEs sind in Anhang B dargestellt.

Tabelle 4: Ergebnisse (Mortalität, Morbidität, Gesundheitsbezogene Lebensqualität, Nebenwirkungen) – RCT, direkter Vergleich: Sotatercept vs. Placebo (mehrseitige Tabelle)

Studie Endpunktkategorie Endpunkt	Sotatercept		Placebo		Sotatercept vs. Placebo RR [95 %-KI]; p-Wert ^a
	N	Patientinnen und Patienten mit Ereignis n (%)	N	Patientinnen und Patienten mit Ereignis n (%)	
STELLAR					
Mortalität					
Gesamt mortalität ^b	163	2 (1,2)	160	7 (4,4)	0,28 [0,06; 1,33]; 0,097 ^c
Morbidität					
Symptomatik (PAH-SYMPACT – Verbesserung zu Woche 24 ^d)					
kardiopulmonale Symptome	115	47 (40,9)	117	35 (29,9)	1,35 [0,95; 1,93]; 0,095
kardiovaskuläre Symptome	115	49 (42,6)	117	34 (29,1)	1,48 [1,04; 2,11]; 0,030
Dyspnoe (Borg CR10 Skala – Verbesserung zu Woche 24 ^e)	160	38 (23,8)	159	37 (23,3)	1,02 [0,69; 1,51]; 0,918
Gesundheitszustand (EQ-5D VAS – Verbesserung zu Woche 24 ^f)	124	29 (23,4)	126	20 (15,9)	1,49 [0,89; 2,49]; 0,131
Gesundheitsbezogene Lebensqualität					
PAH-SYMPACT – Verbesserung zu Woche 24 ^d					
körperliche Beeinträchtigungen	117	39 (33,3)	123	31 (25,2)	1,31 [0,87; 1,96]; 0,193
kognitive / emotionale Beeinträchtigung	117	30 (25,6)	123	30 (24,4)	1,04 [0,67; 1,60]; 0,866
Nebenwirkungen					
UEs (ergänzend dargestellt)	163	151 (92,6)	160	149 (93,1)	–
SUEs	163	40 (24,5)	160	47 (29,4)	0,84 [0,58; 1,20]; 0,529 ^c
Abbruch wegen UEs	163	6 (3,7)	160	11 (6,7)	0,54 [0,20; 1,41]; 0,246 ^c
Augenerkrankungen (SOC, UEs) ^g	163	21 (12,9)	160	7 (4,4)	2,94 [1,29; 6,73]; 0,007 ^c
Nasenbluten (PT, UEs)	163	36 (22,1)	160	3 (1,9)	11,78 [3,70; 37,48]; < 0,001 ^c
a. Wenn nichts anderes angegeben: Mantel-Haenszel-Schätzung, stratifiziert nach WHO-Funktionsklasse (Klasse II vs. III) und PAH-Hintergrundtherapie (Mono/Doppel- vs. Dreifachtherapie); p-Wert des Wald-Tests b. Die Ergebnisse zur Gesamt mortalität beruhen auf den Angaben zu tödlich verlaufenen UEs über den gesamten Studienzeitraum. c. eigene Berechnung von RR, 95 %-KI (asymptotisch) und p-Wert (unbedingter exakter Test, CSZ-Methode nach [17])					

Tabelle 4: Ergebnisse (Mortalität, Morbidität, Gesundheitsbezogene Lebensqualität, Nebenwirkungen) – RCT, direkter Vergleich: Sotatercept vs. Placebo (mehrseitige Tabelle)

Studie Endpunktkategorie Endpunkt	Sotatercept		Placebo		Sotatercept vs. Placebo RR [95 %-KI]; p-Wert ^a
	N	Patientinnen und Patienten mit Ereignis n (%)	N	Patientinnen und Patienten mit Ereignis n (%)	
d. Eine Abnahme um $\geq 0,6$ Punkte im Vergleich zum Studienbeginn wird als klinisch relevante Verbesserung angesehen (Wertebereich der Skala: 0 bis 4). e. Eine Abnahme um $\geq 1,5$ Punkte im Vergleich zum Studienbeginn wird als klinisch relevante Verbesserung angesehen (Wertebereich der Skala: 0 bis 10). f. Eine Zunahme um ≥ 15 Punkte im Vergleich zum Studienbeginn wird als klinisch relevante Verbesserung angesehen (Wertebereich der Skala: 0 bis 100). g. häufige aufgetretene Ereignisse im Interventions- vs. Kontrollarm waren verschwommenes Sehen (4 vs. 0) und Katarakt (4 vs. 0)					
6MWT: 6-Minuten-Gehtest; CR10: 10 Punkte Category Ratio Scale; KI: Konfidenzintervall; n: Anzahl Patientinnen und Patienten mit (mindestens 1) Ereignis; N: Anzahl ausgewerteter Patientinnen und Patienten; PAH: pulmonale arterielle Hypertonie; PT: bevorzugter Begriff; RCT: randomisierte kontrollierte Studie; RR: relatives Risiko; SOC: Systemorganklasse; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis; VAS: visuelle Analogskala; WHO: Weltgesundheitsorganisation					

Tabelle 5: Ergebnisse (Morbidität, stetig) – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie Endpunktkategorie Endpunkt	Sotatercept			Placebo			Sotatercept vs. Placebo Hodges-Lehman Location Shift [95 %-KI]; p-Wert
	N ^a	Werte Studien- beginn [m] Median (Q1; Q3)	Mediane Änderung zu Woche 24 [m] MW (Min; Max) ^b	N ^a	Werte Studien- beginn [m] Median (Q1, Q3)	Mediane Änderung zu Woche 24 [m] MW (Min; Max) ^b	
STELLAR							
Morbidität							
Gehfähigkeit (6MWT) zu Woche 24	163	417,0 (348,0; 464,5)	34,3 (33,0; 35,5)	160	427,1 (365,0; 465,0)	1,0 (-1,0; 3,5)	40,40 [27,28; 53,53]; < 0,001
a. Anzahl der Patientinnen und Patienten, die bei der Effektschätzung berücksichtigt wurden, die Werte zu Studienbeginn (und zu Woche 24) können auf anderen Patientenzahlen basieren. b. Mittelwert, Minimum und Maximum der medianen Änderungen zu Woche 24, die sich aus den durch die multiple Imputation erzeugten Imputationsdatensätzen ergeben							
6MWT: 6-Minuten-Gehtest; KI: Konfidenzintervall; m: Meter; Max: Maximum; Min: Minimum; MW: Mittelwert; N: Anzahl ausgewerteter Patientinnen und Patienten; Q1: 1. Quartil; Q3: 3. Quartil; RCT: randomisierte kontrollierte Studie							

Insgesamt ist die Aussagesicherheit zu allen Endpunkten wie in Abschnitt 2.1.2.2 beschrieben eingeschränkt.

Mortalität

Gesamtmortalität

Für den Endpunkt Gesamtmortalität zeigt sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Morbidität

Gehfähigkeit (6MWT)

Für den Endpunkt Gehfähigkeit, erhoben mittels des 6MWT, zeigt sich zu Woche 24 ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zum Vorteil von Sotatercept. Es liegt jedoch eine Effektmodifikation durch das Merkmal WHO-Funktionsklasse vor. Für Patientinnen und Patienten mit PAH der WHO-Funktionsklasse II zeigt sich ein statistisch signifikanter Unterschied zum Vorteil von Sotatercept im Vergleich zu Placebo, der jedoch nicht als klinisch relevant eingestuft wird. Für Patientinnen und Patienten mit PAH der WHO-Funktionsklasse III zeigt sich hingegen ein statistisch signifikanter und klinisch relevanter Unterschied zwischen den Behandlungsgruppen zum Vorteil von Sotatercept (siehe Abschnitt 2.1.2.4).

Symptomatik (PAH-SYMPACT)

Kardiopulmonale Symptome

Für den Endpunkt kardiopulmonale Symptome zeigt sich zu Woche 24 kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Kardiovaskuläre Symptome

Für den Endpunkt kardiovaskuläre Symptome zeigt sich zu Woche 24 ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zum Vorteil von Sotatercept.

Dyspnoe (Borg CR10 Skala)

Für den Endpunkt Dyspnoe, gemessen mittels der CR10 Skala nach Borg, zeigt sich zu Woche 24 kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Gesundheitszustand (EQ-5D VAS)

Für den Endpunkt Gesundheitszustand, gemessen anhand der EQ-5D VAS, zeigt sich zu Woche 24 kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Allerdings liegt eine Effektmodifikation durch das Merkmal WHO-Funktionsklasse vor. Für Patientinnen und Patienten mit PAH der WHO-Funktionsklasse II zeigt sich ein statistisch signifikanter Unterschied zum Vorteil von Sotatercept im Vergleich zu Placebo. Für Patientinnen und Patienten mit PAH der WHO-Funktionsklasse III zeigt sich hingegen kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (siehe Abschnitt 2.1.2.4).

Gesundheitsbezogene Lebensqualität (PAH-SYMPACT)

Die gesundheitsbezogene Lebensqualität wurde anhand der Domänen körperliche Beeinträchtigungen und kognitive / emotionale Beeinträchtigungen des PAH-SYMPACT erhoben.

Für die Endpunkte körperliche Beeinträchtigungen und kognitive / emotionale Beeinträchtigungen zeigt sich zu Woche 24 jeweils kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Nebenwirkungen

SUEs

Für den Endpunkt SUEs zeigt sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Abbruch wegen UEs

Für den Endpunkt Abbruch wegen UEs zeigt sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

Augenerkrankungen (UEs)

Für den Endpunkt Augenerkrankungen (UEs) zeigt sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zum Nachteil von Sotatercept.

Nasenbluten (UEs)

Für den Endpunkt Nasenbluten (UEs) zeigt sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zum Nachteil von Sotatercept.

2.1.2.4 Subgruppen und andere Effektmodifikatoren

Im vorliegenden Addendum werden folgende Subgruppenmerkmale betrachtet:

- Alter (< 65 Jahre vs. ≥ 65 Jahre)
- Geschlecht (männlich vs. weiblich)
- WHO-Funktionsklasse II vs. III

Der pU legt Subgruppenanalysen nach Alter, Geschlecht und WHO-Funktionsklasse für alle im Dossier aufgeführten Endpunkte mit Ausnahme des Endpunkts Gesamtmortalität vor. Der pU begründet dies damit, dass für den Endpunkt Gesamtmortalität die Anzahl an Ereignissen unterhalb des Schwellenwerts von 10 Ereignissen liegt. Das Vorgehen des pU ist sachgerecht.

Interaktionstests werden durchgeführt, wenn mindestens 10 Patientinnen und Patienten pro Subgruppe in die Analyse eingehen. Bei binären Daten müssen darüber hinaus in mindestens 1 Subgruppe mindestens 10 Ereignisse vorliegen.

Es werden nur die Ergebnisse dargestellt, bei denen eine Effektmodifikation mit einer statistisch signifikanten Interaktion zwischen Behandlung und Subgruppenmerkmal (p -Wert $< 0,05$) vorliegt. Zudem werden ausschließlich Subgruppenergebnisse dargestellt, wenn mindestens in einer Subgruppe ein statistisch signifikanter und relevanter Effekt vorliegt.

Tabelle 6 und Tabelle 7 fassen die Subgruppenergebnisse zum Vergleich von Sotatercept mit Placebo bei Patientinnen und Patienten mit PAH mit der WHO-Funktionsklasse II bis III zusammen.

Tabelle 6: Subgruppen (Morbidität, stetig) – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie Endpunkt Merkmal Subgruppe	Sotatercept			Placebo			Sotatercept vs. Placebo
	N ^a	Werte Studien- beginn [m] Median (Q1; Q3)	Mediane Änderung zu Woche 24 [m] MW (Min; Max) ^b	N ^a	Werte Studien- beginn [m] Median (Q1, Q3)	Mediane Änderung zu Woche 24 [m] MW (Min; Max) ^b	Hodges-Lehman Location Shift [95 %-KI]; p-Wert
STELLAR							
Gehfähigkeit (6MWT) zu Woche 24							
Krankheitsschwere							
WHO-FK II	79	k. A.	k. A.	78	k. A.	k. A.	21,6 [6,67; 36,60]; k. A.
WHO-FK III	84	k. A.	k. A.	82	k. A.	k. A.	60,9 [40,46; 81,35]; k. A.
Gesamt	Interaktion:						0,002 ^c
a. Anzahl der Patientinnen und Patienten, die bei der Effektschätzung berücksichtigt wurden, die Werte zu Studienbeginn (und zu Woche 24) können auf anderen Patientenzahlen basieren. b. Mittelwert, Minimum und Maximum der medianen Änderungen zu Woche 24, die sich aus den durch die multiple Imputation erzeugten Imputationsdatensätzen ergeben c. p-Wert des Cochran's Q-Tests 6MWT: 6-Minuten-Gehtest; k. A.: keine Angabe; KI: Konfidenzintervall; m: Meter; Max: Maximum; Min: Minimum; MW: Mittelwert; N: Anzahl ausgewerteter Patientinnen und Patienten; Q1: 1. Quartil; Q3: 3. Quartil; RCT: randomisierte kontrollierte Studie; WHO: Weltgesundheitsorganisation; WHO-FK: WHO-Funktionsklasse							

Tabelle 7: Subgruppen (Morbidity, dichotom) – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie	Sotatercept		Placebo		Sotatercept vs. Placebo	
	N	Patientinnen und Patienten mit Ereignis n (%)	N	Patientinnen und Patienten mit Ereignis n (%)	RR [95 %-KI]	p-Wert
STELLAR						
Morbidity						
Gesundheitszustand (EQ-5D VAS ^a ; Verbesserung zu Woche 24)						
Krankheitsschwere						
WHO-FK II	65	16 (24,6)	65	5 (7,7)	3,20 [1,24; 8,26] ^b	0,016
WHO-FK III	59	13 (22,0)	61	15 (24,6)	0,91 [0,48; 1,73] ^b	0,764
Gesamt					Interaktion: 0,026 ^c	
a. Eine Zunahme um ≥ 15 Punkte im Vergleich zum Studienbeginn wird als klinisch relevante Verbesserung angesehen (Wertebereich der Skala: 0 bis 100) b. RR: Mantel-Haenszel-Schätzung, stratifiziert nach WHO-Funktionsklasse (Klasse II vs. III) und PAH-Hintergrundtherapie (Mono / Doppel- vs. Dreifachtherapie); p-Wert: Wald-Test c. p-Wert des Likelihood-Ratio-Tests, basierend auf einem (laut pU) linearen Modell mit den Kovariaten Behandlung und Subgruppe sowie der Interaktion zwischen Behandlung und Subgruppe, welches stratifiziert ist nach WHO-Funktionsklasse (Klasse II vs. III) und PAH-Hintergrundtherapie (Mono/Doppel- vs. Dreifachtherapie)						
KI: Konfidenzintervall; n: Anzahl Patientinnen und Patienten mit (mindestens 1) Ereignis; N: Anzahl ausgewerteter Patientinnen und Patienten; PAH: Pulmonale arterielle Hypertonie; RCT: randomisierte kontrollierte Studie; RR: relatives Risiko; SOC: Systemorganklasse; UE: unerwünschtes Ereignis; WHO: Weltgesundheitsorganisation; WHO-FK: WHO-Funktionsklasse						

Morbidity

Gehfähigkeit (6MWT)

Für den Endpunkt Gehfähigkeit (6MWT) liegt eine Effektmodifikation durch das Merkmal WHO-Funktionsklasse vor.

Für Patientinnen und Patienten mit PAH der WHO-Funktionsklasse II zeigt sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zum Vorteil von Sotatercept. Die untere Grenze des 95 %-Konfidenzintervalls liegt jedoch lediglich bei 6,7 m, was nicht als klinisch relevant eingestuft wird.

Für Patientinnen und Patienten mit PAH der WHO-Funktionsklasse III zeigt sich ebenfalls ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Die untere Grenze des 95 %-Konfidenzintervalls liegt bei 40,5 m. Dies wird als relevanter Effekt interpretiert. Es ergibt sich ein Vorteil von Sotatercept gegenüber Placebo.

Gesundheitszustand (EQ-5D VAS)

Für den Endpunkt Gesundheitszustand (EQ-5D VAS) liegt eine Effektmodifikation durch das Merkmal WHO-Funktionsklasse vor. Für Patientinnen und Patienten mit PAH der WHO-Funktionsklasse II zeigt sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zum Vorteil von Sotatercept.

Für Patientinnen und Patienten mit PAH der WHO-Funktionsklasse III zeigt sich hingegen kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

2.1.3 Zusammenfassung der Ergebnisse

Auf Basis der Studie STELLAR ergeben sich für erwachsene Patientinnen und Patienten mit PAH mit der WHO-Funktionsklasse II bis III die folgenden Vor- und Nachteile auf Endpunktebene:

- Vorteil von Sotatercept gegenüber Placebo für den Endpunkt Gehfähigkeit (6MWT) zu Woche 24 in der Subgruppe der Patientinnen und Patienten mit PAH der WHO-Funktionsklasse III
- Vorteil von Sotatercept gegenüber Placebo für den Endpunkt Gesundheitszustand (EQ-5D VAS) zu Woche 24 in der Subgruppe der Patientinnen und Patienten mit PAH der WHO-Funktionsklasse II
- Vorteil von Sotatercept gegenüber Placebo für den Endpunkt kardiovaskuläre Symptome (PAH-SYMPACT) zu Woche 24
- Nachteile von Sotatercept gegenüber Placebo für die Endpunkte Augenerkrankungen (UEs) und Nasenbluten (UEs)

Über den Zusatznutzen beschließt der G-BA.

3 Literatur

Das Literaturverzeichnis enthält Zitate des pU, in denen gegebenenfalls bibliografische Angaben fehlen.

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Sotatercept (pulmonale arterielle Hypertonie); Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung [online]. 2024 [Zugriff: 16.12.2024]. URL: <https://doi.org/10.60584/A24-96>.
2. MSD Sharp & Dohme. Sotatercept (WINREVAIR); Dossier zur Nutzenbewertung gemäß § 35a SGB V [online]. 2024 [Zugriff: 20.01.2025]. URL: <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/1125/#dossier>.
3. MSD Sharp & Dohme. Stellungnahme zum IQWiG-Bericht Nr. 1898: Sotatercept (pulmonale arterielle Hypertonie); Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung. [Demnächst verfügbar unter: <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/1125/#beschluesse> im Dokument "Zusammenfassende Dokumentation"].
4. MSD Sharp & Dohme. Stellungnahme zum IQWiG-Bericht Nr. 1898: Sotatercept (pulmonale arterielle Hypertonie); Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung; Zusatzanalyse im Nachgang zur mündlichen Anhörung [unveröffentlicht]. 2025.
5. Deutsche Gesellschaft für Kardiologie, Deutschen Gesellschaft für Pneumologie. Stellungnahme zum IQWiG-Bericht Nr. 1898: Sotatercept (pulmonale arterielle Hypertonie); Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung. [Demnächst verfügbar unter: <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/1125/#beschluesse> im Dokument "Zusammenfassende Dokumentation"].
6. Gemeinsamer Bundesausschuss. Sotatercept (D-1104); mündliche Anhörung gemäß 5. Kapitel § 19 Abs. 2 Verfahrensordnung des Gemeinsamen Bundesausschusses - stenografisches Wortprotokoll [online]. 2025 [Zugriff: 07.02.2025]. URL: https://www.g-ba.de/downloads/91-1031-1125/2025-01-27_Wortprotokoll_Sotatercept_D-1104.pdf.
7. Johnson&Johnson. Uptravi Filmtabletten [online]. 12.2024 [Zugriff: 10.02.2025]. URL: <https://www.fachinfo.de>.
8. Janssen-Cilag. VELETRI 0,5 mg/-1,5mg Pulver zur Herstellung einer Infusionslösung [online]. 11.2022 [Zugriff: 07.02.2025]. URL: <https://www.fachinfo.de>.
9. OMT. REMODULIN 5 mg/ml Infusionslösung [online]. 07.2021 [Zugriff: 07.02.2025]. URL: <https://www.fachinfo.de>.

10. MSD. Adempas 0,5 mg/–1 mg/–1,5 mg/–2 mg/–2,5 mg Filmtabletten [online]. 05.2023 [Zugriff: 07.02.2025]. URL: <https://www.fachinfo.de>.
11. Moutchia J, McClelland RL, Al-Naamani N et al. Minimal Clinically Important Difference in the 6-minute-walk Distance for Patients with Pulmonary Arterial Hypertension. *Am J Respir Crit Care Med* 2023; 207(8): 1070-1079. <https://doi.org/10.1164/rccm.202208-1547OC>.
12. Mathai SC, Puhan MA, Lam D, Wise RA. The minimal important difference in the 6-minute walk test for patients with pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2012; 186(5): 428-433. <https://doi.org/10.1164/rccm.201203-0480OC>.
13. Gabler NB, French B, Strom BL et al. Validation of 6-minute walk distance as a surrogate end point in pulmonary arterial hypertension trials. *Circulation* 2012; 126(3): 349-356. <https://doi.org/10.1161/CIRCULATIONAHA.112.105890>.
14. Gilbert C, Brown MCJ, Cappelleri JC et al. Estimating a minimally important difference in pulmonary arterial hypertension following treatment with sildenafil. *Chest* 2009; 135(1): 137-142. <https://doi.org/10.1378/chest.07-0275>.
15. McCollister D, Shaffer S, Badesch DB et al. Development of the Pulmonary Arterial Hypertension-Symptoms and Impact (PAH-SYMPACT(R)) questionnaire: a new patient-reported outcome instrument for PAH. *Respir Res* 2016; 17(1): 72. <https://doi.org/10.1186/s12931-016-0388-6>.
16. Chin KM, Gomberg-Maitland M, Channick RN et al. Psychometric Validation of the Pulmonary Arterial Hypertension-Symptoms and Impact (PAH-SYMPACT) Questionnaire: Results of the SYMPHONY Trial. *Chest* 2018; 154(4): 848-861. <https://doi.org/10.1016/j.chest.2018.04.027>.
17. Martín Andrés A, Silva Mato A. Choosing the optimal unconditioned test for comparing two independent proportions. *Computat Stat Data Anal* 1994; 17(5): 555-574. [https://doi.org/10.1016/0167-9473\(94\)90148-1](https://doi.org/10.1016/0167-9473(94)90148-1).

Anhang A Gründe, warum bei Studienteilnehmenden mit Mono- oder Zweifachtherapie keine zusätzliche medikamentöse Therapie zu Studienbeginn eingesetzt wurde

Tabelle 8: Gründe, warum bei Studienteilnehmenden mit Mono- oder Zweifachtherapie keine zusätzliche Arzneimitteltherapie eingesetzt wurde – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie Art der Hintergrundtherapie Grund	n (%)	
	Sotatercept N = 163	Placebo N = 160
STELLAR		
Monotherapie	9 (5,5)	4 (2,5)
zusätzliche medikamentöse Therapie nicht empfohlen	2 (1,2)	2 (1,3)
zusätzliche medikamentöse Therapie nicht erstattungsfähig / nicht verfügbar	3 (1,8)	1 (0,6)
Sicherheits-/ Verträglichkeitsbedenken	4 (2,5)	1 (0,6)
Zweifachtherapie	56 (34,4)	56 (35,0)
zusätzliche medikamentöse Therapie von der Patientin / dem Patienten verweigert	4 (2,5)	6 (3,8)
zusätzliche medikamentöse Therapie nicht empfohlen	31 (19,0)	16 (10,0)
zusätzliche medikamentöse Therapie nicht erstattungsfähig / nicht verfügbar	5 (3,1)	14 (8,8)
Sicherheits-/ Verträglichkeitsbedenken	16 (9,8)	20 (12,5)
n: Anzahl Patientinnen und Patienten in der Kategorie; N: Anzahl randomisierter Patientinnen und Patienten; RCT: randomisierte kontrollierte Studie		

Anhang B Ergebnisse zu Nebenwirkungen

In den nachfolgenden Tabellen werden für die Gesamtraten UEs und SUEs Ereignisse für SOCs und PTs gemäß Medizinischem Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung (MedDRA) jeweils auf Basis folgender Kriterien dargestellt:

- Gesamtrate UEs (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10 % der Patientinnen und Patienten in 1 Studienarm aufgetreten sind
- Gesamtrate SUEs: Ereignisse, die bei mindestens 5 % der Patientinnen und Patienten in 1 Studienarm aufgetreten sind
- zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patientinnen und Patienten und bei mindestens 1 % der Patientinnen und Patienten in 1 Studienarm aufgetreten sind

Für den Endpunkt Abbruch wegen UEs erfolgt eine vollständige Darstellung aller Ereignisse (SOCs / PTs), die zum Abbruch geführt haben.

Tabelle 9: Häufige UEs^a – RCT, direkter Vergleich: Sotatercept vs. Placebo (mehrsseitige Tabelle)

Studie SOC ^b PT ^b	Patientinnen und Patienten mit Ereignis n (%)	
	Sotatercept N = 163	Placebo N = 160
STELLAR		
Gesamtrate UEs	151 (92,6)	149 (93,1)
Erkrankungen des Blutes und des Lymphsystems	29 (17,8)	25 (15,6)
Thrombozytopenie	16 (9,8)	3 (1,9)
Herzerkrankungen	30 (18,4)	30 (18,8)
Erkrankungen des Ohrs und des Labyrinths	13 (8,0)	14 (8,8)
Augenerkrankungen	21 (12,9)	7 (4,4)
Erkrankungen des Gastrointestinaltrakts	68 (41,7)	51 (31,9)
Diarrhoe	25 (15,3)	16 (10,0)
Übelkeit	23 (14,1)	19 (11,9)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	61 (37,4)	50 (31,3)
Erschöpfung	23 (14,1)	16 (10,0)
Schmerz an der Injektionsstelle	11 (6,8)	11 (6,9)
Peripheres Ödem	14 (8,6)	12 (7,5)
Erkrankungen des Immunsystems	11 (6,8)	7 (4,4)
Infektionen und parasitäre Erkrankungen	105 (64,4)	89 (55,6)
COVID-19	48 (29,5)	42 (26,3)
Nasopharyngitis	11 (6,8)	13 (8,1)
Infektion der oberen Atemwege	9 (5,5)	11 (6,9)
Harnwegsinfekt	11 (6,8)	6 (3,8)
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	27 (16,6)	19 (11,9)
Untersuchungen	33 (20,3)	21 (13,1)
Erhöhter Hämoglobin-Wert	10 (6,1)	0 (0)
Stoffwechsel- und Ernährungsstörungen	37 (22,7)	33 (20,6)
Hypokaliämie	14 (8,6)	6 (3,8)
Eisenmangel	11 (6,8)	9 (5,6)
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	50 (30,7)	30 (18,8)
Erkrankungen des Nervensystems	62 (38)	40 (25,0)
Schwindel	24 (14,7)	10 (6,3)
Kopfschmerzen	40 (24,5)	28 (17,5)
Psychiatrische Erkrankungen	13 (8,0)	8 (5,0)
Erkrankungen der Nieren und Harnwege	7 (4,3)	12 (7,5)

Tabelle 9: Häufige UEs^a – RCT, direkter Vergleich: Sotatercept vs. Placebo (mehrsseitige Tabelle)

Studie SOC ^b PT ^b	Patientinnen und Patienten mit Ereignis n (%)	
	Sotatercept N = 163	Placebo N = 160
Erkrankungen der Geschlechtsorgane und der Brustdrüse	10 (6,1)	10 (6,3)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	59 (36,2)	48 (30,0)
Dyspnoe	5 (3,1)	17 (10,6)
Nasenbluten	36 (22,1)	3 (1,9)
Nasale Kongestion	10 (6,1)	0 (0)
Erkrankungen der Haut und des Unterhautzellgewebes	65 (39,9)	31 (19,4)
Ausschlag	13 (8,0)	6 (3,8)
Teleangiektasie	27 (16,6)	7 (4,4)
Gefäßerkrankungen	27 (16,6)	16 (10,0)
Spülung	10 (6,1)	4 (2,5)
a. Ereignisse, die bei ≥ 10 Patientinnen und Patienten in mindestens 1 Studienarm aufgetreten sind b. MedDRA-Version 25.0; SOC- und PT-Schreibweise ohne Anpassung aus Modul 4 übernommen MedDRA: Medizinisches Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung; n: Anzahl Patientinnen und Patienten mit mindestens 1 Ereignis; N: Anzahl ausgewerteter Patientinnen und Patienten; PT: bevorzugter Begriff; RCT: randomisierte kontrollierte Studie; SOC: Systemorganklasse; UE: unerwünschtes Ereignis		

Tabelle 10: Häufige SUEs^a – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie SOC ^b PT ^b	Patientinnen und Patienten mit Ereignis n (%)	
	Sotatercept N = 163	Placebo N = 160
STELLAR		
Gesamtrate SUEs	40 (24,5)	47 (29,4)
Herzerkrankungen	6 (3,7)	11 (6,9)
Infektionen und parasitäre Erkrankungen	14 (8,6)	8 (5)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	7 (4,3)	11 (6,9)
a. Ereignisse, die in mindestens 1 Studienarm bei ≥ 5 % der Patientinnen und Patienten aufgetreten sind b. MedDRA-Version 25.0; SOC- und PT-Schreibweise ohne Anpassung aus Modul 4 übernommen MedDRA: Medizinisches Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung; n: Anzahl Patientinnen und Patienten mit mindestens 1 Ereignis; N: Anzahl ausgewerteter Patientinnen und Patienten; PT: bevorzugter Begriff; RCT: randomisierte kontrollierte Studie; SOC: Systemorganklasse; SUE: schwerwiegendes unerwünschtes Ereignis		

Tabelle 11: Abbruch wegen UEs – RCT, direkter Vergleich: Sotatercept vs. Placebo

Studie SOC ^a PT ^a	Patientinnen und Patienten mit Ereignis n (%)	
	Sotatercept N = 163	Placebo N = 160
STELLAR		
Gesamtrate Abbruch wegen UEs	6 (3,7)	11 (6,9)
Herzerkrankungen	1 (0,6)	3 (1,9)
Herzstillstand	0 (0)	2 (1,3)
Rechtsherzversagen	0 (0)	1 (0,6)
Akuter Myokardinfarkt	1 (0,6)	0 (0)
Erkrankungen des Immunsystems	1 (0,6)	0 (0)
Sarkoidose	1 (0,6)	0 (0)
Infektionen und parasitäre Erkrankungen	0 (0)	2 (1,3)
COVID-19-Lungenentzündung	0 (0)	1 (0,6)
Sepsis	0 (0)	1 (0,6)
Stoffwechsel- und Ernährungsstörungen	0 (0)	1 (0,6)
Mangelernährung	0 (0)	1 (0,6)
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	1 (0,6)	0 (0)
Arthralgie	1 (0,6)	0 (0)
Erkrankungen des Nervensystems	1 (0,6)	0 (0)
Intrakranielle Blutung	1 (0,6)	0 (0)
Schwangerschaft, Wochenbett und perinatale Bedingungen	0 (0)	1 (0,6)
Schwangerschaftsabbruch	0 (0)	1 (0,6)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	2 (1,2)	4 (2,5)
Pulmonale arterielle Hypertonie	0 (0)	2 (1,3)
Pulmonale Hypertonie	0 (0)	1 (0,6)
Respiratorische Insuffizienz	0 (0)	1 (0,6)
Nasenbluten	1 (0,6)	0 (0)
Hämoptoe	1 (0,6)	0 (0)
Erkrankungen der Haut und des Unterhautzellgewebes	1 (0,6)	0 (0)
Teleangiektasie	1 (0,6)	0 (0)
a. MedDRA-Version 25.0; SOC- und PT-Schreibweise ohne Anpassung aus Modul 4 übernommen MedDRA: Medizinisches Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung; n: Anzahl Patientinnen und Patienten mit mindestens 1 Ereignis; N: Anzahl ausgewerteter Patientinnen und Patienten; PT: bevorzugter Begriff; RCT: randomisierte kontrollierte Studie; SOC: Systemorganklasse; UE: unerwünschtes Ereignis		