

# Nutzenbewertung



Gemeinsamer  
Bundesausschuss

**von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V**

**Bewertung von Arzneimitteln für seltene Leiden nach § 35a  
Absatz 1 Satz 11 i. V. m. Kapitel 5 § 12 Nr. 1 Satz 2 Verfo**

**Wirkstoff: Tofersen**

**Neues Anwendungsgebiet: Behandlung von Erwachsenen  
mit amyotropher Lateralsklerose, die mit einer Mutation im  
Superoxid-Dismutase-1-Gen assoziiert ist**

Datum der Veröffentlichung: 1. Oktober 2024

Version 1.0

## Inhaltsverzeichnis

Tabellenverzeichnis .....	3
Abkürzungsverzeichnis .....	5
Hintergrund .....	7
1 Fragestellung .....	8
2 Auswahl und Methodik der relevanten Studien .....	9
2.1 Studienbasis für die Nutzenbewertung .....	9
2.2 Beschreibung der eingeschlossenen Studie .....	10
2.3 Endpunkte .....	14
2.3.1 Mortalität .....	15
2.3.2 Morbidität .....	16
2.3.3 Lebensqualität .....	24
2.3.4 Sicherheit .....	24
2.3.5 Erhebungszeitpunkte .....	28
2.4 Statistische Methoden .....	29
2.5 Verzerrungspotential auf Studien- und Endpunktebene .....	33
3 Ergebnisse der eingeschlossenen Studie .....	35
3.1 Studiencharakteristika und Studienmedikation .....	35
3.2 Mortalität .....	38
3.3 Morbidität .....	38
3.4 Lebensqualität .....	42
3.5 Sicherheit .....	42
4 Diskussion der Methodik und Ergebnisse .....	47
4.1 Zulassungsstatus und Zulassungspopulation von Tofersen .....	47
4.2 Design und Methodik der Studie .....	47
4.3 Mortalität .....	49
4.4 Morbidität .....	49
4.5 Lebensqualität .....	50
4.6 Sicherheit .....	51
5 Anforderung an eine qualitätsgesicherte Anwendung .....	52
6 Zusammenfassung der Nutzenbewertung .....	53
Referenzen .....	55

## Tabellenverzeichnis

Tabelle 1: Übersicht über die Studienbasis .....	9
Tabelle 2: Charakterisierung der Studie VALOR (Studie 233AS101 Teil C) .....	10
Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Studie 233AS101 Teil C (VALOR) .....	13
Tabelle 4: Charakterisierung der Intervention Tofersen und Placebo.....	14
Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie VALOR.....	15
Tabelle 6: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie VALOR .....	28
Tabelle 7: Verzerrungspotential der Studie VALOR .....	33
Tabelle 8: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie VALOR.....	33
Tabelle 9: Allgemeine Angaben; Studie VALOR (ITT-Population) .....	35
Tabelle 10: Charakterisierung der Studienpopulation; Studie VALOR (ITT-Population).....	35
Tabelle 11: Ergebnisse für Zeit bis zum Tod oder zur dauerhaften Beatmung zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR, ITT-Population .....	38
Tabelle 12: Ergebnisse für Zeit bis zur dauerhaften Beatmung zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; ITT-Population .....	39
Tabelle 13: Ergebnisse für Veränderung im ALSFRS-R-Gesamtwert um $\geq 15$ % bis zum Datenschnitt 16. Juli 2021 in der Studie VALOR, ITT-Population (ergänzend dargestellt) .....	40
Tabelle 14: Ergebnisse für die Veränderung im FSS-Gesamtwert um $\geq 15$ % zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; ITT-Population (ergänzend dargestellt) .....	41
Tabelle 15: Ergebnisse für Veränderung im VAS-Gesamtwert um $\geq 15$ % zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR, ITT-Population (ergänzend dargestellt) .....	41
Tabelle 16: Ergebnisse für Veränderung im SF-36 um $\geq 15$ % zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR, ITT-Population (ergänzend dargestellt) .....	42
Tabelle 17: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE, die aufgetreten sind, zum Studienende (Datenschnitt: 16. Juli 2021) der Studie Valor; Sicherheitspopulation.....	43
Tabelle 18: UE mit Inzidenz $\geq 10$ % in mindestens einem der Behandlungsarme zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; Sicherheitspopulation.....	43
Tabelle 19: Schwere UE (Grad 3 oder höher), die bei $\geq 5$ % der Teilnehmenden in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; Sicherheitspopulation .....	45

Tabelle 20: Schwerwiegende UE, die bei $\geq 5$ % der Teilnehmenden in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; Sicherheitspopulation.....	46
Tabelle 21: Ergebnisse zum C-SSRS zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; Sicherheitspopulation .....	46
Tabelle 22: Zusammenfassende Darstellung der Ergebnisse zu Woche 28 (Tag 197) der Studie VALOR; ITT-Population (Datenschnitt: 16. Juli 2021) .....	53

## Abkürzungsverzeichnis

ALS	Amyotrophe Lateralsklerose
ALSAQ-5	Amyotrophic Lateral Sclerosis Assessment Questionnaire
ALSFRS-R	Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
C-SSRS	Columbia Suicide Severity Rating Scale
CRF	Case report form
CTCAE	Common Terminology Criteria for Adverse Events
EAC	Event Adjudication Committee
EMA	European Medicines Agency
EOS	End of study
EPAR	European Public Assessment Report
EQ-5D	EuroQoL Five Dimension Questionnaire
FDA	U. S. Food and Drug Administration
FSS	Fatigue Severity Scale
FU	Follow-up
FVC	Forced vital capacity (forcierte Vitalkapazität)
HHD	Handheld dynamometer
HR	Hazard Ratio
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Interactive Response Technology
ITT	Intention to Treat
k. A.	Keine Angabe
KI	Konfidenzintervall
MCS	Psychischer Summenscore
MedDRA	Medical Dictionary for Regulatory Activities
MI	Multiple Imputation
mITT	Modified Intention to Treat
MMSE	Mini Mental State Examination
MW	Mittelwert
N	Anzahl
NfL	Neurofilament light chain (Neurofilament Leichtketten)
OLE	Open-Label-Extension (Langzeit-Extensionsstudie)
PCA	Principal Component Analysis
PCS	Körperlicher Summenscore
PGI-C	Patient Global Impression of Change
PGI-S	Patient Global Impression of Status
PRO	Patient reported outcome
PT	Preferred Term/s

pU	Pharmazeutischer Unternehmer
RR	Relatives Risiko
SAP	Statistischer Analyseplan
SF-36	36-Item Short Form Health Survey
SGB	Sozialgesetzbuch
SOC	Systemorganklasse
SOD1	Superoxid-Dismutase Typ 1
SOP	Standard Operating Procedure
SUE	Schwerwiegendes unerwünschtes Ereignis
SVC	Slow vital capacity (langsame Vitalkapazität)
UE	Unerwünschtes Ereignis
VAS	Visuelle Analogskala
VerfO	Verfahrensordnung des G-BA
WPAI	Work Productivity and Activity Impairment

## Hintergrund

Tofersen ist zugelassen als Arzneimittel zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt der medizinische Zusatznutzen durch die Zulassung als belegt.

Gemäß § 35a Absatz 2 SGB V entscheidet der Gemeinsame Bundesausschuss (G-BA), ob er die Nutzenbewertung selbst durchführt oder das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) beauftragt. Ausgehend von der gesetzlichen Vorgabe in § 35a Absatz 1 Satz 11 1. Halbs. SGB V, dass der Zusatznutzen eines Orphan Drug durch die Zulassung als belegt gilt, hat der G-BA in seiner Sitzung vom 15. März 2012 das Verfahren der Nutzenbewertung von Orphan Drugs dahingehend modifiziert, dass bei Orphan Drugs zunächst keine eigenständige Festlegung einer zweckmäßigen Vergleichstherapie mehr durch den G-BA als Grundlage der insoweit allein rechtlich zulässigen Bewertung des Ausmaßes eines gesetzlich zu unterstellenden Zusatznutzens erfolgt. Vielmehr wird ausschließlich auf der Grundlage der Zulassungsstudien das Ausmaß des Zusatznutzens durch den G-BA bewertet.

Bei Orphan Drugs erfolgt eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann, wenn der Umsatz des betreffenden Arzneimittels die Umsatzschwelle gemäß § 35a Absatz 1 Satz 12 SGB V überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt.

Der G-BA bestimmt gemäß Kapitel 5 § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des G-BA (VerfO) das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Tofersen zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers (pU) heran und bewertet die darin enthaltenen Daten nach Maßgabe der in Kapitel 5 § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Tofersen in seiner Sitzung am 24. September 2024 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 01. Juli 2024 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

- Kosten der Therapie für die gesetzliche Krankenversicherung (GKV)
- Anzahl der GKV-Patientinnen und -Patienten in der Zielpopulation

Die Nutzenbewertung wird am 01. Oktober 2024 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der G-BA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

## 1 Fragestellung

Tofersen (Qalsody) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der G-BA bestimmt bei Orphan Drugs, die einen Umsatz von 30 Millionen Euro in den letzten 12 Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. Die Bewertung bezieht sich auf folgende Indikation [8]:

Behandlung von Erwachsenen mit amyotropher Lateralsklerose (ALS), die mit einer Mutation im Superoxid-Dismutase-1-Gen (SOD1) assoziiert ist.

Gemäß Fachinformation beträgt die zugelassene Dosierung 100 mg Tofersen pro Behandlung. Die Behandlung mit Tofersen sollte mit 3 Aufsättigungsdosen begonnen werden, die im Abstand von jeweils 14 Tagen verabreicht werden. Danach sollte alle 28 Tage eine Erhaltungsdosis verabreicht werden. Tofersen ist zur intrathekalen Anwendung mittels Lumbalpunktion bestimmt.



## 2 Auswahl und Methodik der relevanten Studien

### 2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studienname (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
<b>Studien zum Wirkstoff</b>				
Studie VALOR 233AS101 (NCT02623699) <sup>1)</sup>	Ja	Ja	Ja	<ul style="list-style-type: none"> <li>Teil A und B werden nicht dargestellt.</li> <li>In Teil A wurde eine Einzeldosis Tofersen (N = 15) in einer nicht-zulassungskonformen Dosierung appliziert.</li> <li>In Teil B wurden &gt; 50 % der Patienten und Patientinnen (N = 50) mit einer nicht fachinformationskonformen Dosierung von Tofersen behandelt und gegenüber Placebo verglichen. Die Anwendung von Tofersen war auf maximal 5 Applikationen beschränkt.</li> </ul>
Studie 233AS102 (NCT03070119) OLE <sup>2)</sup>	Ja	Ja	Nein	<ul style="list-style-type: none"> <li>Einarmige Verlängerungsstudie, bei der alle Teilnehmenden der Studie 233AS101 Tofersen erhalten.</li> <li>Keine vergleichenden Langzeitdaten vorhanden.</li> </ul>
Studie 233AS303 <sup>3)</sup>	Ja	Nein	Nein	<ul style="list-style-type: none"> <li>Studie noch nicht beendet. Keine Interimsdaten vorhanden.</li> </ul>

<sup>1)</sup> Zulassungsrelevante Studie gemäß EPAR.

<sup>2)</sup> Im Rahmen des Zulassungsverfahrens bei der EMA ergänzend zu den Zulassungsstudien eingereicht.

<sup>3)</sup> Randomisierte, placebo-kontrollierte Studie mit einer Run-In-Längsschnittstudie zum natürlichen Verlauf und offener Verlängerungsstudie zur Untersuchung von Tofersen bei klinisch präsymptomatischen Erwachsenen mit einer bestätigten SOD1-Mutation.

Abkürzungen: EMA: European Medicines Agency; EPAR: European Public Assessment Report; OLE: Open-Label-Extension; pU: pharmazeutischer Unternehmer; SOD1: Superoxid-Dismutase Typ 1

#### Zur Nutzenbewertung für Tofersen herangezogene Unterlagen

- Dossier des pU zu Tofersen [6,7]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [16]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie VALOR (233AS101 Teil C) [5,9,10,11,12]
- Fachinformation zu Tofersen [8]

## 2.2 Beschreibung der eingeschlossenen Studie

Die Nutzenbewertung von Tofersen im vorliegenden Anwendungsgebiet basiert auf der Zulassungsstudie VALOR (Studie 233AS101 Teil C). Die Studie und die Intervention werden in den Tabellen 2–4 charakterisiert.

Tabelle 2: Charakterisierung der Studie VALOR (Studie 233AS101 Teil C)

Charakteristikum	Beschreibung
<b>Design und Studienablauf</b>	<p>Phase I (Teil A: SAB), Phase II (Teil B: MAB), Phase III (Teil C: VALOR) Studie zur Beurteilung der Wirksamkeit (nur in Phase III bzw. Teil C-VALOR untersucht), Sicherheit, Verträglichkeit, Pharmakokinetik und -dynamik von Tofersen bei erwachsenen Teilnehmenden mit amyotropher Lateralsklerose und bestätigter SOD1-Mutation</p> <p>Phase III (Teil C: pivotale Studie): randomisierte, doppelblinde, placebo-kontrollierte Studie mit einer 9-maligen Applikation von 100 mg Tofersen über ca. 24 Wochen bei SOD1-assoziierter ALS. Die Randomisierung erfolgte 2:1 in die Behandlungsarme Tofersen oder Placebo.</p> <p>Die Randomisierung erfolgte stratifiziert anhand der 3 folgenden Kriterien:</p> <ul style="list-style-type: none"> <li>• Erfüllung der prognostischen Kriterien für eine schnelle Krankheitsprogression (1 der beiden Kriterien trifft zu):             <ol style="list-style-type: none"> <li>a) Abfall in der ALSFRS-R von <math>\geq 0,2</math> pro Monat (berechnet als [48-Baseline-Score]/Zeit seit Symptombeginn) mit folgenden SOD1-Mutationen: p.Ala5Val, p.Ala5Thr, p.Leu39Val, p.Gly42Ser, p.His44Arg, p.Leu85Val, p.Gly94Ala, p.Leu107Val und p.Val149Gly</li> <li>oder</li> <li>b) keine der zuvor genannten SOD1-Mutationen und ein Abfall der ALSFRS-R vor Randomisierung von <math>\geq 0,9</math> pro Monat (berechnet als [48-Baseline-Score]/Zeit seit Symptombeginn)</li> <li>oder</li> <li>c) Kriterien für alle anderen teilnahmeberechtigten Patientinnen und Patienten: SOD1-Mutation, die nicht unter Punkt „a“ aufgeführt ist (Kriterium für Abfall in der ALSFRS-R besteht nicht).</li> </ol> </li> <li>• Einnahme von Edaravon zu Baseline</li> <li>• Einnahme von Riluzol zu Baseline             <ul style="list-style-type: none"> <li>○ Die gleichzeitige Anwendung von beiden Wirkstoffen, Edaravon und Riluzol, wurde nicht als eigenständiger Stratifizierungsfaktor definiert. Bei dieser Konstellation wurde anhand Edaravon stratifiziert.</li> </ul> </li> </ul> <p><u>Studienphasen</u></p> <ul style="list-style-type: none"> <li>• 4-wöchige Screening-Phase (Woche -4 bis Tag 0)</li> <li>• 24-wöchige Behandlungsphase             <ul style="list-style-type: none"> <li>○ 3 Aufsättigungsdosen im Abstand von 2 Wochen</li> <li>○ 5 Erhaltungsdosen im Abstand von 4 Wochen</li> </ul> </li> <li>• 4–8-wöchige FU-Phase             <ul style="list-style-type: none"> <li>○ Visite Woche 28 (Tag 197): „End of Study“ (EOS)-Visite der Studie VALOR bei Teilnehmenden, die sofort in die Verlängerungsstudie OLE übergangen. EOS-Visite bei vorzeitigem Abbruch der Studie VALOR innerhalb von 4 Wochen nach Erhalt der letzten Dosis der Studienmedikation.</li> <li>○ Visite Woche 32 (Tag 225): Alternative EOS-Visite zur Erhebung der Sicherheit und Begleitmedikation bzw. Prozeduren bei Teilnehmenden, die</li> </ul> </li> </ul>

Charakteristikum	Beschreibung
	<p>verzögert an der Verlängerungsstudie OLE teilnehmen. Visite kann zu Woche 32 (Tag 225) oder 2 Tage vor Einschluss in die Verlängerungsstudie OLE erfolgen. Die Visite erfolgte zudem für Teilnehmende, die nicht an der Verlängerungsstudie OLE teilnehmen. Die Visite erfolgte 8 Wochen nach Erhalt der letzten Dosis der Studienmedikation.</p> <p>Die Studie ist abgeschlossen.</p>
<b>Population</b>	<p><b>Wesentliche Einschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Alter <math>\geq 18</math> Jahre zum Zeitpunkt der Einverständniserklärung</li> <li>• Schwäche, die auf ALS und eine bestätigte SOD1-Mutation zurückzuführen ist.             <ol style="list-style-type: none"> <li>a) Bestätigung einer SOD1-Mutation während der Screening-Visite durch ein Zentrallabor</li> <li>b) Zusätzlich:                 <ul style="list-style-type: none"> <li>○ Prognostische Kriterien für eine schnelle Krankheitsprogression (1 der folgenden 2 Kriterien trifft zu):                     <ol style="list-style-type: none"> <li>i. Eine der folgenden SOD1-Mutationen: p.Ala5Val, p.Ala5Thr, p.Leu39Val, p.Gly42Ser, p.His44Arg, p.Leu85Val, p.Gly94Ala, p.Leu107Val und p.Val149Gly und ein Abfall in der ALSFRS-R von <math>\geq 0,2</math> pro Monat (berechnet als <math>[48\text{-Baseline-Score}]/\text{Zeit seit Symptombeginn}</math>) oder</li> <li>ii. SOD1-Mutation, die nicht unter Punkt „i“ aufgeführt ist, mit einem Abfall in der ALSFRS-R vor Randomisierung von <math>\geq 0,9</math> pro Monat (berechnet als <math>[48\text{-Baseline-Score}]/\text{Zeit seit Symptombeginn}</math>)</li> </ol> </li> <li>○ Kriterien für alle anderen teilnahmeberechtigten Personen: SOD1-Mutation, die nicht unter Punkt „i“ aufgeführt ist (Kriterium für Abfall in der ALSFRS-R besteht nicht).</li> </ul> </li> </ol> </li> <li>• Für Patienten und Patientinnen, die die prognostischen Kriterien für eine schnelle Krankheitsprogression erfüllen: langsame Vitalkapazität (slow vital capacity, SVC) <math>\geq 65</math> % des Vorhersagewerts, adjustiert für Geschlecht, Alter und Körpergröße (aus der sitzenden Position). Für alle anderen teilnahmeberechtigten Patienten und Patientinnen SVC <math>\geq 50</math> % des Vorhersagewerts, adjustiert für Geschlecht, Alter und Körpergröße (aus der sitzenden Position). Es sollten mindestens 3 Tests akzeptable Ergebnisse aufweisen, von denen mindestens 2 Tests mit der bestmöglichen und zweitbestmöglichen Anstrengung eine Vitalkapazität innerhalb von 150 ml aufweisen.</li> <li>• Bei Einnahme von Riluzol: <math>\geq 30</math> Tage vor Tag 1 mit stabiler Dosis und voraussichtlich gleichbleibender Dosis bis zur letzten Studienvisite</li> <li>• Bei Einnahme von Edaravon: Dosisstart <math>\geq 60</math> Tage (2 Behandlungszyklen) vor Tag 1 und voraussichtlich gleichbleibende Dosis bis zur letzten Studienvisite. Sofern das Prüfpersonal entscheidet, dass Edaravon aus medizinischen Gründen abgesetzt werden sollte, war es in diesem Fall nicht erlaubt, eine Behandlung während der Studie wieder zu beginnen. Es war ebenfalls nicht erlaubt, Edaravon an den Dosierungstagen dieser Studie zu verabreichen.</li> </ul> <p><b>Wesentliche Ausschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Anamnese oder positives Testergebnis für Humanes Immundefizienz-Virus (HIV)</li> <li>• Aktuelle Hepatitis-C-Infektion (definiert als positive Hepatitis-C-Virus [HCV] positive Antikörper (Ak) und nachweisbare HCV-Ribonukleinsäure (RNA)). Patienten und Patientinnen mit positivem HCV-Ak und nicht nachweisbarer HCV-RNA waren berechtigt, an der Studie teilzunehmen.</li> </ul>

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> <li>• Aktuelle Hepatitis-B (HB)-Infektion (definiert als positiv für HB-Ak und/oder Anti-Hepatitis-B-Core-Antigen [HBc]). Personen mit Immunität gegen Hepatitis B durch vorherige natürliche Infektion (definiert als negative HB-Ak, positive Anti-HBc und positive Anti-HBs) oder Impfung (definiert als negative HB-Ak, negative Anti-HBc und positive Anti-HBs) waren berechtigt, an der Studie teilzunehmen.</li> <li>• Behandlung mit einem anderen Prüfpräparat (einschließlich Prüfpräparaten für ALS durch Compassionate-Use-Programme), biologischen Wirkstoff oder Prozedur innerhalb von 1 Monat oder 5 Halbwertszeiten des Studienpräparats, je nachdem, welcher Zeitraum länger war. Insbesondere war keine vorherige Behandlung mit small interfering RNA, Stammzelltherapie oder Gentherapie erlaubt.</li> <li>• Nach Ansicht des Prüfpersonals aktueller oder erwarteter Bedarf von einem Zwerchfelschrittmacher während des Studienzeitraums.</li> <li>• Nach Ansicht des Prüfpersonals aktuelle oder kürzliche (innerhalb von 1 Monat) Verwendung oder erwarteter Bedarf von Kupfer (II) (Diacetyl-Bis(N4-methylthiosemicarbazon)) oder Pyrimethamin.</li> <li>• Vorliegen einer unbehandelten oder unzureichend behandelten aktiven Infektion, die eine systemische antivirale oder antimikrobielle Therapie erfordert, zu irgendeinem Zeitpunkt während der Screening-Phase.</li> <li>• Anhaltender medizinischer Zustand (z. B. Auszehrung oder Kachexie, schwere Anämie), der nach Ansicht des Prüfpersonals die Durchführung oder Bewertung der Studie beeinträchtigen würde.</li> <li>• Vorgeschichte von Allergien gegen eine Vielzahl von Anästhetika.</li> <li>• Das Vorhandensein eines Risikos für erhöhte oder unkontrollierte Blutungen und/oder das Risiko einer Blutung, die nicht optimal behandelt werden kann. Dazu gehören unter anderem anatomische Faktoren an oder in der Nähe der Lumbalpunktion-Stelle (z. B. vaskuläre Anomalien, Neoplasmen oder andere Anomalien) und zugrunde liegende Störungen der Gerinnungskaskade, der Thrombozytenfunktion oder der Thrombozytenzahl (z. B. Hämophilie, Von-Willebrand-Krankheit, Lebererkrankung).</li> <li>• Nach Ansicht des Prüfpersonals erwarteter Bedarf für die Verabreichung von Thrombozytenaggregationshemmern oder gerinnungshemmenden Medikamenten vor und/oder nach der Lumbalpunktion.</li> <li>• Vorhandensein eines implantierten Shunts zur Drainage von Liquor oder eines implantierten Zentrales Nervensystem (ZNS)-Katheters.</li> <li>• Nach Ansicht des Prüfpersonals klinisch signifikante Anomalien in hämatologischen oder klinisch-chemischen Parametern.</li> <li>• Nach Ansicht des Prüfpersonals klinisch signifikante 12-Kanal-Elektrokardiogramm-Anomalien, einschließlich korrigiertem QT-Intervall unter Verwendung der Fridericia-Korrekturmethode von &gt; 450 ms für Männer und &gt; 470 ms für Frauen.</li> </ul>
<b>Intervention und Zahl der Patientinnen und Patienten</b>	<p>Gescreent gesamt: N = 162  Randomisiert/Eingeschlossen: N = 108  Intervention: N = 72  Kontrolle: N = 36</p>
<b>Ort und Zeitraum der Durchführung; Datenschnitte</b>	<p><b>Studienzentren</b>  32 Studienzentren in 9 Ländern (Belgien, Kanada, Dänemark, Deutschland, Italien, Japan, Vereinigtes Königreich, Vereinigte Staaten von Amerika)</p>

Charakteristikum	Beschreibung
	<b>Studienzeitraum</b> <ul style="list-style-type: none"> <li>• Erste/r Patient/in, erste Visite: 27. März 2019</li> <li>• Letzte/r Patient/in, letzte Visite: 16. Juli 2021</li> </ul>
<b>Endpunkte gemäß Studienprotokoll</b>	<b>Primärer Endpunkt</b> Veränderung des ALSFRS-R-Gesamtscore von Baseline zu Woche 28  <b>Sekundäre Endpunkte</b> <ul style="list-style-type: none"> <li>• Veränderung der SVC zu Woche 28 gegenüber Baseline</li> <li>• Veränderung vom Handdynamometer (handheld dynamometer, HHD)-Megascore zur Beurteilung der Muskelkraft zu Woche 28 gegenüber Baseline</li> <li>• Zeit bis zum Tod oder zur dauerhaften Beatmung (<math>\geq 22</math> Stunden mechanische Beatmung [invasiv oder nichtinvasiv] pro Tag an <math>\geq 21</math> aufeinanderfolgenden Tagen)</li> <li>• Zeit bis zum Tod</li> <li>• Sicherheit (UE, MMSE, C-SSRS)</li> </ul> <b>Explorative Endpunkte zur Lebensqualität</b> <ul style="list-style-type: none"> <li>• Amyotrophic Lateral Sclerosis Assessment Questionnaire (ALSAQ-5)</li> <li>• Fatigue Severity Scale (FSS)</li> <li>• EuroQol Five Dimension Questionnaire (EQ-5D-5L)</li> <li>• Work Productivity and Activity Impairment (WPAI)</li> <li>• Zarit Burden Interview (ZBI)</li> <li>• 36-Item Short Form Health Survey (SF-36)</li> <li>• Patient Global Impression of Change (PGI-C)</li> <li>• Clinician Global Impression of Change (CGI-C)</li> <li>• Patient Global Impression of Status (PGI-S)</li> <li>• Clinician Global Impression of Status (CGI-S)</li> </ul>

Abkürzungen: ALS: Amyotrophe Lateralsklerose; ALSAQ-5: Amyotrophic Lateral Sclerosis Assessment Questionnaire; ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; C-SSRS: Columbia Suicide Severity Rating Scale; EOS: End of study; EQ-5D: EuroQol Five Dimension Questionnaire; FSS: Fatigue Severity Scale; FU: Follow-up; HHD: handheld dynamometer; MMSE: Mini Mental State Examination; MUNIX: motor unit number index; OLE: Open-Label-Extension; PGI-C: Patient Global Impression of Change; PGI-S: Patient Global Impression of Status; SF-36: 36-Item Short Form Health Survey; SOD1: Superoxid-Dismutase Typ 1; SVC: Slow vital capacity; UE: Unerwünschtes Ereignis; WPAI: Work Productivity and Activity Impairment; ZBI: Zarit Burden Interview

### Protokolländerungen

Es wurden 8 Änderungen des Originalprotokolls vom 24. September 2015 vorgenommen. Es werden die für die Nutzenbewertung relevanten Protokolländerungen seit Einschluss des/der ersten Patienten/Patientin (27. März 2019) dargestellt.

*Tabelle 3: Für die Nutzenbewertung relevante Protokolländerungen der Studie 233AS101 Teil C (VALOR)*

Amendment	Wesentliche Änderungen
Version 6 vom 19.09.2019 (Anzahl der bis dahin eingeschlossenen Personen: k. A.)	<ul style="list-style-type: none"> <li>• Änderung des primären Endpunktes von ALSFRS-R-Steigung zu ALFSR-R-Gesamtscore (Änderung zu Woche 28 gegenüber Baseline) und Auswertung mittels Joint-Rank-Analyse (Kombination des ALSFRS-R-Scores und der Überlebenszeit). Dieser Ansatz minimierte die Abhängigkeit einer Linearitätsannahme (ALSFRS-R-Steigungsabfall).</li> </ul>

Amendment	Wesentliche Änderungen
	<ul style="list-style-type: none"> <li>• Erhöhung der Fallzahl von ursprünglich 60 Personen auf 99 Personen aufgrund des geänderten primären Endpunktes, des zweiseitigen <math>\alpha</math>-Fehlers von 0,05, überarbeiteter Annahmen zum Überleben aufgrund von Daten zum natürlichen Verlauf und den Interimsdaten aus Teil B (82 % vs. 90 % Überleben unter Placebo bzw. Tofersen).</li> <li>• Gemäß SAP V.2.0 konnte eine geschulte Person als Vertretung den Endpunkt ALSFRS-R erheben.</li> </ul>

Abkürzungen: ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; k. A.: keine Angabe; SAP: Statistischer Analyseplan

## Charakterisierung der Intervention

Tabelle 4: Charakterisierung der Intervention Tofersen und Placebo

Intervention	Kontrolle
<b>Tofersen</b> <ul style="list-style-type: none"> <li>• Dosierung: 100 mg einmal täglich intrathekal appliziert als Bolus (Gesamtvolumen 15 ml) über 1–3 Minuten</li> <li>• Applikation von insgesamt 8 Dosen: <ul style="list-style-type: none"> <li>○ 3 Aufsättigungsdosen am Tag 1, 15 und 29</li> <li>○ im Anschluss Erhaltungsdosis alle 4 Wochen</li> </ul> </li> <li>• Dosisanpassungen waren nicht vorgesehen.</li> </ul>	<b>Placebo</b> <ul style="list-style-type: none"> <li>• Placebo (künstlicher Liquor) wurde mit gleichem Dosierungsschema wie Tofersen ebenfalls durch intrathekalen Bolus über 1–3 Minuten verabreicht.</li> </ul>
<b>Nicht erlaubte Begleitmedikation</b> <ul style="list-style-type: none"> <li>• Thrombozytenaggregationshemmer und gerinnungshemmende Medikamente, die vor und/oder nach einem LP-Eingriff gemäß den lokalen oder institutionellen Richtlinien und/oder nach Einschätzung des ärztlichen Prüfpersonals nicht sicher ausgesetzt werden können.</li> <li>• Verwendung eines Zwerchfell-Schrittmachersystems, Kupfer (II) (Diacetyl-bis(N4-methylthiosemicarbazon) oder Pyrimethamin</li> <li>• Off-Label-Anwendung von krankheitsmodifizierenden Behandlungen der ALS</li> </ul> <b>Erlaubte Begleitmedikation</b> <ul style="list-style-type: none"> <li>• Riluzol in einer unveränderten Dosierung mindestens 30 Tage vor Tag 1</li> <li>• Edaravon in einer unveränderten Dosierung mit Behandlungsbeginn mindestens 60 Tage vor Tag 1</li> <li>• Tägliche Einnahme von Vitaminen und Nahrungsergänzungsmitteln in unveränderter Dosierung mindestens 14 Tage vor Tag 1. Tägliche Einnahme von Kreatinin in einer Dosierung von <math>\leq 5</math> g und Vitamin E unter <math>\leq 1.000</math> IU</li> <li>• Medikation zur Behandlung von Symptomen im Ermessen des ärztlichen Prüfpersonals</li> </ul>	

Abkürzungen: ALS: Amyotrophe Lateralsklerose; LP: Lumbalpunktion

## 2.3 Endpunkte

In diesem Kapitel wird die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. Alle Endpunkte der eingeschlossenen Studie (siehe Tabelle 2) wurden einer Prüfung unterzogen. Endpunkte, die in der nachfolgenden Tabelle 5 nicht gelistet sind, wurden weder vom pU noch im Rahmen der Nutzenbewertung als patientenrelevant bzw. bewertungsrelevant eingestuft. Tabelle 5 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie VALOR

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Gesamtüberleben	Mortalität	Ja	Ja
Zeit bis zum Tod oder zur dauerhaften Beatmung	Mortalität/Morbidität	Ja	Ja
ASLFRS-R <sup>1)</sup>	Morbidität	Ja	Ergänzend <sup>2)</sup>
FSS		Ja	Ergänzend <sup>2)</sup>
EQ-5D-5L VAS		Ja	Ergänzend <sup>2)</sup>
ALSAQ-5 <sup>3)</sup>		Ja	Nein
WPAI-Frage 6 <sup>3)</sup>		Ja	Nein
SVC		Ja	Nein
HHD-Megascor		Ja	Nein
PGI-C <sup>4)</sup>		Ja	Nein
PGI-S <sup>4)</sup>		Ja	Nein
SF-36		Lebensqualität	Ja
Unerwünschte Ereignisse <sup>5)</sup>	Sicherheit	Ja	Ja
C-SSRS		Ja	Ja
MMSE		Ja	Nein

<sup>1)</sup> Primärer Endpunkt

<sup>2)</sup> Der Endpunkt wird ergänzend dargestellt aufgrund unklarer Definition und Berechnung der Responderanalysen (siehe Abschnitt 2.4).

<sup>3)</sup> Der Endpunkt wird, anders als im Dossier Modul 4, nicht der Lebensqualität, sondern der Kategorie Morbidität zugeordnet (siehe Ausführungen unter Abschnitt 2.3.2)

<sup>4)</sup> Der Endpunkt wurde nicht bewertet aufgrund zu geringer Rücklaufquoten (< 70 %) zu allen Messzeiten post-Baseline.

<sup>5)</sup> Der pU berichtet Ergebnisse zum Endpunkt „Vermeidung von Stürzen“ unter der Kategorie „Morbidität“. Der Endpunkt wurde im Rahmen der Sicherheit als PT „Sturz“ erhoben und für das Dossier Modul 4 post-hoc als Endpunkt „Vermeidung von Stürzen“ der Kategorie „Morbidität“ zugeordnet.

Abkürzungen: ALSAQ-5: Amyotrophic Lateral Sclerosis Assessment Questionnaire; ASLFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; C-SSRS: Columbia Suicide Severity Rating Scale; EQ-5D: EuroQol Five Dimension Questionnaire; FSS: Fatigue Severity Scale; HHD: handheld dynamometer; MMSE: Mini Mental State Examination; PGI-C: Patient Global Impression of Change; PGI-S: Patient Global Impression of Status; PT: Preferred Term; pU: pharmazeutischer Unternehmer; SF-36: 36-Item Short Form Health Survey; SVC: slow vital capacity; VAS: Visuelle Analogskala; WPAI: Work Productivity and Activity Impairment

### 2.3.1 Mortalität

#### Gesamtüberleben

Der Endpunkt Gesamtüberleben wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Das Gesamtüberleben wird als „Zeit bis zum Tod“ analysiert. Der Endpunkt ist ebenfalls Bestandteil des Endpunktes „Zeit bis zum Tod oder zur dauerhaften Beatmung“.

Die Nachbeobachtung des Endpunktes endet zum letzten Kontakt (persönliches Erscheinen zur Visite, Hausbesuch oder telefonischer Kontakt) mit dem Patienten oder der Patientin in

der Studie, unabhängig davon, ob die vollständige Behandlungsdauer erreicht, die Studie regulär oder frühzeitig beendet wurde. Bei vorzeitigem Abbruch der Studie bzw. am Studienende fand eine Visite 4–8 Wochen später statt.

Personen wurden auf Grundlage des letzten bekannten Lebenszeitpunkts zensiert. Für die Berechnung wurde das Datum des Erhalts der ersten Dosis verwendet.

#### *Bewertung*

Die Operationalisierung ist nachvollziehbar.

#### Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen.

#### Validität

Die Erhebung des Endpunktes wird als valide eingeschätzt.

### **2.3.2 Morbidität**

#### **Zeit bis zum Tod oder zur dauerhaften Beatmung**

Der Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“ wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Die Zeit bis zum Tod oder zur dauerhaften Beatmung ist definiert als die Zeit bis zum Auftreten eines der folgenden Ereignisse:

- Tod
- Dauerhafte Beatmung ( $\geq 22$  Stunden/Tag invasiv oder nicht-invasiv durchgehend für  $\geq 21$  Tage)

Der Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung wurde verblindet durch ein zentrales, unabhängiges Entscheidungskomitee (EAC) ausgewertet. Das Verfahren und die Beurteilung der Ereignisse sind in einer Satzung des EAC beschrieben.

**Zeit bis zur dauerhaften Beatmung:** Gemäß Definition kann die dauerhafte Beatmung eine beliebige Kombination aus invasiver (IV) und nicht-invasiver Beatmung (NIV) sein, solange sie für die angegebene Anzahl von Stunden und Tagen erfolgt. Die Beatmung wurde während der Studie über das tägliche Beatmungstagebuch (Papier oder elektronisch) erfasst. Neben dem Beatmungstagebuch können weitere Quelldokumente verwendet werden. Es liegt im Ermessen des Prüfarztes oder der Prüfarztin, die Anzahl der Beatmungstunden pro Tag und die Art der verwendeten Beatmungsunterstützung zu bestimmen.

Als Beatmungsunterstützung werden folgende Optionen angegeben: invasive Beatmung (über Tracheostomie oder Endotrachealtubus) oder nicht-invasive Beatmung (biphasischer positiver Atemwegsdruck (Biphasic Positive Airway Pressure, BiPAP) mittels Atemmaske, kontinuierlicher Atemwegsüberdruck (Continuous Positive Airway Pressure, CPAP) mittels Atemmaske, Beatmung mit intermittierend positivem Druck (Intermittent Positive Pressure Ventilation, IPPV), druckkontrollierte Beatmung (pressured control, PC), Druckunterstützung mit gesichertem Durchschnittsvolumen (Average Volume Assured Pressure Support, AVAPS), Proportional Assist Ventilation (PAV), andere. Bei einer nicht-invasiven Beatmung war der Einsatz von nasalen und Mundmasken, Mundstücken, Nasenkissen, Ganzgesichtsmasken bzw. Helmgeräten vorgesehen.

Die Daten der verschiedenen Quellen werden, mit Ausnahme der elektronischen Tagebuchdaten, die als eigene Quelle dienen, in das Beatmungsprotokoll des CRF eingetragen.



Gemäß CRF soll das Studienpersonal die Plausibilität und Vollständigkeit des Patiententagebuchs überprüfen und, sofern eine Beatmung notwendig war, angeben, ob diese Information aus den medizinischen Unterlagen bzw. aus dem Patiententagebuch stammt. Die Überprüfung und Beurteilung der Ereignisse, die zu einer dauerhaften Beatmung führten, sollten gemäß einer Satzung des EAC durchgeführt werden.

Eine Auswertung der Zeit bis zum Tod oder zur dauerhaften Beatmung war zum Studienende zu Visite Woche 28 (Tag 197) vorgesehen.

**Zeit bis zum Tod:** Die Komponente „Zeit bis zum Tod“ wird beim Endpunkt „Gesamtüberleben“ in der Nutzenbewertung beschrieben (siehe Abschnitt 2.3.1.).

#### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Es handelt sich um einen kombinierten Endpunkt, der sich aus Teilkomponenten der Kategorie Mortalität und Morbidität zusammensetzt. Die Überprüfung und Beurteilung der Ereignisse, die zu einer dauerhaften Beatmung führten, sollten gemäß einer Satzung des EAC durchgeführt werden. Die Satzung konnte in den eingereichten Unterlagen nicht identifiziert werden.

#### Patientenrelevanz

Der Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“ ist ein kombinierter Endpunkt aus „Zeit bis zum Tod“ (Mortalität) und „zur dauerhaften Beatmung“ (Morbidität). Beide Teilkomponenten werden als patientenrelevant angesehen.

#### Validität

Der Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“ ist ein kombinierter Endpunkt aus „Zeit bis zum Tod“ (Mortalität) und „Zeit bis zur dauerhaften Beatmung“ (Morbidität). Zur Beschreibung des Endpunktes „Gesamtüberleben“ siehe Abschnitt 2.3.1.

Der pU erhebt sowohl die invasive und die nicht-invasive Beatmung und wertet diese gemeinsam aus. Eine invasive Beatmung sollte unter strenger Indikationsstellung und u. a. unter Berücksichtigung des individuellen Krankheitsverlaufs vorgenommen werden [14]. Die invasive und die nicht-invasive Beatmung unterscheiden sich dahingehend, dass für eine nicht-invasive Beatmung keine Intubation oder Tracheotomie erforderlich sind. Die invasive Beatmung kann z. B. mit tubusbedingten Komplikationen (wie z. B. Schleimhautläsionen) oder ventilatorassoziierter Pneumonie (VAP) einhergehen und ist somit schwerwiegender anzusehen als die nicht-invasive Beatmung [26]. Eine differenzierte Auswertung von einer invasiven oder nicht-invasiven Beatmung war nicht vorgesehen.

Gemäß Dossier Modul 4 liegt die Art und Dauer der Beatmung im Ermessen des Prüfarztes bzw. der Prüfarztin. Diese kann jedoch nach Land und Region variieren. Eine standardisierte Erhebung wird zum Teil empfohlen [15]. Inwieweit eine Standardisierung mittels Überprüfung der Ereignisse durch das EAC vorgenommen wurde, ist unklar.

Aufgrund von qualitativer und quantitativer Heterogenität werden die Teilkomponenten des Endpunktes zusätzlich einzeln betrachtet.

#### **Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised (ALSFERS-R)**

Der Endpunkt Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Der „Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised“ (ALSFERS-R) wurde entwickelt, um das Ausmaß der Einschränkungen in der Progression von amyotropher

Lateralsklerose zu messen. Der Fragebogen verfügt über 4 Subskalen mit jeweils 3 Items, die die Aspekte Grobmotorik, Feinmotorik, bulbäre Symptome (Beeinträchtigung von Sprache, Schlucken oder Speichelfluss) und Atemfunktion betreffen [13]. Die Beurteilung findet durch geschultes Klinikpersonal statt. Die insgesamt 12 Items werden mit einer 5-Punkte-Likert-Skala bewertet. 4 Punkte können bei normaler Funktion bzw. keiner Einschränkung vergeben werden, 0 Punkte entsprechend bei vollständigem Funktionsverlust. Anhand dieser Funktionsskala (0–48 Punkte, 48 Punkte = beste motorische Funktion) kann die Krankheitsschwere beurteilt werden. In jeder Subdomäne konnte eine Punktzahl zwischen 0–12 Punkten erreicht werden. Eine hohe Punktzahl spricht als Ergebnis für eine bessere Funktion. Bewertet wird der gegenwärtige Zeitpunkt.

Für japanische Teilnehmende wird der japanische (Ohashi) ALSFRS-R verwendet, außer für Item 5a (Schneiden von Lebensmitteln und Umgang mit Utensilien ohne Magensonde) und Item 11 (Orthopnea), bei dem der japanisch übersetzte Global ALSFRS-R verwendet wurde.

In der Studie VALOR wird der ALSFRS-R an jedem Studienstandort durch die gleiche Person, die in der Durchführung geschult und qualifiziert ist, durchgeführt. Es war zudem möglich, eine weitere Person des Studienzentrums als Vertretung zu ernennen, die ebenfalls in der Durchführung des ALSFRS-R geschult ist.

Es werden sowohl die Veränderung zu Baseline als auch Responderanalysen zur Verschlechterung und Verbesserung um  $\geq 15\%$  der Skalenspannweite berichtet.

Die Durchführung kann bei einer Visite bis zu 2 Tage vor der Einnahme des Studienmedikaments erfolgen.

#### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Die Operationalisierung der Responderanalysen ist zum Teil nicht nachvollziehbar. Es ist nicht definiert, ob das Ansprechen für die Analyse tatsächlich zum Zeitpunkt Woche 28 oder ggf. über die Zeit berücksichtigt wurde.

#### Patientenrelevanz

Eine Beurteilung der ALS-Symptomatik wird als patientenrelevant angesehen.

#### Validität

Der ALSFRS-R stellt eine Erweiterung des ALSFRS-Fragebogens (Amyotrophic Lateral Sclerosis Functional Rating Scale) dar. Der ALSFRS-Fragebogen besteht aus 4 Subskalen. Da im ALSFRS nur 1 Item die Atemfunktion abfragt, finden die anderen 3 Subskalen mit jeweils 3 Items entsprechend stärkere Berücksichtigung. Im ALSFRS-R wurden daher Items hinzugefügt mit dem Ziel, den Verlauf der respiratorischen Dysfunktion bei ALS besser zu untersuchen. Die revidierte Version des ALSFRS wurde initial anhand von 387 Patientinnen und Patienten mit ALS validiert [13]. Es bestehen Limitationen bezüglich des Aufbaus des Instruments. Es ist unklar, wie die 2 neu hinzugefügten Items zur respiratorischen Funktion des ALSFRS-R entwickelt wurden und inwiefern hierbei eine Patientenbeteiligung vorgesehen war. In der ursprünglichen Entwicklungsstudie konnte mittels PCA-Analyse eine Bildung von 4 Subscores, nicht jedoch eines Gesamtscore, gezeigt werden [1,13]. Auch Folgeuntersuchungen ergaben ein heterogenes Bild hinsichtlich der Bildung eines Gesamtscores und der Bildung von 3 bzw. 4 Domänen. Bei der beschriebenen Validierungsstudie wurde jedoch nicht die nicht-linear verlaufende 5-Punkte-Ordinalskala verwendet, sondern diese für die Analysen in eine linearisierte Skala umgewandelt [17,18,29]. Es wird empfohlen, neben dem Gesamtscore die Ergebnisse der Subskalen zu berichten [29]. Für das Messinstrument existieren Übersetzungen in mehrere Sprachen. Der ALSFRS-R wird häufig in klinischen Studien angewandt [15].

Insgesamt zeigten sich in mehreren Studien an über 1.000 Patientinnen und Patienten akzeptable Ergebnisse zur internen Konsistenz [13] sowie zur Inter-/Intrarater-Reliabilität des ALSFRS-R [2,13]. Untersuchungen deuten auf Unterschiede in der Interpretation und Erhebung der Items bei klinischen Studien zwischen den Studienzentren hin, die auf ein Fehlen einer einheitlichen SOP hinweisen [25]. Der pU legt für die Studie VALOR ein Manual zur Durchführung des ALSFRS-R vor, das zu einer einheitlichen Durchführung der Tests zwischen den Studienzentren beitragen soll.

Erste Informationen zur Konstruktvalidität und known-groups validity liegen für den ALSFRS-R-Gesamtscore [2,13] vor. Es zeigte sich z. T. nur eine geringe bis moderate Korrelation zwischen respiratorischen Parametern, wie dem FVC% bzw. SVC%, mit dem Gesamtscore und der Subdomäne „respiratorische Funktion“ [13,19,24]. Die Änderungssensitivität wurde in Longitudinalstudien untersucht [3,4,28].

Berücksichtigt werden Responder-Auswertungen mit einer Verbesserung/Verschlechterung um  $\geq 15\%$  der Skalenspannweite.

### **Fatigue Severity Scale (FSS)**

Der Endpunkt FSS wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Die Fatigue Severity Scale (FSS) ist ein selbstberichteter Fragebogen zur Bestimmung der wahrgenommenen Erschöpfung. Er umfasst 9 Fragen zur Schwere der Erschöpfung und ihre Auswirkungen auf Aktivitäten bei Personen mit einer Vielzahl von verschiedenen Erkrankungen (bspw. Beeinträchtigung der täglichen Aktivitäten, wie Arbeit, Familie oder soziales Leben) mit einem Bezugszeitraum von 2 Wochen. Die Bewertung erfolgt auf einer 7-Punkt-Likert-Skala, von 1 = „Ich stimme gar nicht zu“ bis 7 = „Ich stimme vollkommen zu“. Es kann ein Gesamtwert von 63 Punkten erreicht werden. Höhere Werte deuten auf eine schwerere Fatigue-Symptomatik hin.

Die Durchführung kann bei einer Visite bis zu 2 Tage vor der Einnahme des Studienmedikaments erfolgen.

Es werden sowohl die Veränderung zu Baseline als auch Responderanalysen zur Verschlechterung und Verbesserung um  $\geq 15\%$  der Skalenspannweite dargestellt.

##### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Die Operationalisierung der Responderanalysen ist zum Teil nicht nachvollziehbar. Es ist nicht definiert, ob das Ansprechen für die Analyse tatsächlich zum Zeitpunkt Woche 28 oder ggf. über die Zeit berücksichtigt wurde.

#### Patientenrelevanz

Der Endpunkt wird in der vorliegenden Indikation als patientenrelevant angesehen.

#### Validität

Der Endpunkt gilt als validiert und wurde bereits in vorherigen Nutzenbewertungsverfahren bewertet. Die FSS wird in zahlreichen klinischen Studien zur ALS eingesetzt. Studien zur Änderungssensitivität konnten nicht identifiziert werden. Berücksichtigt werden Responder-Auswertungen mit einer Verbesserung/Verschlechterung um  $\geq 15\%$  der Skalenspannweite.

### **Allgemeiner Gesundheitszustand (EQ-5D-5L VAS)**

Der Endpunkt „Allgemeiner Gesundheitszustand“ erhoben mittels EQ-5D-5L VAS wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Mit der EQ-5D-5L VAS bewerten die Teilnehmenden bezogen auf den aktuellen Tag den eigenen Gesundheitszustand auf einer Skala von 0 (schlechtester vorstellbarer Gesundheitszustand) bis 100 (bester vorstellbarer Gesundheitszustand) [27].

Die Durchführung kann bei einer Visite bis zu 2 Tage vor der Einnahme des Studienmedikaments erfolgen.

Es werden sowohl die Veränderung zu Baseline als auch Responderanalysen zur Verschlechterung und Verbesserung um  $\geq 15\%$  der Skalenspannweite berichtet.

##### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Die Operationalisierung der Responderanalysen ist zum Teil nicht nachvollziehbar. Es ist nicht definiert, ob das Ansprechen für die Analyse tatsächlich zum Zeitpunkt Woche 28 oder ggf. über die Zeit berücksichtigt wurde.

#### Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen.

#### Validität

Die Validität der EQ-5D-5L VAS wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als valide eingeschätzt. Berücksichtigt werden Responderauswertungen mit einer Verbesserung/Verschlechterung um  $\geq 15\%$  der Skalenspannweite.

### **Amyotrophic Lateral Sclerosis Assessment Questionnaire-5**

Der Endpunkt ALSAQ-5 wird in der Nutzenbewertung nicht berücksichtigt aufgrund unklarer Patientenrelevanz.

#### Operationalisierung

##### *Beschreibung*

Der „Amyotrophic Lateral Sclerosis Assessment Questionnaire“ (ALSAQ) ist ein PRO-Fragebogen, der speziell zur Messung des subjektiven Wohlbefindens von Patienten und Patientinnen mit ALS eingesetzt wird.

Der ALSAQ ist in einer Langform (ALSAQ-40) und einer Kurzform (ALSAQ-5) erhältlich. In der Studie VALOR wurde die Kurzversion ALSAQ-5 eingesetzt. Der Fragebogen als Version zur Selbstbeurteilung umfasst Fragen, die den 5 Dimensionen des ALSAQ-40 entsprechen sollen: körperliche Mobilität (Frage 1: „Ich habe es als schwierig empfunden aufzustehen“), Einschränkungen durch ALS (Frage 2: „Ich habe Schwierigkeiten meine Arme und Hände zu verwenden“), Aktivitäten des täglichen Lebens (Frage 3: „Ich habe Schwierigkeiten feste Nahrung zu mir zu nehmen“), Kommunikation (Frage 4: „Ich habe den Eindruck, mich versteht man beim Reden schlecht“) und emotionale Funktion (Frage 5: „Ich habe mich hoffnungslos bezüglich der Zukunft gefühlt“). Die Fragen werden auf einer 5-Punkte-Likert-Skala von 0 (nie) bis 4 (immer) beantwortet. Der Bezugsrahmen der Fragen umfasst die letzten 2 Wochen. Mit

dem ALSAQ-5-Gesamtscore können zwischen 0 und 100 Punkte erreicht werden, wobei eine niedrigere Punktzahl einen besseren gesundheitsbezogenen Status darstellt. [20,22]

In der Studie VALOR konnte die Erhebung während einer Visite 2 Tage vor Einnahme der Studienmedikation erfolgen.

Es werden sowohl die Veränderung zu Baseline als auch Responderanalysen zur Verschlechterung und Verbesserung um  $\geq 15\%$  der Skalenspannweite berichtet.

#### *Bewertung*

Der ALSAQ-5 erhebt vorwiegend Alltagsaktivitäten im Sinne motorischer Einschränkungen (sich bewegen: aufstehen, Arme und Hände bewegen, feste Nahrung zu sich nehmen), Kommunikation sowie Emotionen. Motorik sowie Sprache werden mit dem Fragebogen ALSFRS-R bereits detailliert erfasst, die emotionale Funktion über den SF-36. Es handelt sich hierbei um eine Doppelerhebung der beschriebenen Funktionen.

Darüber hinaus ist nicht definiert, wann das Ansprechen für die Analyse berücksichtigt wurde (bspw. Response bis zur ersten Verbesserung/Verschlechterung oder zu jeder Studienvisite).

#### Patientenrelevanz

Der Endpunkt ALSAQ-5 wird der Kategorie Morbidität zugeordnet, da der 5 Item umfassende Fragebogen nicht alle Dimensionen der gesundheitsbezogenen Lebensqualität adäquat abbilden kann.

Die Patientenrelevanz des Endpunktes ist in der vorliegenden Operationalisierung unklar.

#### Validität

Für den Endpunkt ALSAQ-5 liegen Validierungsstudien anhand von Personen mit ALS vor [20,21,23,24].

Es liegen für den Endpunkt Informationen zur Inhaltsvalidität, Reliabilität (interne Konsistenz), Konstruktvalidität und Änderungssensitivität vor. Angaben zur Test-Retest-Reliabilität und known groups validity konnten nicht identifiziert werden. Bei dem Endpunkt handelt es sich um eine Kurzversion des ALSAQ-40. Für den ALSAQ-5 wurden Items ausgewählt, die eine hohe Korrelation ( $r_s > 0,79$ ) zur gehörigen Subskala aufwiesen und nicht statistisch signifikant zum Gesamtscore waren. Jedoch konnte in einer weiteren Validierungsstudie dies nur bedingt bestätigt werden [20,21,23,24].

Neben den Limitationen bezüglich der Validität des Fragebogens deckt die Kurzversion des ALSAQ-40 nicht hinreichend alle Aspekte der gesundheitsbezogenen Lebensqualität ab und die Fragen fokussieren sich auf Aspekte der Morbidität. Motorische und sprachliche Einschränkungen werden mittels ALSFRS-R und die Lebensqualität mittels SF-36 erfasst. Auf eine Darstellung des Endpunktes in der Nutzenbewertung wird daher verzichtet.

#### **Work Productivity and Activity Impairment Questionnaire (WPAI)**

Der Endpunkt WPAI wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Patientenrelevanz.

Bei dem Endpunkt „Work Productivity and Activity Impairment Questionnaire“ (WPAI) handelt es sich um einen Fragebogen zur Erfassung primär gesundheitsökonomischer Aspekte in Bezug auf die Beeinträchtigung der Arbeitsproduktivität und Aktivitäten. Der Fragebogen umfasst 6 Fragen, die die Arbeitsproduktivität insgesamt und die Einschränkung täglicher Aktivitäten abdecken, und kann an eine spezifische Erkrankung angepasst werden. Der Fragebogen misst den selbstberichteten Produktivitätsverlust im Zusammenhang mit der Erkrankung während der letzten 7 Tage. Er besteht aus Fragen zur Abwesenheit von der Arbeit aufgrund der Erkrankung, zu den am Arbeitsplatz verbrachten Stunden, zu der auf die

Erkrankung zurückzuführende Verminderung in der Produktivität. Item 6 fragt Einschränkungen bei der Ausübung normaler Tätigkeiten ab.

WPAI-Ergebnisse werden als Beeinträchtigungs-Prozentsätze ausgedrückt, wobei höhere Zahlen auf eine größere Beeinträchtigung und eine geringere Produktivität hinweisen.

#### *Bewertung*

Die Operationalisierung ist nachvollziehbar.

#### Patientenrelevanz

Es handelt sich um ein Instrument zur Messung der Beeinträchtigung der Arbeitsproduktivität und Aktivitäten. Mit dem Fragebogen werden nicht alle Dimensionen der gesundheitsbezogenen Lebensqualität abgedeckt, weshalb der Fragebogen, anders als im Dossier Modul 4, nicht der Kategorie Lebensqualität, sondern der Kategorie Morbidität zugeordnet wird.

Die Beeinträchtigung täglicher Aktivitäten durch die Erkrankung (Frage 6) adressiert im Gegensatz zu den anderen 5 Fragen einen patientenrelevanten Aspekt. Diese werden jedoch bereits über den Fragebogen SF-36 abgedeckt, weshalb auf eine Darstellung der Ergebnisse zu Item 6 in der Nutzenbewertung verzichtet wird.

#### Validität

Da der Endpunkt aufgrund nicht-gegebener Patientenrelevanz für die meisten Fragen des Instruments nicht in der Nutzenbewertung dargestellt wird, wird auf eine Bewertung der Validität verzichtet.

### **Langsame Vitalkapazität (SVC)**

Der Endpunkt SVC wird in der Nutzenbewertung nicht dargestellt aufgrund unklarer Patientenrelevanz.

#### Operationalisierung

##### *Beschreibung*

Die langsame Vitalkapazität (slow vital capacity, SVC) misst die respiratorische Muskelfunktion und wird in der aufrechten Position gemessen. Es sollten 3–5 Messungen vorgenommen werden, die den Kriterien der „American Thoracic Society“ und der „European Respiratory Society“ folgten. Ein Bewerter oder eine Bewerterin überprüfen, ob diese Kriterien (mindestens 3 akzeptable Tests, davon 2 Tests mit der jeweils größten und zweitgrößten Anstrengung durchgeführt innerhalb einer Vitalkapazität von 150 ml) erfüllt sind.

Die Durchführung erfolgte gemäß dem „Pulmonary Procedure Manual“ der Studie [5].

Die Art der Durchführung (z. B. Tragen einer Maske bei der Messung ab Baseline) sollte über den gesamten Studienzeitraum gleichmäßig stattfinden.

Die Durchführung kann bei einer Visite bis zu 2 Tage vor der Einnahme des Studienmedikaments erfolgen.

##### *Bewertung*

Die Operationalisierung ist zum Teil nachvollziehbar. Es ist unklar, wie in der Studie VALOR die Auswertung der mehrfach durchgeführten Messungen erfolgte (z. B. Bildung eines Mittelwerts, Verwendung des bestmöglichen Ergebnisses).

#### Patientenrelevanz

Bei dem Endpunkt SVC handelt es sich um einen Laborparameter, der nicht per se patientenrelevant ist. Es konnten keine geeigneten Studien zur Surrogatvalidierung für eine

patientenrelevanten Endpunkt identifiziert werden. Die respiratorische Funktion wird zudem mittels ALSFRS-R untersucht.

#### Validität

Aufgrund fehlender Patientenrelevanz im vorliegenden Anwendungsgebiet wird auf eine Bewertung der Validität verzichtet.

#### **Handheld Dynamometry Megascore (HHD)**

Der Endpunkt „Handheld Dynamometry“ wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Patientenrelevanz.

#### Operationalisierung

##### *Beschreibung*

Ein Endpunkt zur Beurteilung der neuromuskulären Beeinträchtigung ist die Messung der Kraftfähigkeiten mittels „Handheld Dynamometry“ (HHD). Hierbei wird die quantitative Muskelkraft bewertet, welche die isometrische Stärke mehrerer Muskeln unter Verwendung der Standardpositionierung des Probanden bzw. der Probandin testet. Ungefähr 8 Muskelgruppen werden (pro Seite) sowohl in den oberen als auch in den unteren Extremitäten untersucht. Eine genaue Beschreibung der Testdurchführung wird im Manual beschrieben.

Der HHD-Megascore wird als Durchschnitt der normalisierten Z-Scores der insgesamt 16 Muskeln berechnet. Eine höhere Punktzahl oder eine positive Veränderung bedeuten eine Verbesserung.

Für das Dossier Modul 4 wird post-hoc die Änderung zu Tag 196 gegenüber Baseline dargestellt.

##### *Bewertung*

Die Operationalisierung ist nachvollziehbar.

#### Patientenrelevanz

Die mittels HHD gemessene Muskelkraft wird an 8 Muskelgruppen pro Seite quantitativ bestimmt. Der pU führt aus, dass die Muskelkraft eine wichtige Determinante sowohl für die Funktion als auch für das Überleben bei ALS sei, ohne dies weiter mit Quellen zu belegen. Es handelt sich um eine mechanische Messung der Muskelkräfte an einzelnen Muskelgruppen, bei der die unmittelbare Patientenrelevanz unklar ist. Eine Untersuchung von motorischen Einschränkungen wird bereits durch den ALSFRS-R vorgenommen.

#### Validität

Aufgrund fehlender Patientenrelevanz im vorliegenden Anwendungsgebiet wird auf eine Bewertung der Validität verzichtet.

#### **PGI-C und PGI-S**

Die Endpunkte „Patient Global Impression of Change“ (PGI-C) und „Patient Global Impression of Severity“ (PGI-S) werden in der Nutzenbewertung nicht berücksichtigt, da die Rücklaufquoten zu Baseline und den weiteren Studienvisiten in beiden Behandlungsarmen bei < 70 % lagen.

### 2.3.3 Lebensqualität

#### SF-36

Der Endpunkt SF-36 wird in der Nutzenbewertung berücksichtigt.

##### Operationalisierung

###### *Beschreibung*

Zur Erhebung der Lebensqualität wurde die Version 2 des SF-36 eingesetzt.

Der SF-36-Fragebogen beinhaltet 8 Dimensionen: körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperlicher Schmerz, allgemeine Gesundheitswahrnehmung, Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden sowie zwei Summenskalen (PCS und MCS). Die PCS schließt die Domänen-Skalen für körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperliche Schmerzen und allgemeine Gesundheitswahrnehmung ein. In die MCS gehen die Domänen-Skalen für Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden ein.

Es können Werte von 0 bis 100 angenommen werden, wobei ein höherer Wert einen besseren Gesundheitszustand widerspiegelt.

Die Erhebung des Endpunktes kann bei einer Visite bis zu 2 Tage vor der Einnahme des Studienmedikaments erfolgen.

Es werden sowohl die Veränderung zu Baseline als auch Responderanalysen zur Verschlechterung und Verbesserung um  $\geq 15\%$  der Skalenspannweite berichtet.

###### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Die Operationalisierung der Responderanalysen ist zum Teil nicht nachvollziehbar. Es ist nicht definiert, ob das Ansprechen für die Analyse tatsächlich zum Zeitpunkt Woche 28 oder ggf. über die Zeit berücksichtigt wurde.

##### Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen.

##### Validität

Die Validität des Endpunktes SF-36 wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als valide eingeschätzt.

### 2.3.4 Sicherheit

#### Unerwünschte Ereignisse

Der Endpunkt Unerwünschte Ereignisse wird in der Nutzenbewertung berücksichtigt.

##### Operationalisierung

###### *Beschreibung*

Als UE galt jegliches unerwünschte medizinische Ereignis, welches während der klinischen Studie auftrat oder sich verschlechterte. Ein UE konnte jegliches unangenehme und ungewollte körperliche Anzeichen, Symptom oder auch ein Laborparameter sein, welches während der Studie auftrat oder sich verschlechterte, unabhängig davon, ob dies mit der Studienmedikation in Verbindung gebracht werden konnte.

UE wurden von der ersten Dosis bis zur letzten Follow-up-Visite im Rahmen jeder Studienvsitede und allen telefonischen Follow-up-Kontakten aktiv von den Prüffärztinnen und Prüffärzten erfasst.



Alle erfassten UE wurden ebenfalls nach Systemorganklasse (System Organ Class, SOC) und Preferred Term (PT) gemäß MedDRA (Version 24.0) kodiert.

Ein schwerwiegendes unerwünschtes Ereignis (SUE) ist ein unerwünschtes medizinisches Ereignis, das bei einer beliebigen Dosis

- zum Tod führt
- nach Ansicht des medizinischen Prüfpersonals den Patienten oder die Patientin in unmittelbare Lebensgefahr bringt (lebensbedrohliches Ereignis); dies schließt jedoch kein Ereignis ein, das, wenn es in einer schwereren Form aufgetreten wäre, zum Tod hätte führen können
- einen stationären Krankenhausaufenthalt oder die Verlängerung eines bestehenden Krankenhausaufenthalts erfordert
- zu anhaltender oder erheblicher Behinderung/Invalidität führt
- zu einer kongenitalen Anomalie / einem Geburtsfehler führt
- ein medizinisch bedeutsames Ereignis ist, das nach Ansicht des ärztlichen Prüfpersonals den Teilnehmenden gefährden kann oder ein Eingreifen erfordert, um eines der oben aufgeführten Ereignisse zu verhindern

UE von besonderem Interesse wurden nicht erhoben.

**Vermeidung von Stürzen:** Gemäß Dossier Modul 4 wurde der PT „Sturz“ als „Vermeidung von Stürzen“ separat in der Kategorie Morbidität ausgewertet.

#### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Es konnten keine Angaben identifiziert werden, ob Ereignisse der Grunderkrankung herausgerechnet wurden.

**Vermeidung von Stürzen:** Für den Endpunkt „Vermeidung von Stürzen“ konnten keine Angaben zur Operationalisierung und Erhebung identifiziert werden.

#### Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen.

**Vermeidung von Stürzen:** Die Vermeidung von Stürzen wird als patientenrelevant angesehen.

#### Validität

Die Erhebung der Unerwünschten Ereignisse wird als valide angesehen. Es konnten keine Angaben identifiziert werden, ob UE der Grunderkrankung herausgerechnet wurden.

**Vermeidung von Stürzen:** Der PT „Sturz“ wurde als „Vermeidung von Stürzen“ der Kategorie Morbidität zugeordnet. Eine Begründung für die Auswahl bzw. Rationale hierfür konnte nicht identifiziert werden. Es wurde zudem nicht dargelegt, wie die Erhebung des Endpunktes „Vermeidung von Stürzen“ aus dem PT „Sturz“ in der multizentrischen Studie vorgenommen wurde. Der Endpunkt wird daher als UE PT „Sturz“ der Kategorie „Sicherheit“ zugeordnet.

### **Columbia Suicide Severity Rating Scale (C-SSRS)**

Der Endpunkt Columbia Suicide Severity Rating Scale wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Die C-SSRS ist ein standardisiertes klinisches Interview zur systematischen Bewertung und strengen Überwachung des Auftretens, der Intensität und der Häufigkeit von Suizid-

bezogenen Gedanken und Verhaltensweisen. Die einzelnen Fragen wurden ausgewählt basierend auf Forschungsergebnissen, die gezeigt haben, welche Aspekte (vergangener) suizidaler Gedanken und Verhaltensweisen zukünftiges suizidales Verhalten vorhersagen. Der Bezugszeitraum kann dabei variieren, je nach Forschungskontext, in dem die Skala eingesetzt wird.

In der Studie VALOR wird die Version vom 14. Januar 2009 verwendet, welche aus 3 Domänen (Suizidgedanken (5 Items) und suizidales Verhalten (6 Items) und selbstverletzendes Verhalten ohne suizidale Absichten (1 Item)) besteht. Die Items werden in einem dichotomen Antwortformat („ja/nein“) beantwortet. Zusätzlich wird durch ein weiteres Item abgefragt, ob die Person einen Suizidversuch unternommen hat. Zu Baseline werden suizidale Absichten ohne Bezugszeitraum für die gesamte Lebenszeit abgefragt und für die Post-Baseline-Visiten für die Zeit nach der letzten Visite.

#### Suizidgedanken

- Frage 1: Wunsch zu sterben
- Frage 2: Nicht-spezifische aktive Suizidgedanken
- Frage 3: Aktive Selbstmordgedanken mit beliebigen Methoden (nicht geplant) ohne Absicht zu handeln
- Frage 4: Aktive Selbstmordgedanken mit einer gewissen Handlungsabsicht, ohne spezifischen Plan
- Frage 5: Aktive Selbstmordgedanken mit konkretem Plan und Absicht

#### Suizidales Verhalten

- Frage 6: Vorbereitende Handlungen oder Verhaltensweisen
- Frage 7: Abgebrochener Versuch
- Frage 8: Unterbrochener Versuch
- Frage 9: Tatsächlicher Versuch
- Frage 10: Suizidales Verhalten
- Frage 11 (nur post-Baseline): Suizid

#### Selbstverletzendes Verhalten ohne suizidale Absichten

- Frage 12: Selbstverletzendes Verhalten ohne suizidale Absichten

In der Version vom 14. Januar 2009 enthält der Fragebogen neben Items zu suizidalen Gedanken zudem Items zur Intensität der Suizidgedanken. Diese sind in der Version, die in der Studie VALOR angewandt wird, nicht enthalten.

Wenn die Person eine der ersten 5 Fragen mit „Ja“ beantwortet, wird bei der Person von suizidalen Gedanken ausgegangen. Nur Personen, die Frage 2 mit „Ja“ beantwortet haben, beantworten auch die Fragen 3, 4 und 5. Für alle Teilnehmenden, die Frage 2 mit „Nein“ beantwortet haben, wird auch für die Fragen 3, 4 und 5 die Antwort „Nein“ angenommen. Es wird ein Score aus den ersten 5 Items zur Beurteilung der Ausprägung suizidaler Gedanken gebildet (0–5 Punkte). Der oder die Teilnehmende weist behandlungsbedingte Suizidgedanken auf, sofern nach Baseline die Person neue oder sich verschlimmernde Suizidgedanken angibt. Bei einem Score von 0 hat die Person keine suizidalen Gedanken.

Die Beantwortung einer der Fragen 6–10 mit „Ja“ bei einer Visite post-Baseline weist auf die Gefahr eines suizidalen Verhaltens hin.

Die Erhebung des Items 12 wird nicht im Dossier Modul 4 beschrieben.

Die Beantwortung des Fragebogens kann bei einer Visite bis zu 2 Tage vor der Einnahme des Studienmedikaments erfolgen.

#### *Bewertung*

Die Operationalisierung ist zum Teil nachvollziehbar.

Es finden sich keine Angaben, inwiefern die Patientinnen und Patienten selbst die Fragen beantworteten oder ob eine Befragung durch das Studienpersonal stattfand.

Im Dossier Modul 4 wird die Erhebung des Items 12 nicht beschrieben.

#### Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen.

#### Validität

Die Validität der C-SSRS wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrunde liegenden Anwendungsgebiet als valide angesehen.

Es bestehen Unklarheiten, inwiefern die Durchführung des Interviews durch geschultes Personal erfolgte.

#### **Mini Mental State Examination (MMSE)**

Der Endpunkt Mini Mental Status Test wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht-gegebener Patientenrelevanz im vorliegenden Anwendungsgebiet

#### Operationalisierung

##### *Beschreibung*

Der MMSE Standardversion 2.0 ist ein 5- bis 10-minütiger Test mit 11 Fragen, der verschiedene Bereiche der kognitiven Funktion untersucht: „Orientierung“, „Aufmerksamkeit“, „Sofortiges Erinnern“, „Kurzzeit-Erinnern“, „Sprache“ und „Fähigkeit, einfachen verbalen und schriftlichen Befehlen zu folgen“.

Es handelt sich primär um ein Screening-Verfahren zur Feststellung kognitiver Defizite.

Der Test wird in 2 Teilen durchgeführt und erfordert zuerst sprachliche Antworten auf gestellte Fragen und im zweiten Teil das Befolgen verbaler und schriftlicher Anweisungen, das Schreiben eines Satzes und Kopieren einer geometrischen Figur. Es kann ein Gesamtscore von 0–30 Punkten erreicht werden, wobei niedrigere Werte Einschränkungen der kognitiven Fähigkeiten repräsentieren.

Für einige Teilnehmende erfolgte die Erhebung zu allen Visiten z. T. vor oder auch nach Erhalt der Studienmedikation, während andere Teilnehmende die Erhebung nur zu einzelnen Visiten (Tag 85, 169, 197) absolvierten.

##### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Für einige Teilnehmende erfolgte die Erhebung zu allen Visiten z. T. vor oder auch nach Erhalt der Studienmedikation, während andere Teilnehmende die Erhebung nur zu einzelnen Visiten absolvierten.

#### Patientenrelevanz

Personen mit ALS können zu Beginn der Erkrankung leichte kognitive Einschränkungen aufweisen oder diese im Verlauf der Erkrankung entwickeln. In der Regel spielen diese jedoch keine entscheidende Rolle für die Patienten und Patientinnen, da andere Beschwerden im Vordergrund stehen. Darüber hinaus können die Symptome kognitiver Einschränkungen auch auf andere Begleiterkrankungen (z. B. Depressionen) zurückzuführen sein. Die Patientenrelevanz des Endpunktes wird im vorliegenden Anwendungsgebiet als unklar eingeschätzt.

Aufgrund dessen wird der Endpunkt in der vorliegenden Operationalisierung und im Anwendungsgebiet als nicht patientenrelevant eingestuft.

## Validität

Aufgrund nicht gegebener Patientenrelevanz im vorliegenden Anwendungsgebiet wird die Validität des Endpunktes nicht bewertet.

### 2.3.5 Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 6.

*Tabelle 6: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie VALOR*

Studienvisite (Woche) Endpunkt	Baseline	Woche 2	Woche 4	Woche 8, 12, 16, 20, 24	Woche 28 EOS <sup>1)</sup>	Woche 32 Sicherheit/ alternative EOS <sup>2)</sup>
Mortalität <sup>3)</sup>	x					
Zeit bis zur dauerhaften Beatmung <sup>3)4)</sup>	x					
Zeit bis zum Tod oder zur dauerhaften Beatmung <sup>4)5)</sup>	x					
ALSFRS-R	x	x	x	x	x	
FSS	x		x	x	x	
EQ-5D-5L VAS	x		x	x	x	
SF-36	x			x	x	
Unerwünschte Ereignisse <sup>6)</sup>	x					
C-SSRS	x			x	x	

<sup>1)</sup> Visite zu Woche 28 oder innerhalb von 4 Wochen nach Einnahme der letzten Dosis bei Teilnehmenden mit vorzeitigem Abbruch der Studie.

<sup>2)</sup> Visite zu Woche 32 oder 8 Wochen nach der letzten Dosis für Teilnehmende, die nicht an der Verlängerungsstudie OLE teilnahmen. Für Teilnehmende nach Woche 28 mit Verzögerung des Einschusses in die Extensionsstudie OLE wird diese Visite entweder persönlich oder telefonisch durchgeführt. Die Visite kann zu Woche 32 oder 2 Tage vor Einschluss in die Extensionsstudie OLE erfolgen.

<sup>3)</sup> Bestandteil des kombinierten Endpunktes „Zeit bis zum Tod oder zur dauerhaften Beatmung“.

<sup>4)</sup> Es erfolgte eine tägliche Dokumentation der Beatmung durch den Patienten bzw. die Patientin im Beatmungstagebuch (auf Papier oder elektronisch). Neben dem Beatmungstagebuch können weitere Quellendokumente verwendet werden. Es liegt im Ermessen des Prüfarztes oder der Prüfarztin, die Anzahl der Beatmungsstunden pro Tag und die Art der verwendeten Beatmungsunterstützung zu bestimmen. Die Daten der verschiedenen Quellen werden, mit Ausnahme der elektronischen Tagebuchdaten, die als eigene Quelle dienen, in das Beatmungsprotokoll des CRF eingetragen. Das ärztliche Prüfpersonal überprüft zu jeder Visite die Eintragungen des Beatmungstagebuchs.

<sup>5)</sup> Eine Erfassung von Todesfällen war bis Woche 32 bzw. 8 Wochen nach der letzten Dosis bei Teilnehmenden ohne Teilnahme an der Verlängerungsstudie OLE vorgesehen.

<sup>6)</sup> Nach der Visite zur Dosisverabreichung fand mit den Patientinnen und Patienten ein telefonischer Kontakt statt.

Abkürzungen: ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; C-SSRS: Columbia Suicide Severity Rating Scale; CRS: Case report form; EOS: End of study; EQ-5D EuroQoL Five Dimension Questionnaire; FSS: Fatigue Severity Scale; OLE: Open-Label-Extension; SF-36: 36-Item Short Form Health Survey; VAS: Visual Analogue Scale

## 2.4 Statistische Methoden

Für die Studie VALOR liegt der SAP V2.0 (14. August 2021) vor, der nach der letzten Studiervisite und Datenschnitt vom 16. Juli 2021 erstellt wurde. Dieser umfasst u. a. Änderungen zur primären Wirksamkeitsanalyse des ALSFRS-R sowie zur Zeit bis zum Tod oder zur dauerhaften Beatmung auf Wunsch der FDA.

Darüber hinaus werden weitere post-hoc durchgeführte Auswertungen zu den Wirksamkeitsendpunkten (ALSFRS-R, Zeit bis zum Tod oder zur dauerhaften Beatmung) beschrieben, die Analysen u. a. ohne eine Adjustierung des ersten Stratifizierungsfaktors umfasste.

Die Randomisierung erfolgte anhand folgender Stratifizierungsfaktoren (Details siehe Abschnitt 2.2)

- Erfüllung der prognostischen Kriterien für eine schnelle Krankheitsprogression (1 der ersten 2 Kriterien trifft zu):
  - Abfall in der ALSFRS-R von  $\geq 0,2$  pro Monat (berechnet als [48-Baseline-Score]/Zeit seit Symptombeginn) mit im Protokoll definierten SOD1-Mutationen
  - oder
  - keine der zuvor genannten SOD1-Mutationen und ein Abfall der ALSFRS-R vor Randomisierung von  $\geq 0,9$  pro Monat (berechnet als [48-Baseline-Score]/Zeit seit Symptombeginn)“
  - oder
  - Kriterien für alle anderen teilnahmeberechtigten Personen: SOD1-Mutation, die nicht im Protokoll definiert ist (Kriterium für Abfall in der ALSFRS-R besteht nicht).
- Einnahme von Edaravon zu Baseline
- Einnahme von Riluzol zu Baseline
  - Die gleichzeitige Anwendung von beiden Wirkstoffen, Edaravon und Riluzol, wurde nicht als eigenständiger Stratifizierungsfaktor definiert. Bei dieser Konstellation wurde anhand Edaravon stratifiziert.

Der pU führt hierzu im Dossier Modul 4 aus, dass die für die Randomisierung ausgewählten Stratifizierungsfaktoren zur Prognose einer Krankheitsprogression (definierte SOD1-Mutationen und ein Abfall in der ALSFRS-R von  $\geq 0,2$  pro Monat oder nicht-protokoll-definierte SOD1-Mutation mit einem Abfall in der ALSFRS-R vor Randomisierung von  $\geq 0,9$  pro Monat) nicht ausreichend geeignet waren, um die Studienarme hinsichtlich Krankheitsprogression zu balancieren (siehe Abschnitt 2.5 und Abschnitt 4.2.).

### Analysepopulationen

- ITT-Population: Alle Personen, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben.
- Sicherheits-Population: entspricht der ITT-Population.

Die primären Analysen der Wirksamkeitsendpunkte waren zunächst für die mITT-Population (definiert als alle Teilnehmenden der ITT-Population, die die Kriterien einer rapiden Krankheitsprogression erfüllen) vorgesehen. Auswertungen anhand der ITT-Populationen waren als Sekundär- oder Sensitivitätsanalysen beschrieben.

### Datenschnitte

Finaler Datenschnitt: 16. Juli 2021

### Präspezifizierte Subgruppenanalysen

Für den primären Endpunkt wurden folgende Subgruppenanalysen gemäß SAP präspezifiziert:

- Geschlecht (männlich, weiblich)
- Baseline-Krankheitsdauer seit Symptombeginn (Terzile bzw. Mediane bei unzureichender Datenlage)
- Baseline-NfL-Plasmakonzentrationen (Terzile bzw. Mediane bei unzureichender Datenlage)
- Lokalisation (bulbär, andere): Wenn bei einem Teilnehmenden die Krankheit an mehreren Körperregionen beginnt und diese bulbäre Körperstellen umfasst, wird die Person der bulbären Körperstelle zugeordnet.
- Geographische Region (Nordamerika, Europa, Asien-Pazifik): Kanada wird Nordamerika zugeordnet.
- Riluzol- oder Edaravon-Einnahme. Sofern Edaravon bei < 5 Personen angewandt wird, wird der Gebrauch von Riluzol und Edaravon kombiniert (versus kein Gebrauch von Riluzol und Edaravon).

Es konnten für die im Dossier Modul 4 post-hoc ausgewerteten Subgruppenanalysen keine für die Nutzenbewertung relevanten Effektmodifikationen identifiziert werden.

### Präspezifizierung geplanter und durchgeführter Analysen

Sowohl für den kombinierten Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“, den Endpunkt „Gesamtüberleben“ und „Zeit bis zur dauerhaften Beatmung“ wurden Überlebenszeitanalysen durchgeführt. Mittels Cox-Proportional-Hazard-Methode wurde unter Berücksichtigung der Krankheitsdauer seit Symptombeginn, des ALSFRS-R-Gesamtscore zu Beginn der Studie und der Verwendung von Riluzol oder Edaravon Hazard Ratio [+ 95%-KI] und p-Wert in der ITT-Population berechnet. Kaplan-Meier-Kurven wurden dargestellt. Die Darstellung des p-Wertes erfolgte post-hoc für das Dossier Modul 4. Anstelle einer Adjustierung anhand des Stratifikationsfaktors „Erfüllung der prognostischen Kriterien für eine schnelle Krankheitsprogression“, der die Steigung des ALSFRS-R umfasste, wurde der ALSFRS-R-Gesamtscore zu Baseline verwendet.

Für den kombinierten Endpunkt „Zeit bis zum Tod oder Zeit bis zur dauerhaften Beatmung“ bzw. der „Zeit bis zur dauerhaften Beatmung“ wird angegeben, dass Patienten und Patientinnen Ereignisse aufweisen können, die in der Studie VALOR begannen und in der Verlängerungsstudie OLE fortbestanden. Es konnten jedoch keine eindeutigen Angaben identifiziert werden, ob Ereignisse in der Analyse berücksichtigt wurden, bei denen die Beatmung in der VALOR-Studie begonnen wurde, jedoch die Dauer, gemäß Endpunktdefinition, erst in der Verlängerungsstudie OLE erreicht wurde, da bei Nicht-Erfüllen der Endpunktdefinition Personen zum letzten Kontakt der Studie VALOR zensiert wurden.

Der pU reicht mit dem Dossier für die Endpunkte ALSFRS-R, FSS, „allgemeiner Gesundheitszustand“ erhoben über EQ-5D VAS, SF-36 Responderanalysen für eine Verbesserung oder Verschlechterung um  $\geq 15\%$  der Skalenspannweite zum Studienende Visite Woche 28 (Tag 197) ein. Vom pU wird jeweils der durchschnittliche Anteil an Personen mit entsprechender Response angegeben. Eine Angabe zu den tatsächlichen Ereigniszahlen fehlt und ein Rückschluss ist anhand der durchschnittlichen Anteile nicht möglich. Es ist unklar, wie der durchschnittliche Anteil berechnet wurde, z. B. über Strata oder ggf. über verschiedene Zeitpunkte hinweg und ob es sich um prozentuale Anteile oder dezimale Anteile handelt.

Das Relative Risiko [+ 95%-KI] und der p-Wert wurden mittels Cochran-Mantel-Haenszel-Modell unter Berücksichtigung der prognostischen Kriterien für schnelles Fortschreiten der Erkrankung berechnet.

Für die Sicherheitsdaten und den Endpunkt C-SSRS waren nur deskriptive Berechnungen präspezifiziert. Für das Dossier Modul 4 wurden das Relative Risiko [95%-KI] und der p-Wert berechnet, wobei unklar ist, inwiefern eine Adjustierung nach u. a. Stratifizierungsfaktoren vorgenommen wurde.

### Sensitivitätsanalysen

Für den Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“ wurden Überlebenszeitanalysen für die ITT-Population durchgeführt auf Grundlage imputierter Werte auf Basis von Angaben des ALSFRS-R Frage 12 zur respiratorischen Insuffizienz und zur Beatmung in der Nacht oder am Tag, anhand der Einträge in der klinischen Datenbank, unabhängig der Bewertung durch das EAC.

### **Fehlende Werte und Daten-Imputation**

Für die Endpunkte, die aus Skalenwerten gebildet wurden (ALSFRS-R, FSS, SF-36), sollten zunächst fehlende Einzelitems innerhalb von Visiten nach prädefinierten Algorithmen zur Bildung der Skalen-/Domän-Summenscores imputiert werden.

In der Analyse der Endpunkte (ALSFRS-R, FSS, SF-36, „allgemeiner Gesundheitszustand“ mittels EQ-5D-5L) zu den vorgesehenen Zeitpunkten wurden Multiple Imputationen (MI) verwendet. Eine MI wurde mit der Methode Markov Chain Monte Carlo (MCMC) unter Annahme multivariater Normalität zur Imputation von Post-Baseline-Werten jeweils für jeden Studienarm angewandt. In das MI-Modell gingen die Behandlungsgruppe sowie der Einsatz von Riluzol und Edaravon, der Baselinewert des Endpunktes sowie alle verfügbaren Post-Baseline-Werte ein. Imputierte Werte außerhalb des erwartbaren Bereichs der Skala werden zurückgesetzt. Die Plausibilität der Annahme (voraussichtlich Missing At Random) wird nicht diskutiert.

Fehlende Werte zu Baseline werden ebenfalls imputiert, sofern kein Wert zur Visite vor Einnahme der ersten Dosis der Studienmedikation verfügbar ist. Fehlt der Wert für Tag 1, wird für die Imputation der Zeitpunkt gewählt, der der ersten Dosis am nächsten liegt und davor liegt, als Baseline verwendet.

### **Einschätzung der statistischen Auswertungen**

Anstelle der im SAP spezifizierten Auswertungen zur mITT werden die für das Dossier post-hoc durchgeführten Auswertungen zur ITT herangezogen.

### Zeit bis zum Tod oder zur dauerhaften Beatmung

Die präspezifizierten Überlebenszeitanalysen und Berechnung von Hazard Ratio (+ 95%-KI] und p-Werten für den kombinierten Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“ und der Einzelkomponenten werden als geeignet angesehen. Es wird angegeben, dass Patienten und Patientinnen Ereignisse aufweisen konnten, die in der Studie VALOR begannen und in der Verlängerungsstudie OLE fortbestanden. Es konnten jedoch keine eindeutigen Angaben identifiziert werden, ob diese Ereignisse für die Analyse berücksichtigt wurden, bei denen die Beatmung in der VALOR-Studie begonnen wurde, jedoch die Dauer, gemäß Endpunktdefinition, erst in der Verlängerungsstudie OLE erreicht wurde, da bei Nicht-Erfüllen der Endpunktdefinition Personen zum letzten Kontakt der Studie VALOR zensiert werden sollten.

### Responderanalysen zu ALSFRS-R, FSS, SF-36, „allgemeiner Gesundheitszustand“ mittels EQ-5D-5L

Für die Auswertung der Endpunkte ALSFRS-R, FSS, SF-36, „allgemeiner Gesundheitszustand“ mittels EQ-5D-5L anhand der ITT-Population wurden fehlende Werte mittels multipler Imputation ersetzt. Die Plausibilität der Annahme (voraussichtlich „Missing At Random“) wird nicht diskutiert. Aufgrund fehlender Angaben zu den Annahmen wird davon ausgegangen, dass diese unter der Annahme eines zufälligen Fehlens ersetzt wurden. Inwieweit diese Annahme bei Studienabbrechern (11 % im Interventions- und 8 % im Kontrollarm) angemessen ist, wird nicht argumentiert. Andere Analysen, wie Sensitivitätsanalysen mit Imputation fehlender Werte als Non-Responder, wurden mit dem Dossier nicht eingereicht.

Für die Responderanalysen der Wirksamkeitsendpunkte ALSFRS-R, FSS, EQ-5D-5L VAS sowie SF-36 wird im Dossier Modul 4 der durchschnittliche Anteil mit entsprechender Response angegeben, ohne Angaben zur Berechnung des durchschnittlichen Anteils. Eine Angabe zu den tatsächlichen Ereigniszahlen fehlt und ein Rückschluss ist anhand der durchschnittlichen Anteile nicht möglich. Es ist unklar, wie der durchschnittliche Anteil berechnet wurde, z. B. über Strata oder ggf. über verschiedene Zeitpunkte hinweg, und ob es sich um prozentuale oder dezimale Anteile handelt. Es ist nicht definiert, ob das Ansprechen für die Analyse tatsächlich zum Zeitpunkt Woche 28 (Tag 197) oder ggf. über die Zeit berücksichtigt wurde. Auf eine Darstellung des durchschnittlichen Anteils an Respondern mit einer Verschlechterung oder Verbesserung um  $\geq 15$  % der Skalenspannweite wird verzichtet. Relatives Risiko und p-Werte werden ergänzend dargestellt. Allerdings kann die Plausibilität der Ergebnisse aufgrund der fehlenden Ereigniszahlen abschließend nicht beurteilt werden.

Die Endpunkte werden ergänzend dargestellt.

### Sicherheit

Es wurden post-hoc für die Sicherheitsendpunkte das relative Risiko und p-Werte berechnet, wobei Angaben zum Analyseverfahren und ob eine Adjustierung nach Stratifizierungsfaktoren durchgeführt wurde, fehlen.



## 2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 7: Verzerrungspotential der Studie VALOR

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patientin/Patient	Verblindung Behandlungsperson	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
VALOR	Ja	Ja	Ja	Ja	Nein	Unklar <sup>1)2)3)</sup>	Unklar

<sup>1)</sup> Die Erstellung der Randomisierungssequenz mittels zentralem IRT-System und verdeckte Zuteilung zu den Studienarmen wird als adäquat angesehen. Die Randomisierung erfolgte stratifiziert nach Einnahme von Riluzol zu Baseline, Einnahme von Edaravon zu Baseline sowie Erfüllen prognostischer Faktoren (definierte SOD1-Mutationen und ein Abfall in der ALSFRS-R von  $\geq 0,2$  pro Monat oder nicht-protokoll-definierte SOD1-Mutation mit einem Abfall in der ALSFRS-R vor Randomisierung von  $\geq 0,9$  pro Monat). Seitens des pU wird angemerkt, dass der Stratifizierungsfaktor „Erfüllen prognostischer Faktoren“ nicht ausreichend geeignet war, um die Studienarme hinsichtlich einer Krankheitsprogression zu balancieren.

<sup>2)</sup> Die vorliegenden Daten deuten auf Imbalancen in den Baseline-Charakteristika hin (u. a. Einschluss um im Median 4 Jahre jüngeren Personen im Interventions- gegenüber dem Kontrollarm, ALSFRS-R-Run-In-Steigung (Screening Tag 0 bis -15) erhöht um -0,3; NfL-Plasma-Konzentration um ca. 10 pg/ml höher gegenüber dem Kontrollarm). Es liegen jedoch nur bedingt Angaben zu den betroffenen Körperregionen sowie zur Krankheitsschwere (u. a. ALS Milano-Torinos Staging und King's staging) zu Baseline vor. Die Baseline-Charakteristika hinsichtlich der Stratifizierungsfaktoren zeigen keine Imbalancen. Eine Verzerrung zugunsten einer der beiden Behandlungsgruppen ist möglich, kann jedoch abschließend nicht beurteilt werden.

<sup>3)</sup> Im Dossier Modul 4 finden sich keine Angaben zu den Protokollverletzungen. Eine abschließende Beurteilung ist daher nicht möglich.

Abkürzungen: ALS: Amyotrophe Lateralsklerose; ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; IRT: Interactive Response Technology; NfL: Neurofilament light chain; SOD1: Superoxid dismutase 1

Das Verzerrungspotential auf Studienebene wird als unklar eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 8 dargestellt.

Tabelle 8: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie VALOR

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Gesamtüberleben	Ja	Ja	Nein	Nein	Niedrig
Zeit bis zum Tod oder zur dauerhaften Beatmung	Ja	Ja	Nein	Unklar <sup>1)</sup>	Unklar
ALSFRS-R	Ja	Ja <sup>2)</sup>	Nein	Unklar <sup>3)</sup>	Unklar
FSS	Ja	Ja <sup>2)</sup>	Nein	Unklar <sup>3)</sup>	Unklar
EQ-5D VAS	Ja	Ja <sup>2)</sup>	Nein	Unklar <sup>3)</sup>	Unklar

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
SF-36	Ja	Ja <sup>2)</sup>	Nein	Unklar <sup>3)</sup>	Unklar
Sicherheit	Ja	Ja	Nein	Unklar <sup>4)</sup>	Unklar
C-SSRS	Ja	Ja	Nein	Unklar <sup>4)</sup>	Unklar

<sup>1)</sup> Es wurden nur wenige Ereignisse bis zum Datenschnitt (16. Juli 2021) beobachtet (Intervention: n = 3 (4 %) vs. Kontrolle n = 2 (6 %)). Bei jeweils 1 Person in den Behandlungsgruppen betrug die Beatmungsdauer nur 1 bzw. 2 Tage und entspricht somit nicht den im Protokoll definierten Vorgaben (dauerhafte Beatmung für mindestens 22 Stunden an konsekutiv mindestens 21 Tagen). Es konnten jedoch keine Angaben identifiziert werden, ob Ereignisse in der Analyse berücksichtigt wurden, bei denen die Beatmung in der VALOR-Studie begonnen wurde, jedoch die Dauer, gemäß Endpunktdefinition, erst in der Verlängerungsstudie OLE erreicht wurde, da bei Nicht-Erfüllen der Endpunktdefinition Personen zum letzten Kontakt der Studie VALOR zensiert werden sollten.

<sup>2)</sup> Für die Analysen wurde für die Auswertung die ITT-Population zugrunde gelegt. Fehlende Werte wurden mittels multipler Imputation ersetzt. Aufgrund fehlender Angaben zu den Annahmen wird davon ausgegangen, dass diese unter der Annahme eines zufälligen Fehlens ersetzt wurden. Inwieweit diese Annahme bei Studienabbruchern (11 % im Interventions- und 8 % im Kontrollarm) angemessen ist, wird nicht argumentiert. Es werden zudem keine Angaben zu Anzahl und Gründen fehlender Werte vorgenommen.

<sup>3)</sup> Für die Responderanalysen wird der durchschnittliche Anteil mit entsprechender Response angegeben, ohne Angaben zur Berechnung des durchschnittlichen Anteils. Ein Rückschluss auf die Ereigniszahlen ist nicht möglich. Es ist nicht definiert, ob das Ansprechen für die Analyse tatsächlich zum Zeitpunkt Woche 28 oder ggf. über die Zeit berücksichtigt wurde. Auf eine Darstellung des durchschnittlichen Anteils an Respondern mit einer Verschlechterung oder Verbesserung um  $\geq 15\%$  der Skalenspannweite wird verzichtet. Das Relative Risiko und p-Werte werden ergänzend dargestellt. Allerdings kann die Plausibilität der Ergebnisse aufgrund der fehlenden Ereigniszahlen abschließend nicht beurteilt werden.

<sup>4)</sup> Es konnten keine Informationen identifiziert werden, inwiefern post-hoc für das Dossier Modul 4 adjustierte Effektschätzer (Relatives Risiko und 95%-KI) berechnet wurden.

Abkürzungen: ALS: Amyotrophe Lateralsklerose; ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; C-SSRS: Columbia Suicide Severity Rating Scale; EQ-5D: EuroQol Five Dimension Questionnaire; FSS: Fatigue Severity Scale; ITT: Intention-to-treat; KI: Konfidenzintervall; OLE: Open-Label-Extension; SOD1: Superoxid dismutase 1; SF-36: 36-Item Short Form Health Survey; VAS: Visuelle Analogskala

### 3 Ergebnisse der eingeschlossenen Studie

Es werden in der vorliegenden Nutzenbewertung die Ergebnisse zum Studienende der Studie VALOR für den finalen Datenschnitt (16. Juli 2021) dargestellt.

#### 3.1 Studiencharakteristika und Studienmedikation

In Tabelle 9 und in Tabelle 10 sind die Studiencharakteristika der Studie VALOR beschrieben.

*Tabelle 9: Allgemeine Angaben; Studie VALOR (ITT-Population)*

Studie	Tofersen N = 72 <sup>1)</sup>	Placebo N = 36 <sup>1)</sup>
ITT-Population <sup>2)</sup> , n (%)	72 (100,0)	36 (100,0)
Safety-Population <sup>2)</sup> , n (%)	72 (100,0)	36 (100,0)
Abbruch der Einnahme der Studienmedikation, n (%)	9 (12,5)	3 (8,3)
Aufgrund von:		
UE	3 (4,2)	0 (0,0)
Entzug der Einverständniserklärung	1 (1,4)	1 (2,8)
Tod	1 (1,4)	0 (0,0)
Krankheitsprogression	3 (4,2)	2 (5,6)
Andere	1 (1,4)	0 (0,0)
Abbruch der Studie, n (%)	8 (11,1)	3 (8,3)
Aufgrund von:		
UE	2 (2,8)	0 (0,0)
Entzug der Einverständniserklärung	2 (2,8)	1 (2,8)
Tod	1 (1,4)	0 (0,0)
Krankheitsprogression	3 (4,2)	2 (5,6)
Mediane Beobachtungsdauer <sup>3)4)5)</sup> Tage (min; max)	k. A.	k. A.
Zeit bis zum Abbruch der Studienmedikation, Tage (min; max)	k. A.	k. A.

<sup>1)</sup> Alle randomisierten Patientinnen und Patienten

<sup>2)</sup> Definition siehe Kapitel 2.4 Statistische Methoden

<sup>3)</sup> Die Beobachtungsdauer wurde wie folgt berechnet: Letzter Tag in der Studie – Zeitpunkt Erhalt der ersten Dosis der Studienmedikation + 1. Letzter Tag in der Studie war definiert als der Zeitpunkt der letzten Studienvisite oder Untersuchung, des telefonischen Kontakts, Zeitpunkts des Versterbens.

<sup>4)</sup> Die Dauer der Exposition mit dem Studienmedikament entspricht, aufgrund der langen Halbwertszeit des Arzneimittels, der Beobachtungsdauer.

<sup>5)</sup> Im Dossier Modul 4 konnten keine Angaben zur Beobachtungsdauer identifiziert werden. Die Beobachtungsdauer war bis zum Studienende zu Woche 28 (Tag 197) vorgesehen. Eine Visite zur Nachbeobachtung konnte zu Woche 32 erfolgen bei verspätetem Eintritt in die Verlängerungsstudie OLE oder falls keine Teilnahme an dieser erfolgte. Nach interner Prüfung ergibt sich hinsichtlich der medianen Beobachtungsdauer keine Diskrepanz zwischen den Studienarmen.

Abkürzungen: ITT: Intention-to-treat; k. A.: keine Angabe; OLE: Open-Label-Extension; UE: Unerwünschtes Ereignis

*Tabelle 10: Charakterisierung der Studienpopulation; Studie VALOR (ITT-Population)*

Studie	Tofersen 100 mg N = 72	Placebo N = 36
<i>Alter (Jahre)</i>		
Mittelwert (SD)	48,1 (12,64)	51,2 (11,57)
Median (min; max)	47,5 (23; 78)	51,5 (28; 73)

Studie	Tofersen 100 mg N = 72	Placebo N = 36
<i>Altersgruppe (Jahre), n (%)</i>		
18 bis < 35	10 (13,9)	2 (5,6)
35 bis < 50	32 (44,4)	15 (41,7)
50 bis < 65	21 (29,2)	14 (38,9)
≥ 65	9 (12,5)	5 (13,9)
<i>Geschlecht, n (%)</i>		
Männlich	43 (59,7)	19 (52,8)
Weiblich	29 (40,3)	17 (47,2)
<i>Abstammung (genetisch), n (%)</i>		
Asiatisch	72 (100)	36 (100)
Schwarz oder Afroamerikanisch	5 (6,9)	4 (11,1)
Weiß	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
Keine Angabe	44 (61,1)	25 (69,4)
Andere	21 (29,2)	7 (19,4)
	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
<i>Ethnizität, n (%)</i>		
Hispanisch oder Latino	4 (5,6)	1 (2,8)
Nicht-Hispanisch oder -Latino	47 (65,3)	28 (77,8)
Nicht berichtet	21 (29,2)	7 (19,4)
<i>Region, n (%)</i>		
EMEA	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
Nordamerika		
Asien/Pazifik		
<i>Begleitende medikamentöse Therapie (≥ 15 % in einer Behandlungsgruppe), n (%)</i>	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
<i>Begleitende nicht-medikamentöse Therapie (≥ 5 % in einer Behandlungsgruppe), n (%)</i>	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
Krankheitsspezifische Charakteristika		
<i>Zeit seit Symptombeginn bis Randomisierung (Monate), Median (min; max)</i>	11,4 (1,7; 145,7)	14,6 (2,4; 103,2)
<i>Zeit seit ALS-Diagnose bis Randomisierung (Monate), Median (min; max)</i>	4,1 (0,5; 101,8)	5,5 (0,2; 66,8)
<i>Betroffene Körperregion, n (%)</i>		
bulbäre	3 (4,2)	3 (8,3)
untere Extremitäten	46 (63,9)	26 (72,2)
obere Extremitäten	20 (27,8)	7 (19,4)
respiratorisch	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
multiple Regionen	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
<i>ALSFRS-R Steigung vor Randomisierung<sup>2)</sup></i>		
Mittelwert (SD)	-1,08 (1,37)	-1,16 (1,19)
Median (min; max)	-0,75 (-8,30; 0,00)	-0,89 (-4,91; -0,02)
<i>ALSFRS-R Run-In-Steigung<sup>3)</sup></i>		
Mittelwert (SD)	-1,0 (2,19)	-0,7 (3,25)
Median (min; max)	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
<i>ALSFRS-R-Gesamtscore zu Baseline</i>		
Mittelwert (SD)	36,9 (5,91)	37,3 (5,81)
Median (min; max)	k. A. (15; 48)	k. A. (24; 47)

Studie	Tofersen 100 mg N = 72	Placebo N = 36
<i>Bestätigte SOD1-Mutation, n (%)</i>		
Ja	72 (100,0)	36 (100,0)
Nein	0 (0,0)	0 (0,0)
<i>Gemäß Protokoll definierte SOD1-Mutation, n (%)</i>		
Ja	17 (23,6)	8 (22,2)
Nein	55 (76,4)	28 (77,8)
<i>SVC % des Vorhersagewertes</i>		
MW (SD)	82,1 (16,6)	85,1 (16,5)
<i>Plasma-NfL zu Baseline (pg/ml)</i>		
MW (SD)	100,4 (82,8)	89,7 (86,5)
<i>Stratifikationsfaktor: Krankheitsprogression, n (%)</i>		
Schnelle Progression	39 (54,2)	21 (58,3)
Andere	33 (45,8)	15 (41,7)
<i>Stratifikationsfaktor: Riluzol-Vortherapie, n (%)</i>		
Ja	45 (62,5)	22 (61,1)
Nein	27 (37,5)	14 (38,9)
<i>Stratifikationsfaktor: Edaravon-Vortherapie<sup>4)</sup>, n (%)</i>		
Ja	6 (8,3)	3 (8,3)
Nein	66 (91,7)	33 (91,7)
<i>ALS Milano-Torinos Staging (MITOS) zu Baseline, n (%)</i>	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
<i>King's stage zu Baseline, n (%)</i>	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>
<i>Beatmung zu Baseline, n (%)</i>	k. A. <sup>1)</sup>	k. A. <sup>1)</sup>

<sup>1)</sup> Im Dossier Modul 4 werden hierzu keine Angaben vorgenommen.

<sup>2)</sup> Bestandteil des Stratifizierungsfaktors Krankheitsprogression, errechnet wie folgt: (ALSFRS-R-Gesamtscore zu Baseline – ALSFRS-R-Gesamtscore von 48 Punkten) / Dauer der Krankheitssymptomatik (berechnet wie folgt: Zeitpunkt der Messung des ALSFRS-R-Gesamtscores zu Baseline – Zeitpunkt des Beginns der ALS-Krankheitssymptomatik)/30,4375)

<sup>3)</sup> Berechnet für den Zeitraum zwischen Screening und Tag 15.

<sup>4)</sup> Sofern Riluzol und Edaravon gemeinsam appliziert wurden, wurde anhand Edaravon stratifiziert. Die Anzahl an Personen im Interventions- und Kontrollarm entspricht der der Applikation von Edaravon alleine.

Abkürzungen: ALS: Amyotrophe Lateralsklerose; ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; EMEA: Europa, Naher Osten, Afrika; ITT: Intention-to-treat; k. A.: keine Angabe; MW: Mittelwert; NfL: Neurofilament light chain; SD: Standardabweichung; SOD: Superoxid-Dismutase Typ 1; SVC: Slow vital capacity

### Protokollverletzungen

Im Dossier Modul 4 finden sich keine Angaben zu den Protokollverletzungen. Eine abschließende Beurteilung ist daher nicht möglich.

### Begleitmedikation

Im Dossier Modul 4 werden keine Informationen zu den nicht-medikamentösen und medikamentösen Begleittherapien gegeben. Nach interner Prüfung ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

## 3.2 Mortalität

### Gesamtüberleben

In der Studie VALOR (Teil C) ist 1 Person (1,4 %) im Tofersen-Arm verstorben, im Placeboarm ist keine Person verstorben. Aufgrund der geringen Anzahl an Ereignissen konnte die mediane Zeit bis zum Ereignis nicht bestimmt werden. Ein Effektschätzer wurde nicht berechnet.

## 3.3 Morbidität

### Zeit bis zum Tod oder zur dauerhaften Beatmung

Die Ergebnisse zum kombinierten Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“ und „Zeit bis zur dauerhaften Beatmung“ sind in Tabelle 11 bzw. Tabelle 12 dargestellt.

Die mediane Dauer der Beatmung für mindestens 22 Stunden betrug für Personen des Interventionsarms (n = 3 (4,2 %)), die den Endpunkt erreicht haben, 21 (min; max: 1; 38) Tage. Für die 2 Personen des Kontrollarms, die den Endpunkt erreichten, wird eine Dauer von 2 und 28 Tagen angegeben. Für jeweils 1 Person in beiden Behandlungsarmen entspricht die Dauer nicht den Kriterien einer dauerhaften Beatmung (mindestens 21 Tage). Im Dossier Modul 4 wird darauf verwiesen, dass Patienten und Patientinnen Ereignisse aufweisen konnten, die in der Studie VALOR begannen und in der Verlängerungsstudie OLE fortbestanden.

### Zeit bis zum Tod oder zur dauerhaften Beatmung

*Tabelle 11: Ergebnisse für Zeit bis zum Tod oder zur dauerhaften Beatmung zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR, ITT-Population*

Studie VALOR (Teil C) Zeit bis zum Tod oder zur dauerhaften Beatmung <sup>1)</sup>	Tofersen N = 72 n (%)	Placebo N = 36 n (%)	Hazard Ratio [95%-KI] <sup>2)</sup> ; p-Wert <sup>2)</sup>
Ereignisse <sup>3/4)</sup>	4 (5,6)	2 (5,6)	0,97 [0,17; 5,71]; 0,98
Tod	1 (1,4)	0	
Dauerhafte Beatmung <sup>3)</sup>	3 (4,2)	2 (5,6)	
Anzahl Zensierungen <sup>5)</sup>	68 (94,4)	34 (94,4)	

<sup>1)</sup> Endpunkt definiert als Tod oder dauerhafte Beatmung ( $\geq 22$  Stunden/Tag invasiv oder nicht-invasiv durchgehend für  $\geq 21$  Tage) (je nachdem, was als erstes auftritt). Für jeweils 1 Person in beiden Behandlungsarmen entspricht die Dauer nicht den Kriterien einer dauerhaften Beatmung (mindestens 21 Tage).

<sup>2)</sup> Cox-Proportional-Hazard-Modell adjustiert für die Krankheitsdauer seit Symptombeginn, den ALSFRS-R-Gesamtscore zu Beginn der Studie und die Verwendung von Riluzol oder Edaravon.

<sup>3)</sup> Es werden nur Ereignisse berücksichtigt, die vom Endpunkt-Beurteilungsausschuss (EAC) beurteilt wurden.

<sup>4)</sup> Aufgrund der geringen Anzahl an Ereignissen konnte die mediane Zeit bis zum Ereignis nicht bestimmt werden.

<sup>5)</sup> Patientinnen und Patienten, die die Endpunktdefinition nicht erfüllen, werden zum letzten Kontakt in der Studie VALOR zensiert.

Abkürzungen: ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; KI: Konfidenzintervall; ITT: Intention-to-treat; N: Anzahl Personen in der Auswertung; n: Anzahl Personen mit Ereignis.

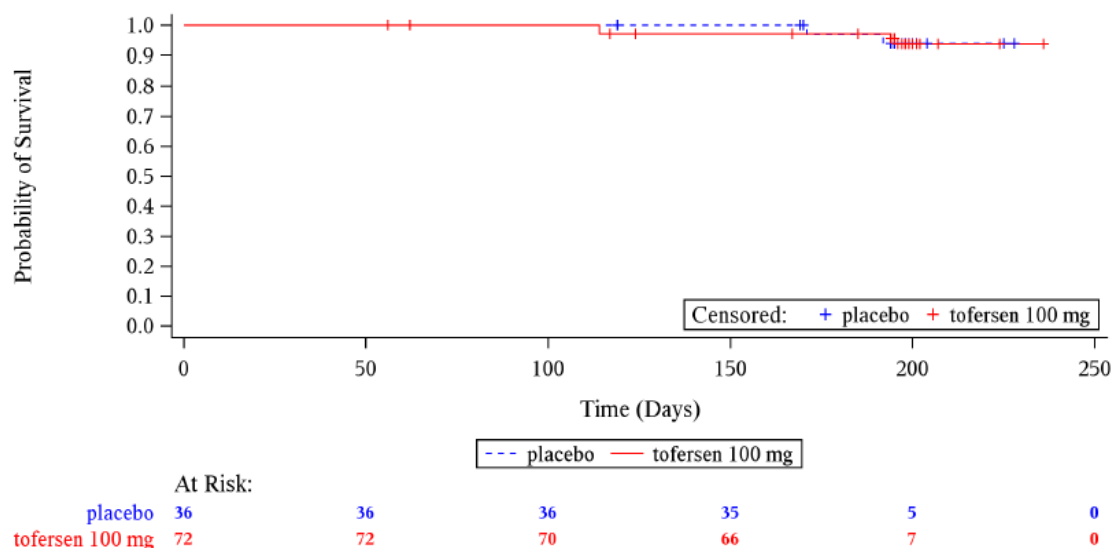


Abbildung: Kaplan-Meier-Kurve für den Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung, Studie VALOR (Datenschnitt: 16. Juli 2021); ITT-Population [7]

### Zeit bis zur dauerhaften Beatmung

Tabelle 12: Ergebnisse für Zeit bis zur dauerhaften Beatmung zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; ITT-Population

Studie VALOR (Teil C) Zeit bis zur dauerhaften Beatmung <sup>1)</sup>	Tofersen N = 72 n (%)	Placebo N = 36 n (%)	Hazard Ratio [95%-KI]; p-Wert <sup>2)</sup>
Ereignisse <sup>3)4)</sup>	3 (4,2)	2 (5,6)	0,82 [0,13; 5,34];
Zensierungen <sup>5)</sup>	69 (95,8)	34 (94,4)	0,83

<sup>1)</sup> Endpunkt definiert als dauerhafte Beatmung ( $\geq 22$  Stunden/Tag invasiv oder nicht-invasiv durchgehend für  $\geq 21$  Tage)

<sup>2)</sup> Cox-Proportional-Hazard-Modell adjustiert für die Krankheitsdauer seit Symptombeginn, den ALSFRS-R-Gesamtscore zu Beginn der Studie und die Verwendung von Riluzol oder Edaravon.

<sup>3)</sup> Aufgrund der geringen Anzahl an Ereignissen konnte die mediane Zeit bis zum Ereignis nicht bestimmt werden.

<sup>4)</sup> Es werden nur Ereignisse berücksichtigt, die vom Endpunkt-Beurteilungsausschuss (EAC) beurteilt wurden.

<sup>5)</sup> Patientinnen und Patienten, die die Endpunktdefinition nicht erfüllen, werden zum letzten Kontakt in der Studie VALOR zensiert.

Abkürzungen: ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; ITT: Intention-to-treat; KI: Konfidenzintervall; N: Anzahl Personen in der Auswertung; n: Anzahl Personen mit Ereignis

### **Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised (ALSFRS-R) – ergänzend dargestellt)**

Für den ALSFRS-R werden die Ergebnisse zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR für die ITT-Population in Tabelle 13 dargestellt.

**Tabelle 13: Ergebnisse für Veränderung im ALSFRS-R-Gesamtwert um  $\geq 15$  % bis zum Datenschnitt 16. Juli 2021 in der Studie VALOR, ITT-Population (ergänzend dargestellt)**

<b>Studie VALOR (Teil C) Amyotrophic Lateral Sclerosis Functional Rating Scale - Revised<sup>1)</sup></b>	<b>Tofersen N = 72 n/N (%)<sup>2)</sup></b>	<b>Placebo N = 36 n/N (%)<sup>2)</sup></b>	<b>Relatives Risiko [95%-KI]<sup>3)4)</sup>; p-Wert</b>
<b>Gesamtscore<sup>5)</sup></b>			
Verschlechterung $\geq 15$ %	k. A.	k. A.	0,95 [0,48; 1,87] 0,87
Verbesserung $\geq 15$ % <sup>6)</sup>	k. A.	k. A.	0,60 [0,13; 2,76] 0,51
<b>Bulbäre Domäne<sup>7)</sup></b>			
Verschlechterung $\geq 15$ %	k. A.	k. A.	0,88 [0,35; 2,22] 0,79
Verbesserung $\geq 15$ % <sup>6)</sup>	k. A.	k. A.	0,22 [0,05; 0,91] 0,037
<b>Domäne Feinmotorik<sup>7)</sup></b>			
Verschlechterung $\geq 15$ %	k. A.	k. A.	0,77 [0,49; 1,22] 0,27
Verbesserung $\geq 15$ % <sup>6)</sup>	k. A.	k. A.	1,15 [0,12; 11,29] 0,91
<b>Domäne Grobmotorik<sup>7)</sup></b>			
Verschlechterung $\geq 15$ %	k. A.	k. A.	0,87 [0,54; 1,41] 0,58
Verbesserung $\geq 15$ % <sup>6)</sup>	k. A.	k. A.	1,53 [0,20; 11,74] 0,68
<b>Respiratorische Domäne<sup>7)</sup></b>			
Verschlechterung $\geq 15$ %	k. A.	k. A.	1,25 [0,64; 2,41] 0,52
Verbesserung $\geq 15$ % <sup>6)</sup>	k. A.	k. A.	0,39 [0,07; 2,13] 0,28

<sup>1)</sup> Es können 0 bis 48 Punkte erreicht werden, wobei 48 Punkte der besten motorischen Funktion ohne Einschränkungen entspricht. Für japanische Teilnehmende wird der japanische (Ohashi) ALSFRS-R verwendet, außer für Item 5a (Schneiden von Lebensmitteln und Umgang mit Utensilien ohne Magensonde) und Item 11 (Orthopnea), bei dem der japanisch übersetzte Global ALSFRS-R verwendet wird.

<sup>2)</sup> Es liegen keine Angaben zur Anzahl Responder pro Studienarm vor. Vom pU wird jeweils der durchschnittliche Anteil an Personen mit entsprechender Response zum Studienende Visite Woche 28 (Tag 197) angegeben, ohne Angaben zur Berechnung des durchschnittlichen Anteils. Aufgrund der fehlenden Angaben zu den Respondern kann eine abschließende Beurteilung der Effektschätzer nicht vorgenommen werden.

<sup>3)</sup> Tofersen vs. Placebo: Cochran-Mantel-Haenszel-adjustiertes RR, berechnet unter Berücksichtigung der prognostischen Kriterien für schnelles Fortschreiten der Erkrankung und der Verwendung von Riluzol oder Edaravon.

<sup>4)</sup> Für fehlende Daten wird eine Multiple Imputation durchgeführt, die die Verwendung von Riluzol oder Edaravon sowie die entsprechenden Baseline- und Post-Baselinerwerte für den jeweiligen Endpunkt miteinschließt. Angaben zu fehlenden Werten der Responderanalysen liegen nicht vor.

<sup>5)</sup> Zu Baseline lag der ALSFRS-R-Gesamtwert (MW (SD)) im Tofersen-Arm bei 36,9 (5,9) und im Placebo-Arm bei 37,3 (5,8).

<sup>6)</sup> Die Verbesserung um  $\geq 15$  % im ALSFRS-R-Gesamtwert bzw. der Subdomäne wird aus Transparenzgründen berichtet. In der vorliegenden Indikation ist nicht von einer Verbesserung der entsprechenden Funktionen auszugehen.

<sup>7)</sup> Es konnten keine Angaben zu den Baseline-Werten der Subdomänen identifiziert werden.

Abkürzungen: ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; ITT: Intention-to-treat; k. A.: keine Angabe; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SD: Standardabweichung



### Fatigue Severity Scale (FSS) – ergänzend dargestellt

Für den Endpunkt FSS werden die Ergebnisse für die ITT-Population zum Studienende (Datenschnitt: 16. Juli 2021) in Tabelle 14 dargestellt.

*Tabelle 14: Ergebnisse für die Veränderung im FSS-Gesamtwert um  $\geq 15\%$  zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; ITT-Population (ergänzend dargestellt)*

Studie VALOR (Teil C) Fatigue Severity Scale <sup>1)</sup>	Tofersen N = 72 n/N (%) <sup>2)</sup>	Placebo N = 36 n/N (%) <sup>2)</sup>	Relatives Risiko [95%-KI] <sup>3)4)</sup> ; p-Wert
Verschlechterung $\geq 15\%$	k. A.	k. A.	0,88 [0,51; 1,51] 0,30
Verbesserung $\geq 15\%$	k. A.	k. A.	1,02 [0,32; 3,24] 0,98

<sup>1)</sup> Es kann ein Gesamtwert von 63 Punkten erreicht werden. Höhere Werte deuten auf eine schwerere Fatigue-Symptomatik hin. Zu Baseline lag der FSS-Gesamtwert (MW (SD)) im Tofersen-Arm bei 37,1 (13,8) und im Placebo-Arm bei 37,7 (16,3).

<sup>2)</sup> Es liegen keine Angaben zur Anzahl Responder pro Studienarm vor. Vom pU wird jeweils der durchschnittliche Anteil an Personen mit entsprechender Response zum Studienende Visite Woche 28 (Tag 197) angegeben, ohne Angaben zur Berechnung des durchschnittlichen Anteils. Aufgrund der fehlenden Angaben zu den Respondern kann eine abschließende Beurteilung der Effektschätzer nicht vorgenommen werden.

<sup>3)</sup> Tofersen vs. Placebo: Cochran-Mantel-Haenszel-adjustiertes RR, berechnet unter Berücksichtigung der prognostischen Kriterien für schnelles Fortschreiten der Erkrankung und der Verwendung von Riluzol oder Edaravon.

<sup>4)</sup> Für fehlende Daten wird eine Multiple Imputation durchgeführt, die die Verwendung von Riluzol oder Edaravon sowie die entsprechenden Baseline- und Post-Baselinerwerte für den jeweiligen Endpunkt miteinschließt. Angaben zu fehlenden Werten der Responderanalysen liegen nicht vor.

Abkürzungen: FSS: Fatigue Severity Scale; ITT: Intention-to-treat; k. A.: keine Angabe; KI: Konfidenzintervall, MW: Mittelwert; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SD: Standardabweichung

### Allgemeiner Gesundheitszustand mittels EQ-5D-5L VAS – ergänzend dargestellt

Für den Endpunkt Allgemeiner Gesundheitszustand werden die Ergebnisse für die ITT-Population zum Studienende (Datenschnitt: 16. Juli 2021) in Tabelle 15 dargestellt.

*Tabelle 15: Ergebnisse für Veränderung im VAS-Gesamtwert um  $\geq 15\%$  zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; ITT-Population (ergänzend dargestellt)*

Studie VALOR (Teil C) EQ-5D-5L VAS <sup>1)</sup>	Tofersen N = 72 n/N (%) <sup>2)</sup>	Placebo N = 36 n/N (%) <sup>2)</sup>	Relatives Risiko [95%-KI] <sup>3)4)</sup> ; p-Wert
Verschlechterung $\geq 15\%$	k. A.	k. A.	0,73 [0,40; 1,33] 0,64
Verbesserung $\geq 15\%$	k. A.	k. A.	0,90 [0,16; 5,15] 0,91

<sup>1)</sup> Es können 0 (schlechtester vorstellbarer Gesundheitszustand) bis 100 (bester vorstellbarer Gesundheitszustand) erreicht werden. Zu Baseline lag der VAS-Gesamtwert (MW (SD)) im Tofersen-Arm bei 66,8 (19,6) und im Placebo-Arm bei 73,8 (16,9).

<sup>2)</sup> Es liegen keine Angaben zur Anzahl Responder pro Studienarm vor. Vom pU wird jeweils der durchschnittliche Anteil an Personen mit entsprechender Response zum Studienende Visite Woche 28 (Tag 197) angegeben, ohne Angaben zur Berechnung des durchschnittlichen Anteils. Aufgrund der fehlenden Angaben zu den Respondern kann eine abschließende Beurteilung der Effektschätzer nicht vorgenommen werden.

<sup>3)</sup> Tofersen vs. Placebo: Cochran-Mantel-Haenszel-adjustiertes RR, berechnet unter Berücksichtigung der prognostischen Kriterien für schnelles Fortschreiten der Erkrankung und der Verwendung von Riluzol oder Edaravon.

<sup>4)</sup> Für fehlende Daten wird eine Multiple Imputation durchgeführt, die die Verwendung von Riluzol oder Edaravon sowie die entsprechenden Baseline- und Post-Baselinerwerte für den jeweiligen Endpunkt miteinschließt. Angaben zu fehlenden Werten der Responderanalysen liegen nicht vor.

Abkürzungen: EQ-5D: EuroQol Five Dimension Questionnaire ITT: Intention-to-treat; k. A.: keine Angabe; KI: Konfidenzintervall, MW: Mittelwert; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SD: Standardabweichung; VAS: Visuelle Analogskala

### 3.4 Lebensqualität

#### SF-36 – ergänzend dargestellt

Für den Endpunkt SF-36 werden in Tabelle 16 die Ergebnisse für die ITT-Population zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR dargestellt.

*Tabelle 16: Ergebnisse für Veränderung im SF-36 um  $\geq 15$  % zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR, ITT-Population (ergänzend dargestellt)*

Studie VALOR (Teil C) SF-36 <sup>1)</sup>	Tofersen N = 72 n/N (%) <sup>2)</sup>	Placebo N = 36 n/N (%) <sup>2)</sup>	Relatives Risiko [95%-KI] <sup>3)4)</sup> ; p-Wert
Psychischer Summenscore (MCS) <sup>5)</sup>			
Verschlechterung $\geq 15$ %	k. A.	k. A.	0,73 [0,30; 1,75] 0,48
Verbesserung $\geq 15$ %	k. A.	k. A.	1,35 [0,41; 4,46] 0,63
Körperlicher Summenscore (PCS) <sup>6)</sup>			
Verschlechterung $\geq 15$ %	k. A.	k. A.	0,75 [0,37; 1,53] 0,43
Verbesserung $\geq 15$ %	k. A.	k. A.	0,54 [0,08; 3,47] 0,52

<sup>1)</sup> Es können Werte von 0 bis 100 angenommen werden, wobei ein höherer Wert einen besseren Gesundheitszustand widerspiegelt.

<sup>2)</sup> Es liegen keine Angaben zur Anzahl Responder pro Studienarm vor. Vom pU wird jeweils der durchschnittliche Anteil an Personen mit entsprechender Response zum Studienende Visite Woche 28 (Tag 197) angegeben, ohne Angaben zur Berechnung des durchschnittlichen Anteils. Aufgrund der fehlenden Angaben zu den Respondern kann eine abschließende Beurteilung der Effektschätzer nicht vorgenommen werden.

<sup>3)</sup> Tofersen vs. Placebo: Cochran-Mantel-Haenszel-adjustiertes RR, berechnet unter Berücksichtigung der prognostischen Kriterien für schnelles Fortschreiten der Erkrankung und der Verwendung von Riluzol oder Edaravon.

<sup>4)</sup> Für fehlende Daten wird eine Multiple Imputation durchgeführt, die die Verwendung von Riluzol oder Edaravon sowie die entsprechenden Baseline- und Post-Baselinerwerte für den jeweiligen Endpunkt miteinschließt. Angaben zu fehlenden Werten der Responderanalysen liegen nicht vor.

<sup>5)</sup> Für den Psychischen Summenscore (MCS) lag zu Baseline in der Interventionsgruppe ein mittlerer Score von 51,4 (SD 12,3) und im Kontrollarm von 50,9 (10,8) vor.

<sup>6)</sup> Für den Körperlichen Summenscore (PCS) lag zu Baseline in der Interventionsgruppe ein mittlerer Score von 35,4 (SD 9,6) und im Kontrollarm von 36,0 (10,8) vor.

Abkürzungen: ITT: Intention-to-treat; k. A.: keine Angabe; KI: Konfidenzintervall; MCS: Psychischer Summenscore; PCS: Körperlicher Summenscore; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SD: Standardabweichung; SF-36: 36-Item Short Form Health Survey

### 3.5 Sicherheit

Es werden die Ergebnisse zur Sicherheit für den Vergleich Tofersen gegenüber Placebo der Studie VALOR zum Studienende, Visite zu Woche 28 (Tag 197) dargestellt. Im Dossier Modul 4 konnten keine Angaben zur medianen Behandlungszeit identifiziert werden. Nach interner Prüfung sind die medianen Behandlungszeiten zwischen den Behandlungsarmen gleich.

*Tabelle 17: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE, die aufgetreten sind, zum Studienende (Datenschnitt: 16. Juli 2021) der Studie Valor; Sicherheitspopulation*

Personen mit mindestens einem ...	Tofersen <sup>1)</sup> N = 72 n (%)	Placebo <sup>1)</sup> N = 36 n (%)	Relatives Risiko [95%-KI] <sup>2)</sup> p-Wert <sup>2)</sup>
UE (ergänzend dargestellt)	69 (95,8)	34 (94,4)	-
UE CTCAE-Grad ≥ 3	12 (16,7)	4 (11,1)	1,50 [0,52; 4,32] 0,45
SUE	13 (18,1)	5 (13,9)	1,30 [0,50; 3,36] 0,59
UE, das zum Abbruch der Studienmedikation führte	k. A.	k. A.	4,56 [0,25; 82,5] <sup>3)</sup> 0,30 <sup>3)</sup>

<sup>1)</sup> Unerwünschte Ereignisse wurden vom ersten Tag des Erhalts der Studienmedikation bis zur Follow-Up-Visite Woche 28 bei Teilnahme bzw. Woche 32 bei keiner Teilnahme an der Verlängerungsstudie OLE erhoben.

<sup>2)</sup> Es konnten keine Informationen identifiziert werden, wie das relative Risiko [95%-KI] und der p-Wert berechnet wurde und ob eine Adjustierung (nach Stratifizierungsmerkmalen) erfolgte.

<sup>3)</sup> Der pU berichtet im Dossier Modul 4 keine Ereigniszahlen. Der angegebene Effektschätzer für den Endpunkt wird nach interner Prüfung als adäquat eingeschätzt.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; k. A.: keine Angabe; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis

Der pU legte keine zusätzlichen Auswertungen unter Nichtberücksichtigung von erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung vor. Es ist nicht auszuschließen, dass Ereignisse der Grunderkrankung mit in die UE-Erfassung eingingen.

### Unerwünschte Ereignisse

Unerwünschte Ereignisse mit Inzidenz ≥ 10 % sind in Tabelle 18 dargestellt. Signifikante Unterschiede zwischen den Studienarmen zeigen sich für die Preferred Terms Diarrhö und Nasopharyngitis sowie SOC Infektionen und parasitäre Erkrankungen (jeweils zugunsten von Tofersen).

*Tabelle 18: UE mit Inzidenz ≥ 10 % in mindestens einem der Behandlungsarme zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; Sicherheitspopulation*

MedDRA-Systemorganklasse Preferred Term	Tofersen <sup>1)</sup> N = 72 n (%)	Placebo <sup>1)</sup> N = 36 n (%)	Relatives Risiko [95%-KI] <sup>2)</sup> p-Wert <sup>2)</sup>
<b>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</b>	<b>49 (68,1)</b>	<b>29 (80,6)</b>	<b>0,84 [0,67; 1,06] 0,14</b>
Schmerzen während eines Eingriffs	41 (56,9)	21 (58,3)	0,98 [0,69; 1,37] 0,89
Sturz	17 (23,6)	15 (41,7)	0,57 [0,32; 0,99] 0,05
Post-Lumbalpunktion-Syndrom	13 (18,1)	11 (30,6)	0,59 [0,30; 1,19] 0,14
Komplikation nach einem Eingriff	3 (4,2)	4 (11,1)	0,38 [0,09; 1,59] 0,18

<b>MedDRA-Systemorganklasse</b> Preferred Term	<b>Tofersen<sup>1)</sup></b> N = 72 n (%)	<b>Placebo<sup>1)</sup></b> N = 36 n (%)	<b>Relatives Risiko</b> [95%-KI] <sup>2)</sup> p-Wert <sup>2)</sup>
<b>Erkrankungen des Nervensystems</b>	<b>47 (65,3)</b>	<b>22 (61,1)</b>	<b>1,07 [0,78; 1,5] 0,68</b>
Kopfschmerzen	33 (45,8)	16 (44,4)	1,03 [0,66; 1,61] 0,89
Parästhesie	6 (8,3)	6 (16,7)	0,50 [0,17; 1,44] 0,20
<b>Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen</b>	<b>40 (55,6)</b>	<b>18 (50,0)</b>	<b>1,11 [0,76; 1,64] 0,59</b>
Schmerz in einer Extremität	19 (26,4)	6 (16,7)	1,58 [0,69; 3,62] 0,28
Rückenschmerzen	15 (20,8)	2 (5,6)	3,75 [0,91; 15,5] 0,07
Arthralgie	10 (13,9)	2 (5,6)	2,50 [0,58; 10,8] 0,22
Myalgie	10 (13,9)	2 (5,6)	2,50 [0,58; 10,8] 0,31
Muskuläre Schwäche	4 (5,6)	4 (11,1)	0,50 [0,13; 1,89] 0,31
Nackenschmerzen	4 (5,6)	4 (11,1)	0,50 [0,13; 1,89] 0,31
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	<b>24 (33,3)</b>	<b>9 (25,0)</b>	<b>1,33 [0,69; 2,56] 0,39</b>
Ermüdung	12 (16,7)	2 (5,6)	3,00 [0,71; 12,7] 0,16
<b>Erkrankungen des Gastrointestinaltrakts</b>	<b>22 (30,6)</b>	<b>16 (44,4)</b>	<b>0,69 [0,42; 1,14] 0,15</b>
Übelkeit	9 (12,5)	6 (16,7)	0,75 [0,29; 1,94] 0,55
Obstipation	6 (8,3)	4 (11,1)	0,75 [0,23; 2,49] 0,64
Diarrhö	1 (1,4)	5 (13,9)	0,10 [0,01; 0,82] 0,03
<b>Erkrankungen der Atemwege, des Brustraums und Mediastinums</b>	<b>18 (25,0)</b>	<b>13 (36,1)</b>	<b>0,69 [0,38; 1,25] 0,22</b>
Dyspnoe	4 (5,6)	5 (13,9)	0,40 [0,11; 1,40] 0,15
<b>Infektionen und parasitäre Erkrankungen</b>	<b>15 (20,8)</b>	<b>15 (41,7)</b>	<b>0,50 [0,28; 0,91] 0,02</b>
Nasopharyngitis	2 (2,8)	7 (19,4)	0,14 [0,03; 0,65] 0,01
Untersuchungen	15 (20,8)	3 (8,3)	2,50 [0,77; 8,08] 0,13

MedDRA-Systemorganklasse Preferred Term	Tofersen <sup>1)</sup> N = 72 n (%)	Placebo <sup>1)</sup> N = 36 n (%)	Relatives Risiko [95%-KI] <sup>2)</sup> p-Wert <sup>2)</sup>
<b>Erkrankungen der Haut und des Unterhautzellgewebes</b>	<b>11 (15,3)</b>	<b>3 (8,3)</b>	<b>1,83 [0,55; 6,16] 0,33</b>
<b>Psychiatrische Erkrankungen</b>	<b>9 (12,5)</b>	<b>6 (16,7)</b>	<b>0,75 [0,29; 1,94] 0,55</b>
<b>Stoffwechsel- und Ernährungsstörungen</b>	<b>6 (8,3)</b>	<b>5 (13,9)</b>	<b>0,60 [0,20; 1,8] 0,37</b>

<sup>1)</sup> Unerwünschte Ereignisse wurden vom ersten Tag des Erhalts der Studienmedikation bis zur Follow-Up-Visite Woche 28 bei Teilnahme bzw. Woche 32 bei keiner Teilnahme an der Verlängerungsstudie OLE erhoben.

<sup>2)</sup> Es konnten keine Informationen identifiziert werden, wie das relative Risiko [95%-KI] und der p-Wert berechnet wurde und ob eine Adjustierung (nach Stratifizierungsmerkmalen) erfolgte.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; KI: Konfidenzintervall; UE: Unerwünschtes Ereignis

### Schwere UE

In Tabelle 19 sind die schweren UE (CTCAE-Grad  $\geq 3$ ) mit Inzidenz  $\geq 5\%$ , die bis zum finalen Datenschnitt (16. Juli 2021) in der Studie VALOR aufgetreten sind, dargestellt.

*Tabelle 19: Schwere UE (Grad 3 oder höher), die bei  $\geq 5\%$  der Teilnehmenden in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; Sicherheitspopulation*

MedDRA-Systemorganklasse Preferred Term	Tofersen <sup>1)</sup> N = 72 n (%)	Placebo <sup>1)</sup> N = 36 n (%)	Relatives Risiko [95%-KI] <sup>2)</sup> p-Wert <sup>2)</sup>
<b>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</b>	<b>k. A.</b>	<b>k. A.</b>	<b>4,56 [0,25; 82,5]<sup>3)</sup> 0,30<sup>3)</sup></b>
<b>Erkrankungen der Atemwege, des Brustraums und Mediastinums</b>	<b>4 (5,6)</b>	<b>4 (11,1)</b>	<b>0,50 [0,13; 1,9] 0,31</b>

<sup>1)</sup> Unerwünschte Ereignisse wurden vom ersten Tag des Erhalts der Studienmedikation bis zur Follow-Up-Visite Woche 28 bei Teilnahme bzw. Woche 32 bei keiner Teilnahme an der Verlängerungsstudie OLE erhoben.

<sup>2)</sup> Es konnten keine Informationen identifiziert werden, wie das relative Risiko [95%-KI] und der p-Wert berechnet wurde und ob eine Adjustierung (nach Stratifizierungsmerkmalen) erfolgte.

<sup>3)</sup> Der pU berichtet im Dossier Modul 4 keine Ereigniszahlen. Der angegebene Effektschätzer für den Endpunkt wird nach interner Prüfung als plausibel eingeschätzt.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; k. A.: keine Angabe; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; UE: Unerwünschtes Ereignis

### Schwerwiegende UE

In Tabelle 20 sind die schwerwiegenden UE mit Inzidenz  $\geq 5\%$ , die bis zum finalen Datenschnitt (16. Juli 2021) in der Studie VALOR aufgetreten sind, dargestellt.

**Tabelle 20: Schwerwiegende UE, die bei  $\geq 5$  % der Teilnehmenden in mindestens einem Studienarm aufgetreten sind, nach Systemorganklasse und Preferred Term zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; Sicherheitspopulation**

MedDRA-Systemorganklasse Preferred Term	Tofersen <sup>1)</sup> N = 72 n (%)	Placebo <sup>1)</sup> N = 36 n (%)	Relatives Risiko [95%-KI] <sup>2)</sup> p-Wert <sup>2)</sup>
<b>Erkrankungen der Atemwege, des Brustraums und Mediastinums</b>	<b>5 (6,9)</b>	<b>4 (11,1)</b>	<b>0,63 [0,18; 2,2] 0,46</b>
Dyspnoe	k. A.	k. A.	0,10 [0,01; 2,06] <sup>3)</sup> 0,14 <sup>3)</sup>

<sup>1)</sup> Unerwünschte Ereignisse wurden vom ersten Tag des Erhalts der Studienmedikation bis zur Follow-Up-Visite Woche 28 bei Teilnahme bzw. Woche 32 bei keiner Teilnahme an der Verlängerungsstudie OLE erhoben.

<sup>2)</sup> Es konnten keine Informationen identifiziert werden, wie das relative Risiko [95%-KI] und der p-Wert berechnet wurde und ob eine Adjustierung (nach Stratifizierungsmerkmalen) erfolgte.

<sup>3)</sup> Der pU berichtet im Dossier Modul 4 keine Ereigniszahlen. Der angegebene Effektschätzer für den Endpunkt wird nach interner Prüfung als plausibel eingeschätzt.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; k. A.: keine Angabe; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; UE: Unerwünschtes Ereignis

### UE, die zum Abbruch der Studienmedikation führten

Es traten keine UE, die zum Abbruch der Studienmedikation bei  $\geq 5$  % der Teilnehmenden in mindestens einem Studienarm führten, bis zum finalen Datenschnitt (16. Juli 2021) auf.

### **C-SSRS**

In Tabelle 21 sind die Ergebnisse zum C-SSRS zu suizidalen Gedanken zu Woche 28 (Tag 197) der Studie VALOR beschrieben. Zu keinem Zeitpunkt wurde seitens der Teilnehmenden suizidales Verhalten bzw. selbstverletzendes Verhalten ohne suizidale Absichten beschrieben.

**Tabelle 21: Ergebnisse zum C-SSRS zum Studienende (Datenschnitt: 16. Juli 2021) der Studie VALOR; Sicherheitspopulation**

C-SSRS <sup>1)</sup>	Tofersen N = 72 n (%)	Placebo N = 36 n (%)	Relatives Risiko [95%-KI] <sup>2)</sup> p-Wert <sup>2)</sup>
Suizidalgedanken (Item 1–5) <sup>3)</sup> zum Studienende <sup>4)5)6)</sup>	5 (6,9)	2 (5,6)	1,25 [0,26; 6,13] 0,78

<sup>1)</sup> Der Endpunkt besteht aus 2 Domänen (Suizidgedanken und suizidales Verhalten) mit jeweils 6 Items sowie einer weiteren Domäne zu „Selbstverletzendes Verhalten ohne suizidale Absichten“ mit 1 Item, die in einem dichotomen Antwortformat („ja/nein“) beantwortet werden.

<sup>2)</sup> Es konnten keine Informationen identifiziert werden, wie das relative Risiko [95%-KI] und der p-Wert berechnet wurden und ob eine Adjustierung (nach Stratifizierungsmerkmalen) erfolgte.

<sup>3)</sup> Zu Baseline gaben 9 (12,5 %) Personen im Interventionsarm und 1 (2,8 %) Person im Kontrollarm Suizidgedanken an.

<sup>4)</sup> Im Dossier Modul 4 wird der Zeitraum bis Woche 52 angegeben, obwohl die Erhebung des Endpunktes bis Woche 24 (bzw. Woche 28 bei fehlender Teilnahme an der Verlängerungsstudie) vorgesehen ist. Da die Angaben aus dem Dossier Modul 4 identisch sind mit denen aus dem Studienbericht der Studie VALOR, wird von einer Erhebung bis zum Ende der Studie VALOR ausgegangen.

<sup>5)</sup> Bei fehlenden Werten wird die letzte nicht fehlende Beobachtung für die Zusammenfassung übernommen.

<sup>6)</sup> Während des Studienverlaufs gaben 4 Personen im Interventionsarm (5,6 %) und 1 Person im Kontrollarm mit einem Baseline-Score von 0 suizidale Gedanken an. Bei keiner Person, die zu Baseline Suizidgedanken angab, wurde eine Zunahme im Score (Verschlechterung) berichtet.

Abkürzungen: C-SSRS: Columbia suicide severity rating scale; KI: Konfidenzintervall

## 4 Diskussion der Methodik und Ergebnisse

### 4.1 Zulassungsstatus und Zulassungspopulation von Tofersen

Tofersen (Qalsody) ist zugelassen zur Behandlung von Erwachsenen mit amyotropher Lateralsklerose (ALS), die mit einer Mutation im Superoxid-Dismutase-1-Gen assoziiert ist. Es handelt sich hierbei um einen neuen Wirkstoff, der in dem vorliegenden Anwendungsgebiet unter besonderen Umständen zugelassen wurde. [16]

Gemäß Fachinformation beträgt die zugelassene Dosierung 100 mg Tofersen pro Behandlung. Die Behandlung mit Tofersen sollte mit 3 Aufsättigungsdosen begonnen werden, die im Abstand von jeweils 14 Tagen verabreicht werden. Danach sollte alle 28 Tage eine Erhaltungsdosis verabreicht werden. Der Bedarf für eine Fortsetzung der Therapie sollte in regelmäßigen Abständen überprüft und je nach klinischem Erscheinungsbild des Patienten oder der Patientin und nach seinem Ansprechen auf die Behandlung individuell abgewogen werden. Tofersen ist zur intrathekalen Anwendung mittels Lumbalpunktion bestimmt. [8]

An der Studie VALOR nahmen erwachsene Patientinnen und Patienten mit einer bestätigten SOD1-Gen-Mutation mit einer Schwäche, die auf ALS zurückzuführen ist, teil. Eine Vorbehandlung mit Riluzol war möglich. In der Studie wurden für den Einschluss Kriterien für eine rapide Krankheitsprogression definiert, jedoch war der Einschluss nicht nur auf diese Patientenpopulation limitiert. Das zugelassene Anwendungsgebiet ist vergleichbar mit der Studienpopulation.

Es werden in der Nutzenbewertung Ergebnisse der pivotalen Studie 233AS101 Teil C (VALOR) des finalen Datenschnitts vom 16. Juli 2021 berücksichtigt. Es nahmen aus dem europäischen Raum ca. 25 % und aus dem amerikanischen Raum ca. 67 % der insgesamt 108 in die Studie eingeschlossenen Patientinnen und Patienten an der Studie teil. Eine Übertragbarkeit auf den deutschen Versorgungskontext ist anzunehmen.

### 4.2 Design und Methodik der Studie

#### Zusammenfassung des Studiendesigns

Bei der Studie VALOR handelt es sich um eine doppelblinde, multizentrische, randomisierte, placebo-kontrollierte Studie, bei der Tofersen 100 mg gegenüber Placebo verglichen wird. Die Randomisierung erfolgte 2:1 in die Behandlungsarme Tofersen (N = 72) oder Placebo (N = 36), stratifiziert anhand folgender Kriterien: Erfüllung der prognostischen Kriterien für eine schnelle Krankheitsprogression (eine protokolldefinierte SOD1-Mutation und vor Randomisierung ein Abfall in der ALSFRS-R von  $\geq 0,2$ /Monat oder eine nicht protokolldefinierte SOD1-Mutation und Abfall in der ALSFRS-R vor Randomisierung von  $\geq 0,9$ /Monat) sowie Einnahme von Edaravon oder Riluzol zu Baseline (bei Anwendung von beiden Wirkstoffen wurde unter Edaravon stratifiziert).

Die Dauer der Studie betrug ca. 32–36 Wochen, inklusive einer 4-wöchigen Screening-Phase, einer 24-wöchigen Behandlungsphase (3 Aufsättigungsdosen im Abstand von 2 Wochen, gefolgt von 5 Erhaltungsdosen im Abstand von 4 Wochen) und einer 4- bis 8-wöchigen Follow-up-Phase, in Abhängigkeit davon, ob der Patient oder die Patientin an der Verlängerungsstudie OLE (Studie 233AS102) teilnimmt.

Im Dossier Modul 4 konnten keine Angaben zur medianen Behandlungszeit identifiziert werden. Nach interner Prüfung sind die medianen Behandlungszeiten gleich.

Der primäre Endpunkt war definiert als ALSFRS-R-Gesamtscore zum Studienende.

Es werden Auswertungen anhand der ITT-Population (Intervention: N = 72 vs. Kontrolle: N = 36) für die Wirksamkeits- und anhand der Safety-Population (Intervention: N = 72 vs. Kontrolle: N = 36) für die Sicherheitsdaten dargestellt.

### Studienpopulation

Die Studie VALOR brachen 11 % der Personen im Interventions- und 8 % im Kontrollarm ab. Als häufigster Grund wird in beiden Behandlungsarmen Krankheitsprogression angegeben (Intervention: 4,2 % vs. Kontrolle: 5,6 %). Ein Abbruch der Studienmedikation erfolgte bei 13 % der Personen im Interventions- und 8 % im Kontrollarm.

Im Dossier Modul 4 werden keine Informationen zu den nicht-medikamentösen und zu medikamentösen Begleittherapien gegeben. Nach interner Prüfung ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

Im Dossier Modul 4 finden sich keine Angaben zu den Protokollverletzungen. Eine abschließende Beurteilung ist daher nicht möglich.

Im Interventionsarm waren Personen im Median um ca. 4 Jahre jünger und die mediane Zeit zwischen Symptombeginn und Randomisierung um ca. 3 Monate kürzer gegenüber dem Kontrollarm. Der Anteil an männlichen bzw. weiblichen Personen war im Kontrollarm bei ca. 50 %, während im Interventionsarm 60 % männliche und 40 % weibliche Personen eingeschlossen wurden. Der Anteil an Personen mit einer nicht-hispanischen oder -Latino-Ethnizität war im Interventionsarm um ca. 10 % höher gegenüber dem Kontrollarm, wobei Angaben jeweils von 29 % der Personen im Interventions- und 19 % im Kontrollarm fehlen.

Imbalancen konnten zudem beim mittleren Neurofilament-light-chain-Plasmaspiegel zu Baseline beobachtet werden (Intervention: 100,4 pg/ml) vs. Kontrolle: 89,7 pg/ml). In der Interventionsgruppe betrug die mittlere ALSFRS-Run-In-Steigung (Screening Tag 0 bis Tag 15) -1,0 gegenüber -0,7 in der Kontrollgruppe. Beides kann auf eine schnellere Krankheitsprogression in der Interventionsgruppe hindeuten. Vergleichbare Werte wurden hingegen für die mediane ALSFRS-R-Steigung vor Randomisierung beobachtet.

Für die Stratifizierungsfaktoren (Einsatz von Riluzol oder Edaravon, Eignung prognostischer Kriterien einer schnellen Krankheitsprogression (definierte *SOD1*-Mutation, Steigung des ALSFRS-R-Gesamtscore) war der Anteil an Personen zwischen den Behandlungsarmen vergleichbar. Es liegen nur zusammenfassende, nicht jedoch separate Angaben für die definierte *SOD1*-Mutation vor.

Der pU führt zu den Stratifizierungsfaktoren im Dossier Modul 4 aus, dass der Stratifizierungsfaktor „Erfüllung der Prognose einer Krankheitsprogression“ (definierte *SOD1*-Mutationen und ein Abfall im ALSFRS-R von  $\geq 0,2$  pro Monat oder nicht-protokoll-definierte *SOD1*-Mutation und Abfall im ALSFRS-R vor Randomisierung von  $\geq 0,9$  pro Monat) nicht ausreichend geeignet war, um die Studienarme hinsichtlich einer Krankheitsprogression zu balancieren. Im EPAR wird zudem die Auswahl der Stratifizierungsfaktoren zur Untersuchung prognostischer Kriterien einer schnellen Krankheitsprogression als kritisch angesehen, u. a. aufgrund einer fraglichen Eignung der ALSFRS-R-Steigung vor Randomisierung als prognostischen Faktor zur Beurteilung einer Krankheitsprogression und aufgrund der Heterogenität der *SOD1*-ALS und der damit verbundenen Vielzahl an Mutationstypen [16]. Anzumerken ist auch, dass bei der Studie VALOR, mit geringer Anzahl an teilnehmenden Personen, eine Vielzahl an Faktoren in die Stratifizierung eingingen. Es liegen zudem heterogene Ergebnisse zur Eignung der Bildung eines Gesamtscores vor (siehe Abschnitt 2.3.2).



Die vorliegenden Daten deuten auf Imbalancen in den Baseline-Charakteristika hin. Es liegen jedoch nur bedingt Angaben zu den betroffenen Körperregionen sowie zur Krankheitsschwere (ALS Milano-Torinos Staging und King's staging) zu Baseline vor. Die Baseline-Charakteristika hinsichtlich der Stratifizierungsfaktoren zeigen keine Imbalancen. Eine Verzerrung zugunsten einer der beiden Behandlungsgruppen ist möglich, kann jedoch abschließend nicht beurteilt werden.

Das Verzerrungspotential der Studie VALOR ist unklar.

### **4.3 Mortalität**

In der Studie VALOR trat nur 1 Todesfall in der Interventionsgruppe bis zum finalen Datenschnitt (16. Juli 2021) auf.

Das Verzerrungspotential wird als niedrig angesehen. Für das vorliegende Anwendungsgebiet ALS liegen mit einer vorgesehenen Beobachtungszeit von 28 Wochen keine ausreichend langen vergleichenden Daten zur Erhebung des Endpunktes vor. Die Aussagesicherheit der Ergebnisse ist eingeschränkt.

### **4.4 Morbidität**

#### **Zeit bis zum Tod oder zur dauerhaften Beatmung**

Für den Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“ und den Teilendpunkt „Zeit bis zur dauerhaften Beatmung“ liegt kein statistisch signifikanter Unterschied zwischen Tofersen und Placebo vor.

Es konnten bis zum finalen Datenschnitt (16. Juli 2021) für den Teilendpunkt „Zeit bis zur dauerhaften Beatmung“ (invasive und nicht-invasive Beatmung zusammengefasst) in beiden Behandlungsgruppen nur wenig Ereignisse (dauerhafte Beatmung: Intervention: n = 3 vs. Kontrolle n = 2) beobachtet werden. Bei jeweils 1 Person in den Behandlungsgruppen betrug die Beatmungsdauer nur 1 bzw. 2 Tage und entspricht somit nicht den im Protokoll definierten Vorgaben (dauerhafte Beatmung für mindestens 22 Stunden an konsekutiv mindestens 21 Tagen). Es wird angegeben, dass Patienten und Patientinnen Ereignisse aufweisen konnten, die in der Studie VALOR begannen und in der Verlängerungsstudie OLE fortbestanden. Es konnten jedoch keine eindeutigen Angaben identifiziert werden, ob Ereignisse in der Analyse berücksichtigt wurden, bei denen die Beatmung in der VALOR-Studie begonnen wurde, jedoch die Dauer gemäß Endpunktdefinition erst in der Verlängerungsstudie OLE erreicht wurde, da bei Nicht-Erfüllen der Endpunktdefinition Personen zum letzten Kontakt der Studie VALOR zensiert werden sollten.

Der pU erhebt sowohl die invasive und die nicht-invasive Beatmung und wertet diese in einer Kategorie. Eine differenzierte Auswertung war nicht vorgesehen. Jedoch traten nur wenige Ereignisse auf.

Für den kombinierten Endpunkt „Zeit bis zum Tod oder zur dauerhaften Beatmung“ konnten bis zum finalen Datenschnitt (16. Juli 2021) in beiden Behandlungsarmen insgesamt ebenfalls nur wenig Ereignisse beobachtet werden. Für das vorliegende Anwendungsgebiet ALS liegen mit einer vorgesehenen Beobachtungszeit von 28 Wochen keine ausreichend langen vergleichenden Daten zur Erhebung, insbesondere zum Endpunkt Gesamtüberleben, vor.

Das Verzerrungspotential für die Auswertung ist unklar und die Aussagesicherheit der Ergebnisse ist eingeschränkt.

### **ALSFRS-R, FSS, „allgemeiner Gesundheitszustand“ mittels EQ-5D-5L VAS**

Der pU reicht mit dem Dossier Responderanalysen für eine Verbesserung und Verschlechterung um  $\geq 15\%$  der Skalenspannweite für die Endpunkte ALSFRS-R, FSS, EQ-5D-5L VAS ein. Für die Auswertung anhand der ITT-Population wurden fehlende Werte mittels multipler Imputation ersetzt. Die Plausibilität der Annahme (voraussichtlich Missing At Random) wird nicht diskutiert. Aufgrund fehlender Angaben zu den Annahmen wird davon ausgegangen, dass diese unter der Annahme eines zufälligen Fehlens ersetzt wurden. Inwieweit diese Annahme bei Studienabbrechern (11 % im Interventions- und 8 % im Kontrollarm) angemessen ist, wird nicht argumentiert. Andere Analysen, wie Sensitivitätsanalysen mit Imputation fehlender Werte als Non-Responder, wurden mit dem Dossier nicht eingereicht.

Vom pU wird jeweils der durchschnittliche Anteil an Personen mit entsprechender Response angegeben. Eine Angabe zu den tatsächlichen Ereigniszahlen fehlt und ein Rückschluss ist anhand der durchschnittlichen Anteile nicht möglich. Es ist unklar, wie der durchschnittliche Anteil berechnet wurde, z. B. über Strata oder ggf. über verschiedene Zeitpunkte hinweg, und ob es sich um prozentuale Anteile oder dezimale Anteile handelt. Es ist nicht definiert, ob das Ansprechen tatsächlich zum Zeitpunkt Woche 28 oder ggf. über die Zeit berücksichtigt wurde. Auf eine Darstellung des durchschnittlichen Anteils an Respondern mit einer Verschlechterung oder Verbesserung um  $\geq 15\%$  der Skalenspannweite wird verzichtet. Das Relative Risiko und p-Werte werden ergänzend dargestellt. Allerdings kann die Plausibilität der Ergebnisse aufgrund der fehlenden Ereigniszahlen abschließend nicht beurteilt werden.

Das Verzerrungspotential für die Auswertung ist unklar und die Aussagesicherheit der Ergebnisse ist eingeschränkt.

## **4.5 Lebensqualität**

### **SF-36**

Der pU reicht mit dem Dossier Responderanalysen für eine Verbesserung und Verschlechterung um  $\geq 15\%$  der Skalenspannweite für den Endpunkt SF-36 ein. Für die Auswertung anhand der ITT-Population wurden fehlende Werte mittels multipler Imputation ersetzt. Die Plausibilität der Annahme (voraussichtlich Missing At Random) wird nicht diskutiert. Aufgrund fehlender Angaben zu den Annahmen wird davon ausgegangen, dass diese unter der Annahme eines zufälligen Fehlens ersetzt wurden. Inwieweit diese Annahme bei Studienabbrechern (11 % im Interventions- und 8 % im Kontrollarm) angemessen ist, wird nicht argumentiert. Andere Analysen, wie Sensitivitätsanalysen mit Imputation fehlender Werte als Non-Responder, wurden mit dem Dossier nicht eingereicht.

Vom pU wird jeweils der durchschnittliche Anteil an Personen mit entsprechender Response angegeben. Eine Angabe zu den tatsächlichen Ereigniszahlen fehlt und ein Rückschluss ist anhand der durchschnittlichen Anteile nicht möglich. Es ist unklar, wie der durchschnittliche Anteil berechnet wurde, z. B. über Strata oder ggf. über verschiedene Zeitpunkte hinweg, und ob es sich um prozentuale Anteile oder dezimale Anteile handelt. Es ist nicht definiert, ob das Ansprechen tatsächlich zum Zeitpunkt Woche 28 oder ggf. über die Zeit berücksichtigt wurde. Auf eine Darstellung des durchschnittlichen Anteils an Respondern mit einer Verschlechterung oder Verbesserung um  $\geq 15\%$  der Skalenspannweite wird verzichtet. Das Relative Risiko und p-Werte werden ergänzend dargestellt. Allerdings kann die Plausibilität der Ergebnisse aufgrund der fehlenden Ereigniszahlen abschließend nicht beurteilt werden.

Das Verzerrungspotential für die Auswertung ist unklar und die Aussagesicherheit der Ergebnisse kann nicht abschließend bewertet werden.

## 4.6 Sicherheit

In der vorliegenden Nutzenbewertung werden die unerwünschten Ereignisse ab dem ersten Tag der Studienmedikation bis zum Ende der Studie / vorzeitigen Studienabbruch erhoben. Bei vorzeitigem Abbruch der Studie bzw. keinem Übergang in die Verlängerungsstudie OLE erfolgte eine zusätzliche Visite. Es werden Ergebnisse zur Sicherheit zur Studienvsitede zu Woche 28 (Tag 197) dargestellt. Im Dossier Modul 4 konnten keine Angaben zur medianen Behandlungszeit identifiziert werden. Nach interner Prüfung sind die medianen Behandlungszeiten zwischen den Behandlungsarmen gleich. Es werden die Ergebnisse für die Sicherheitspopulation dargestellt.

Angaben, inwiefern Auswertungen unter Nichtberücksichtigung von erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung durchgeführt wurden, konnten nicht identifiziert werden.

Bei den Gesamtraten liegt kein statistisch signifikanter Unterschied zwischen Tofersen und Placebo vor.

Beim PT „Diarrhö“ (RR: 0,10 [95%-KI: 0,01; 0,82], beim PT „Nasopharyngitis“ (RR: 0,14 [95%-KI: 0,03; 0,65] und SOC „Infektionen und parasitäre Erkrankungen“ (RR: 0,50 [95%-KI: 0,28; 0,91] liegt ein statistisch signifikanter Unterschied zugunsten von Tofersen gegenüber Placebo vor.

Für die Gesamtrate UE, die zum Abbruch der Studienmedikation führten, UE CTCAE-Grad  $\geq 3$  SOC „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“, SUE „Dyspnoe“ werden im Dossier Modul 4 keine Ereigniszahlen berichtet, jedoch werden die angegebenen Effektschätzer nach interner Prüfung als plausibel eingeschätzt.

### C-SSRS

Für die Subdomäne „Suizidgedanken“ liegt kein statistisch signifikanter Unterschied zwischen Tofersen und Placebo vor.

Zu keinem Zeitpunkt wurde seitens der Teilnehmenden suizidales Verhalten bzw. selbstverletzendes Verhalten ohne suizidale Absichten beschrieben.

Das Verzerrungspotential kann nicht abschließend beurteilt werden. Es ist unklar, wie die Effektschätzer und p-Werte berechnet wurden und inwiefern eine adjustierte Auswertung nach den Stratifizierungsfaktoren vorgenommen wurde.

Aussagen zu langfristigen Effekten auf die Sicherheit lassen sich anhand der Studie nicht ableiten.

## **5 Anforderung an eine qualitätsgesicherte Anwendung**

Die Angaben der Fachinformation sind zu berücksichtigen.

## 6 Zusammenfassung der Nutzenbewertung

Tofersen ist zugelassen für die Behandlung von Erwachsenen mit amyotropher Lateralsklerose (ALS), die mit einer Mutation im Superoxid-Dismutase-1-Gen assoziiert ist. Die Nutzenbewertung von Tofersen basiert auf der zulassungsbegründenden Studie VALOR (Studie 233AS101 Teil C). Es handelt sich bei der Studie 233AS101 Teil C (VALOR) um eine randomisierte, doppelblinde, placebo-kontrollierte Studie bei *SOD1*-assoziierter ALS. Die Randomisierung erfolgte 2:1 in die Behandlungsarme Tofersen (N = 72) oder Placebo (N = 36). Die Studie umfasst eine 4-wöchige Screening-, eine 24-wöchige Behandlungs- (Erhalt von 3 Aufsächtigsdosen im Abstand von 2 Wochen und 5 Erhaltungsdosen im Abstand von 4 Wochen) und eine 4–8-wöchige Nachbeobachtungsphase. Patientinnen und Patienten konnten nach Abschluss der Studie VALOR in die Verlängerungsstudie 233AS102 (OLE) wechseln. Die Studie VALOR ist abgeschlossen und es liegt der finale Datenschnitt vom 16. Juli 2021 vor.

Die Ergebnisse der in der Nutzenbewertung berücksichtigten Endpunkte der Studie werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Tofersen ein statistisch signifikant positiver Effekt bei hohem oder unklarem ( $\uparrow$ ) oder niedrigem ( $\uparrow\uparrow$ ) Verzerrungspotential bzw. ein statistisch signifikant negativer Effekt bei hohem oder unklarem ( $\downarrow$ ) oder niedrigem ( $\downarrow\downarrow$ ) Verzerrungspotential bzw. kein Unterschied ( $\leftrightarrow$ ) gezeigt werden konnte.

Weitergehende, über die Angaben in der nachfolgenden zusammenfassenden Tabelle hinausgehende Erläuterungen zur Methodik und den Ergebnissen der Endpunkte finden sich in den jeweiligen Kapiteln 2.3 bis 2.5 und 3.2 bis 3.4 der Nutzenbewertung.

Tabelle 22: Zusammenfassende Darstellung der Ergebnisse zu Woche 28 (Tag 197) der Studie VALOR; ITT-Population (Datenschnitt: 16. Juli 2021)

Studie VALOR Darstellung der Ergebnisse	Tofersen 100 mg N = 72 <sup>1)</sup>	Placebo N = 36 <sup>1)</sup>	Tofersen 100 mg vs. Placebo	Effekt
<b>Mortalität</b>				
	<i>Personen mit Ereignis, n (%)</i>	<i>Personen mit Ereignis, n (%)</i>	<i>Hazard Ratio [95%-KI]<sup>2)</sup> p-Wert<sup>2)</sup></i>	
<b>Gesamtüberleben</b>				
Gesamtüberleben <sup>3)</sup>	1 (1,4)	0 (0,0)	<i>nicht berechenbar<sup>4)</sup></i>	$\leftrightarrow$
<b>Mortalität/Morbidität</b>				
	<i>Personen mit Ereignis, n (%)</i>	<i>Personen mit Ereignis, n (%)</i>	<i>Hazard Ratio [95%-KI]<sup>2)</sup> p-Wert<sup>2)</sup></i>	
Zeit bis zum Tod oder zur dauerhaften Beatmung <sup>5)</sup>				
Zeit bis zum Tod oder zur dauerhaften Beatmung <sup>3)</sup>	4 (5,6)	2 (5,6)	0,97 [0,17; 5,71]; 0,98	$\leftrightarrow$
Zeit bis zur dauerhaften Beatmung <sup>3)</sup>	3 (4,2)	2 (5,6)	0,82 [0,13; 5,34]; 0,83	$\leftrightarrow$

Studie VALOR Darstellung der Ergebnisse	Tofersen 100 mg N = 72 <sup>1)</sup>	Placebo N = 36 <sup>1)</sup>	Tofersen 100 mg vs. Placebo	Effekt
<b>Sicherheit<sup>6)</sup></b>				
	<i>Personen mit Ereignis, n (%)</i>	<i>Personen mit Ereignis, n (%)</i>	<i>Relatives Risiko [95%- KI]<sup>7)</sup>; p-Wert<sup>7)</sup></i>	
<b>Unerwünschte Ereignisse</b>				
UE CTCAE-Grad $\geq$ 3	12 (16,7)	4 (11,1)	1,50 [0,520; 4,32] 0,453	$\leftrightarrow$
SUE	13 (18,1)	5 (13,9)	1,30 [0,50; 3,36] 0,59	$\leftrightarrow$
UE, das zum Abbruch der Studienmedikation führte	k. A.	k. A.	4,56 [0,25; 82,48] <sup>8)</sup> 0,30 <sup>8)</sup>	$\leftrightarrow$
<b>C-SSRS<sup>9)</sup>10)</b>				
Suizidgedanken (Item 1-5) <sup>11)</sup>	5 (6,9) <sup>12)13)</sup>	2 (5,6) <sup>12)13)</sup>	1,25 [0,26; 6,13] 0,78	$\leftrightarrow$

<sup>1)</sup> Die Anzahl entspricht denjenigen Personen, die für die Berechnung der jeweiligen Maßzahlen herangezogen wurden.

<sup>2)</sup> Cox-Proportional-Hazard-Modell adjustiert für die Krankheitsdauer seit Symptombeginn, den ALSFRS-R-Gesamtscore zu Beginn der Studie und die Verwendung von Riluzol oder Edaravon.

<sup>3)</sup> Aufgrund der geringen Anzahl an Ereignissen konnte die mediane Zeit bis zum Ereignis nicht bestimmt werden.

<sup>4)</sup> Aufgrund der geringen Ereigniszahlen wurde kein Effektschätzer berechnet.

<sup>5)</sup> Endpunkt definiert als Tod oder dauerhafte Beatmung ( $\geq$  22 Stunden/Tag invasiv oder nicht-invasiv durchgehend für  $\geq$  21 Tage) (je nachdem, was als erstes auftritt). Es werden nur Ereignisse berücksichtigt, die vom Endpunkt-Beurteilungsausschuss beurteilt wurden. Für jeweils 1 Person in beiden Behandlungsarmen entspricht die Dauer nicht den Kriterien einer dauerhaften Beatmung (mindestens 21 Tage). Patientinnen und Patienten, die die Endpunktdefinition nicht erfüllen, werden zum letzten Kontakt in der Studie VALOR zensiert.

<sup>6)</sup> Sicherheitspopulation entspricht der ITT-Population

<sup>7)</sup> Es liegen keine Angaben vor, wie das Relative Risiko (95%-KI) und der p-Wert berechnet wurden und inwiefern eine Adjustierung nach u. a. Stratifizierungsfaktoren vorgenommen wurde.

<sup>8)</sup> Der pU berichtet im Dossier Modul 4 keine Ereigniszahlen. Der angegebene Effektschätzer für den Endpunkt wird nach interner Prüfung als adäquat eingeschätzt.

<sup>9)</sup> Der Endpunkt besteht aus 3 Domänen (Suizidgedanken, suizidales Verhalten, selbstverletzendes Verhalten ohne Suizidabsichten), die in einem dichotomen Antwortformat („ja/nein“) beantwortet werden.

<sup>10)</sup> Zu keinem Zeitpunkt wurde seitens der Teilnehmenden suizidales Verhalten bzw. selbstverletzendes Verhalten ohne suizidale Absichten beschrieben. Auf eine Darstellung wird daher verzichtet.

<sup>11)</sup> Zu Baseline gaben 9 (12,5 %) Personen im Interventions- und 1 Person (2,8 %) im Kontrollarm Suizidgedanken an. Bei keiner Person, die zu Baseline Suizidgedanken angab, wurde eine Zunahme im Score (Verschlechterung) berichtet.

<sup>12)</sup> Anzahl an Personen, die bei einer Visite post-Baseline mit einem „Ja“ auf eine der Fragen geantwortet haben.

<sup>13)</sup> Von den Teilnehmenden, die zu Baseline keine Suizidgedanken aufwiesen (Score = 0), gaben während des Studienverlaufs 4 (5,6 %) Personen im Interventions- und 1 (2,8 %) Person im Kontrollarm Suizidgedanken an.

Abkürzungen: ALSFRS-R: Amyotrophic Lateral Sclerosis Functional Rating Scale – Revised; C-SSRS: Columbia Suicide Severity Rating Scale; CTCAE: Common Terminology Criteria for Adverse Events; ITT: Intention-to-treat; k. A.: keine Angabe; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis

## Referenzen

1. The Amyotrophic Lateral Sclerosis Functional Rating Scale. Assessment of activities of daily living in patients with amyotrophic lateral sclerosis: the ALS CNTF treatment study (ACTS) phase I-II Study Group. *Arch Neurol* 1996;53(2):141-147.
2. **Bakers JNE, de Jongh AD, Bunte TM, Kendall L, Han SS, Epstein N, et al.** Using the ALSFRS-R in multicentre clinical trials for amyotrophic lateral sclerosis: potential limitations in current standard operating procedures. *Amyotroph Lateral Scler Frontotemporal Degener* 2022;23(7-8):500-507.
3. **Bakker LA, Schröder CD, van Es MA, Westers P, Visser-Meily JMA, van den Berg LH.** Assessment of the factorial validity and reliability of the ALSFRS-R: a revision of its measurement model. *J Neurol* 2017;264(7):1413-1420.
4. **Bedlack RS, Vaughan T, Wicks P, Heywood J, Sinani E, Selsov R, et al.** How common are ALS plateaus and reversals? *Neurology* 2016;86(9):808-812.
5. **Biogen.** Centralized spirometry procedure manual for Cardio Pulmonary Suite (CPS) software: Biogen 233AS101; version 1.0 [unveröffentlicht]. 29.10.2015.
6. **Biogen.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Anhang 4-G zu Modul 4A: Tofersen; Behandlung von Erwachsenen mit amyotropher Lateralsklerose (ALS), die mit einer Mutation im Superoxid-Dismutase 1 (SOD1)-Gen assoziiert ist; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 01.07.2024.
7. **Biogen.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4A: Tofersen; Behandlung von Erwachsenen mit amyotropher Lateralsklerose (ALS), die mit einer Mutation im Superoxid-Dismutase 1 (SOD1)-Gen assoziiert ist; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 01.07.2024.
8. **Biogen.** Qalsody 100 mg Injektionslösung [online]. 05.2024. Berlin. [Zugriff: 17.07.2024]. URL: <http://www.fachinfo.de>.
9. **Biogen.** A study to evaluate the efficacy, safety, tolerability, pharmacokinetics, and pharmacodynamics of BIIB067 administered to adult subjects with amyotrophic lateral sclerosis and confirmed superoxide dismutase 1 mutation; part C; tofersen (233AS101); clinical study protocol; version 8 [unveröffentlicht]. 15.06.2021.
10. **Biogen.** A study to evaluate the efficacy, safety, tolerability, pharmacokinetics, and pharmacodynamics of BIIB067 administered to adult subjects with amyotrophic lateral sclerosis and confirmed superoxide dismutase 1 mutation; part C; tofersen (233AS101); clinical study report [unveröffentlicht]. 12.04.2022.
11. **Biogen.** A study to evaluate the efficacy, safety, tolerability, pharmacokinetics, and pharmacodynamics of BIIB067 administered to adult subjects with amyotrophic lateral sclerosis and confirmed superoxide dismutase 1 mutation; part C; tofersen (233AS101); clinical study report; section 14.2\_14.4 [unveröffentlicht]. 2022.
12. **Biogen.** A study to evaluate the efficacy, safety, tolerability, pharmacokinetics, and pharmacodynamics of BIIB067 administered to adult subjects with amyotrophic lateral sclerosis and confirmed superoxide dismutase 1 mutation; part C; tofersen (233AS101); statistical analysis plan; version 2 [unveröffentlicht]. 14.08.2021.

13. **Cedarbaum JM, Stambler N, Malta E, Fuller C, Hilt D, Thurmond B, et al.** The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function: BDNF ALS Study Group (Phase III). *J Neurol Sci* 1999;169(1-2):13-21.
14. **Deutsche Gesellschaft für Neurologie.** Motoneuronerkrankungen; S1-Leitlinie, Langfassung, Version 9 [online]. AWMF-Registernummer: 030-001. Berlin (GER): Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF); 2021. [Zugriff: 30.07.2024]. URL: [https://register.awmf.org/assets/guidelines/030-001l\\_S1\\_Motoneuronerkrankungen\\_2021-11.pdf](https://register.awmf.org/assets/guidelines/030-001l_S1_Motoneuronerkrankungen_2021-11.pdf).
15. **European Medicines Agency (EMA).** Guideline on clinical investigation of medicinal products for the treatment of amyotrophic lateral sclerosis (ALS) [online]. 19.11.2015. Amsterdam (NED): EMA. [Zugriff: 28.08.2024]. URL: [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-amyotrophic-lateral-sclerosis\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-amyotrophic-lateral-sclerosis_en.pdf).
16. **European Medicines Agency (EMA).** Qalsody: European public assessment report EMEA/H/C/005493/0000 [online]. 22.02.2024. Amsterdam (NED): EMA. [Zugriff: 17.07.2024]. URL: [https://www.ema.europa.eu/en/documents/assessment-report/qalsody-epar-public-assessment-report\\_en.pdf](https://www.ema.europa.eu/en/documents/assessment-report/qalsody-epar-public-assessment-report_en.pdf).
17. **Franchignoni F, Mandrioli J, Giordano A, Ferro S, Group E.** A further Rasch study confirms that ALSFRS-R does not conform to fundamental measurement requirements. *Amyotroph Lateral Scler Frontotemporal Degener* 2015;16(5-6):331-337.
18. **Franchignoni F, Mora G, Giordano A, Volanti P, Chiò A.** Evidence of multidimensionality in the ALSFRS-R Scale: a critical appraisal on its measurement properties using Rasch analysis. *J Neurol Neurosurg Psychiatry* 2013;84(12):1340-1345.
19. **Jackson C, De Carvalho M, Genge A, Heiman-Patterson T, Shefner JM, Wei J, et al.** Relationships between slow vital capacity and measures of respiratory function on the ALSFRS-R. *Amyotroph Lateral Scler Frontotemporal Degener* 2018;19(7-8):506-512.
20. **Jenkinson C, Fitzpatrick R.** Reduced item set for the amyotrophic lateral sclerosis assessment questionnaire: development and validation of the ALSAQ-5. *J Neurol Neurosurg Psychiatry* 2001;70(1):70-73.
21. **Jenkinson C, Fitzpatrick R, Brennan C, Bromberg M, Swash M.** Development and validation of a short measure of health status for individuals with amyotrophic lateral sclerosis/motor neurone disease: the ALSAQ-40. *J Neurol* 1999;246 Suppl 3:l116-21.
22. **Jenkinson C, Fitzpatrick R, Swash M, Jones G.** Comparison of the 40-item Amyotrophic Lateral Sclerosis Assessment Questionnaire (ALSAQ-40) with a short-form five-item version (ALSAQ-5) in a longitudinal survey. *Clin Rehabil* 2007;21(3):266-272.
23. **Jenkinson C, Norquist JM, Fitzpatrick R.** Deriving summary indices of health status from the Amyotrophic Lateral Sclerosis Assessment Questionnaires (ALSAQ-40 and ALSAQ-5). *J Neurol Neurosurg Psychiatry* 2003;74(2):242-245.
24. **Kaufmann P, Levy G, Montes J, Buchsbaum R, Barsdorf AI, Battista V, et al.** Excellent inter-rater, intra-rater, and telephone-administered reliability of the ALSFRS-R in a multicenter clinical trial. *Amyotroph Lateral Scler* 2007;8(1):42-46.
25. **Maier A, Boentert M, Reilich P, Witzel S, Petri S, Grosskreutz J, et al.** ALSFRS-R-SE: an adapted, annotated, and self-explanatory version of the revised amyotrophic lateral sclerosis functional rating scale. *Neurol Res Pract* 2022;4(1):60.



26. **National Institute for Health and Care Excellence (NICE).** Motor neurone disease: assessment and management [online]. Last updated 23.07.2019. London (GBR): NICE; 2016. [Zugriff: 29.08.2024]. (NICE guideline; NG42). URL: <https://www.nice.org.uk/guidance/ng42/resources/motor-neurone-disease-assessment-and-management-pdf-1837449470149>.
27. **Pickard AS, Neary MP, Cella D.** Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health Qual Life Outcomes* 2007;5:70.
28. **Traynor BJ, Zhang H, Shefner JM, Schoenfeld D, Cudkowicz ME, Neals Consortium.** Functional outcome measures as clinical trial endpoints in ALS. *Neurology* 2004;63(10):1933-1935.
29. **Van Eijk RPA, de Jongh AD, Nikolakopoulos S, McDermott CJ, Eijkemans MJC, Roes KCB, et al.** An old friend who has overstayed their welcome: the ALSFRS-R total score as primary endpoint for ALS clinical trials. *Amyotroph Lateral Scler Frontotemporal Degener* 2021;22(3-4):300-307.