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. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO

Wirkstoff: Luspatercept (beta-Thalassämie)

Datum der Veröffentlichung: 2. November 2020

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 Aufbau und Design der eingeschlossenen Studie 10
2.3 Endpunkte 17
2.3.1 Mortalität 18
2.3.2 Morbidität 19
2.3.3 Lebensqualität 26
2.3.4 Sicherheit 30
2.3.5 Übersicht der Erhebungszeitpunkte 31
2.4 Statistische Methoden 32
2.5 Verzerrungspotential auf Studien- und Endpunktebene 34
3 Ergebnisse der eingeschlossenen Studie 36
3.1 Studiencharakteristika und Studienmedikation 36
3.2 Mortalität 42
3.3 Morbidität 42
3.4 Lebensqualität 43
3.5 Sicherheit 45
4 Diskussion der Methodik und Ergebnisse 51
4.1 Zulassungsstatus und Zulassungspopulation von Luspatercept 51
4.2 Design und Methodik der Studie 51
4.3 Mortalität 54
4.4 Morbidität 55
4.5 Lebensqualität 56
4.6 Sicherheit 57
5 Anforderung an eine qualitätsgesicherte Anwendung 59
6 Zusammenfassung der Nutzenbewertung 60
Referenzen 62

Tabellenverzeichnis

Tabelle 1: Übersicht über die Studienbasis 9
Tabelle 2: Charakterisierung der Studie ACE-536-B-THAL-001 (BELIEVE) 10
Tabelle 3: Charakterisierung der Intervention Luspatercept während der (Langzeit-)Behandlungsphase der Studie ACE-536-B-THAL-001 (BELIEVE) 16
Tabelle 4: Zusammenfassung der Endpunktbewertung 18
Tabelle 5: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie ACE-536-B-THAL-001 (BELIEVE) 31
Tabelle 6: Verzerrungspotential der Studie ACE-536-B-THAL-001 (BELIEVE) 34
Tabelle 7: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie ACE-536-B-THAL-001 (BELIEVE) für den Datenschnitt vom 11.05.2018 35
Tabelle 8: Allgemeine Angaben zur Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt: 11.05.2018) 36
Tabelle 9: Charakterisierung der Studienpopulation ACE-536-B-THAL-001 (BELIEVE) zu Baseline (Datenschnitt 11.05.2018) 38
Tabelle 10: Angaben zur Exposition mit der Studienmedikation innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt: 11.05.2018) 39
Tabelle 11: Angaben zur Therapie mit Erythrozytenkonzentraten und Eisenchelatoren innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt: 11.05.2018) 40
Tabelle 12: Begleitmedikation mit einer Häufigkeit von ≥ 10 % in der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 41
Tabelle 13: Gesamtüberleben in der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 42
Tabelle 14: Anzahl Personen mit Hospitalisierungen während der verblindeten Behandlungs- und Langzeitbehandlungsphase in der ITT-Population der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 42
Tabelle 15: Rücklaufquote TranQoL bezogen auf die ITT-Population der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 43
Tabelle 16: Mittlere Veränderung des TranQoL in der HRQoL evaluable Population der Studie ACE-536-B-THAL-001 (BELIEVE) zu Woche 48 (Datenschnitt: 11.05.2018) 44
Tabelle 17: Rücklaufquote des SF-36 bezogen auf die ITT-Population der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 44
Tabelle 18: Responderanalysen für Verbesserungen des PCS und MCS des SF-36 in der HRQoL evaluable Population der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 45

Tabelle 19: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der verblindeten Behandlungs- und Langzeitbehandlungsphase innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 46

Tabelle 20: UE mit Inzidenz ≥ 10 % bzw. ≥ 10 Ereignisse und einer Inzidenz von ≥ 1 % nach MedDRA-Systemorganklasse und Preferred Term innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 47

Tabelle 21: SUE mit Inzidenz ≥ 5 % nach MedDRA-Systemorganklasse und Preferred Term innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 49

Tabelle 22: UE des NCI-CTCAE-Grades ≥ 3 mit Inzidenz ≥ 5 % nach MedDRA-Systemorganklasse und Preferred Term innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 49

Tabelle 23: Anzahl an Patientinnen und Patienten mit UE von besonderem Interesse und weiteren relevanten Sicherheitsereignissen innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018) 50

Tabelle 24: Zusammenfassende Darstellung der Ergebnisse der Studie ACE-536-B-THAL-001 (BELIEVE) 60

Abkürzungsverzeichnis

ADA	Antidrug antibody
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
BSC	Best Supportive Care
CMH	Cochrane-Mantel-Haenszel
DXA	Dual-Röntgen-Absorptiometrie
ECOG	Eastern Cooperative Oncology Group
EK	Erythrozytenkonzentrat
EMA	European Medicines Agency
EPAR	European Public Assessment Report
FDA	Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
GDF11	Growth differentiation factor 11
GKV	Gesetzliche Krankenversicherung
Hb	Hämoglobin
HRQoL	Health-related Quality of Life (gesundheitsbezogenene Lebensqualität)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-To-Treat
KI	Konfidenzintervall
LS	Least-Square
MCS	Mental Component Summary (psychischer Summenscore)
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal Important Difference
MMRM	Mixed-model repeated Measure
MRT	Magnetresonanztomographie
MW	Mittelwert
N/n	Anzahl
NCI-CTCAE	National Cancer Institute – Common Terminology Criteria for Adverse Events
OR	Odds Ratio
PCS	Physical Component Summary (körperlicher Summenscore)
PT	Preferred Term/s
pU	Pharmazeutischer Unternehmer
RCT	Randomized controlled trial/s (randomisierte kontrollierte Studie/n)
RD	Risikodifferenz
RR	Relatives Risiko im Sinne eines Risk Ratios
SAP	Statistischer Analyseplan

SF-36	Short Form 36
SGB	Sozialgesetzbuch
SMQ	Standardised MedDRA Query
SOC	Systemorganklasse/n
SUE	Schwerwiegende/s unerwünschte/s Ereignis/se
TranQoL	Transfusion-dependent Quality of Life Questionaire
UE	Unerwünschte/s Ereignis/se
ULN	Upper Limit of Normal (obere Grenze des Normalbereiches)
VerfO	Verfahrensordnung

Hintergrund

Luspatercept 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.

Dementsprechend hat der G-BA in seiner Sitzung am 15. März 2012 den mit Beschluss vom 1. August 2011 erteilten Auftrag an das IQWiG zur Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen gemäß § 35a Absatz 2 SGB V in der Weise abgeändert, dass bei Orphan Drugs eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann erfolgt, wenn der Umsatz des betreffenden Arzneimittels die gesetzliche Grenze von 50 Millionen Euro überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt (vgl. § 35a Absatz 1 Satz 12 SGB V).

Der G-BA bestimmt gemäß 5. Kapitel § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des G-BA (VerfO) das Ausmaß des Zusatznutzens für die Anzahl der Patientinnen und Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Luspatercept 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 heran und bewertet die darin enthaltenen Daten nach Maßgabe der im 5. Kapitel § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Luspatercept in seiner Sitzung am 27. Oktober 2020 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 29. Juli 2020 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pharmazeutischen Unternehmers in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

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

Die Nutzenbewertung wird am 02. November 2020 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

Luspatercept (Reblozyl®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Abs. 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 50 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. In die Bewertung werden gemäß Zulassungsstatus Patientinnen und Patienten in folgender Indikation eingeschlossen:

- für die Behandlung von erwachsenen Patientinnen und Patienten mit transfusionsabhängiger Anämie, die mit einer β-Thalassämie verbunden ist.

Folgende Hinweise zur Dosierung sind der Fachinformation zu entnehmen:

- Vor jeder Reblozyl-Anwendung muss der Hämoglobinspiegel (Hb-Spiegel) der Patientinnen und Patienten bestimmt werden. Wenn vor der Dosisgabe eine Transfusion mit Erythrozytenkonzentrat (EK) erfolgt, ist der Hb-Spiegel vor der Transfusion als Grundlage für die Dosierung heranzuziehen.
- Die empfohlene Anfangsdosis von Reblozyl ist 1,0 mg/kg, verabreicht einmal alle 3 Wochen. Bei Patientinnen und Patienten, die nach ≥ 2 aufeinanderfolgenden Dosen (6 Wochen) der Anfangsdosis 1,0 mg/kg kein Ansprechen erreichen, wobei Ansprechen als Reduktion der EK-Transfusionslast um mindestens ein Drittel definiert ist, ist die Dosis auf 1,25 mg/kg zu erhöhen. Die Dosis darf nicht über die maximale Dosis von 1,25 mg/kg einmal alle 3 Wochen hinaus erhöht werden. Wenn eine Person kein Ansprechen mehr zeigt (wenn die EK-Transfusionslast nach einem ersten Ansprechen wieder ansteigt), ist die Dosis um eine Dosisstufe zu erhöhen.

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 Ableitung des Zusatznutzens erachtet	Studie relevant für Nutzenbewertung	Gründe, warum Studie nicht relevant für Nutzenbewertung
Studien zum Wirkstoff				
ACE-536-B-THAL-001 / BELIEVE (NCT0260443) [1)]	Ja	Ja	Ja	
A536-04 (NCT01749540)	Ja[2)]	Nein	Nein	• Offene einarmige Phase-II-/Dosisfindungsstudie mit Dosierungen von 0,2; 0,4; 0,6; 0,8; 1,0; 1,25 und 1,5 mg/kg Luspatercept subkutan alle 3 Wochen bei Patientinnen und Patienten mit β-Thalassämie intermedia oder major, Beobachtungszeit: 3 Monate, Studie ist abgeschlossen (November 2015), N = 64 • nicht vergleichende Studie bei Vorhandensein von RCT-Daten (geringere Aussagesicherheit) • Studienpopulation teilweise abweichend von Zulassung (Einschluss von Patientinnen und Patienten ohne Transfusionsabhängigkeit, n = 33)
A536-06 (NCT02268409)	Ja[3)]	Nein	Nein	• Extensionsstudie der Studie A536-04, Beobachtung geplant für 60 Monate, Studie ist abgeschlossen (18. Juni 2020)[4)], N = 51 • Vorliegender Studienbericht beschreibt Interims-Datenschnitt vom 31.08.2017 (Mediane Behandlungszeit (min, max) = 715 (42, 1042) Tage • nicht vergleichende Studie bei Vorhandensein von RCT-Daten (geringere Aussagesicherheit) • Studienpopulation teilweise abweichend von Zulassung (Einschluss von Patientinnen und Patienten ohne Transfusionsabhängigkeit, n = 27)

[1)] Zulassungsrelevante Studie gemäß EPAR
[2)] Es liegt der finale Studienbericht vom 29.08.2018 vor [1]. Studienprotokoll und SAP wurden nicht vorgelegt.
[3)] Es liegt ein Interims-Studienbericht vom 29.08.2018 mit Datenschnitt vom 31.08.2017 vor [2]. Studienprotokoll und SAP wurden nicht vorgelegt.
[4)] Angaben gemäß www.clinicaltrials.gov

Abkürzungen: EPAR: European Public Assessment Report; FI: Fachinformation; pU: pharmazeutischer Unternehmer

Zur Nutzenbewertung für Luspatercept wurden folgende Studien und Daten herangezogen:

- Herstellerdossier zu Luspatercept [4]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [9]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan der Studie ACE-536-B-THAL-001 / BELIEVE (NCT0260443) [5]
- Bericht zu präspezifizierten Analysen der Lebensqualität [6]
- Ergebnisse aus Post-hoc-Analysen für das Dossier für die Datenschnitte vom 11. Mai 2018 [7,8]

2.2 Aufbau und Design der eingeschlossenen Studie

Die Unterlagen zur Nutzenbewertung für Luspatercept basieren auf der Zulassungsstudie ACE-536-B-THAL-001 / BELIEVE (NCT0260443). Die Studie und die Intervention werden in den Tabellen 2–3 charakterisiert.

Tabelle 2: Charakterisierung der Studie ACE-536-B-THAL-001 (BELIEVE)

Charakteris-tikum	Beschreibung
Design	Es handelt sich um eine multizentrische, doppelblinde, randomisierte, placebokontrollierte Phase-III-Studie zur Untersuchung von Wirksamkeit und Sicherheit von Luspatercept bei Patientinnen und Patienten mit transfusionsabhängiger β-Thalassämie. Die Studie beinhaltete eine 12-wöchige Screening-Phase, eine 48-wöchige Behandlungsphase, eine Langzeit-Behandlungsphase, eine offene-Extensionsphase und eine Post-Treatment-Follow-up-Phase. Nach Unterzeichnung der Einverständniserklärung begann die 12-wöchige Screeningphase, in der die Ein- bzw. Ausschlusskriterien überprüft wurden und der Bedarf an Bluttransfusionen monitoriert wurde. Bei Erfüllung der Ein- und Ausschlusskriterien erfolgte die Randomisierung. Randomisiert wurde im Verhältnis 2:1 stratifiziert nach Region (Nordamerika und Europa, Mittlerer Osten und Nordafrika, Asien-Pazifik) auf die beiden Studienarme: Luspatercept + BSC vs. Placebo + BSC. Die Behandlungsphase begann mit der Verabreichung der ersten Dosis und dauerte 48 Wochen. Eine verblindete Fortsetzung der Behandlung war nach Beendigung der 48-wöchigen Behandlungsphase innerhalb der Langzeit-Behandlungsphase im Ermessen des ärztlichen Prüfpersonals solange möglich, bis der letzte Patient bzw. die letzte Patientin der Studie die 48-wöchige Behandlungsphase abgeschlossen hat oder bis zur Entblindung. Die Langzeit-Behandlungsphase endet daher entweder, nachdem alle Patientinnen und Patienten die 48-wöchige Behandlungsphase abgeschlossen oder diese vorzeitig beendet hatten oder mit der Entscheidung zur Entblindung der Studie durch das Data Monitoring Komitee. Nach Entblindung erfolgte die Beurteilung der Sicherheit und Wirksamkeit auf Basis der primären Analyse durch das Data Monitoring Komitee. Aufgrund der positiven Beurteilung war eine Weiterbehandlung mit Luspatercept sowie ein Therapiebeginn für Studienteilnehmende aus dem Placebo-Arm in der offenen Extensionsphase möglich. Eine Nachbeobachtung der Studienteilnehmenden in der offenen Extensionsphase ist für insgesamt 5 Jahre nach der ersten Dosis Luspatercept innerhalb der Behandlungsphase geplant. Studienteilnehmende beider Behandlungsarme, welche die Therapie abbrachen, sollen für 3 Jahre (156 Wochen) nach Abbruch der Therapie innerhalb der Post-Treatment-Follow-up-Phase nachbeobachtet werden.

Charakteris-tikum	Beschreibung
Population	**Wesentliche Einschlusskriterien** • Alter ≥ 18 Jahre • Dokumentierte Diagnose einer β-Thalassämie oder Hämoglobin E / β-Thalassämie (β-Thalassämie mit einer Mutation oder Multiplikation des α-Globin waren erlaubt) • Regelmäßige Transfusionen, definiert als 6–20 EK-Einheiten innerhalb der 24 Wochen vor der Randomisierung und keine transfusionsfreie Periode von ≥ 35 Tagen während dieser Zeit • ECOG-Status von 0 oder 1 **Wesentliche Ausschlusskriterien** • Diagnose einer Hämoglobin S / β-Thalassämie oder α-Thalassämie (z. B. Hämoglobin H) • Aktive Hepatitis-C-Virus-(HCV)-Infektion, nachgewiesen durch positiven HCV-RNA-Test mit ausreichender Sensitivität, oder aktive infektiöse HBV-Infektion, nachgewiesen durch HBV-Oberflächenantigene (HBsAg) und/oder positive HBV-DNA, oder bekannte positive HIV-Infektion (Patientinnen und Patienten, die antiviral behandelt werden, sollten zwei negative HCV-RNA-Tests mit 3 Monaten Abstand voneinander aufweisen) • Tiefe Venenthrombose oder Schlaganfall mit erforderlicher medizinischer Intervention innerhalb der 24 Wochen vor Randomisierung • Chronische Behandlung mit Antikoagulantien (außer die Behandlung wurde ≤ 28 Tage vor Randomisierung beendet oder die Antikoagulantien wurden prophylaktisch vor Operationen oder Hochrisikoeingriffen gegeben oder waren niedrig molekulares Heparin zur Behandlung von oberflächlichen Venenthrombosen oder dauerhafte Therapie mit Acetylsalicylsäure) • Thrombozytenzahl > 1.000 x 10^9/l • Schlecht kontrollierter Diabetes mellitus innerhalb der 24 Wochen vor Randomisierung, definiert als kurzfristige Komplikation (z. B. hyperosmolare oder ketoazidotische Krise) und/oder Vorgeschichte einer diabetischen kardiovaskulären Komplikation (z. B. Schlaganfall oder Herzinfarkt) • Nutzung von Erythropoese-stimulierenden Substanzen oder Hydroxyurea innerhalb von 24 Wochen vor Randomisierung • Eisenchelat-Therapie, wenn innerhalb von 24 Wochen vor Randomisierung begonnen • Unkontrollierte Hypertonie (kontrollierte Hypertonie ist definiert als ≤ Grad 1 nach den NCI-CTCAE Version 4.0) • Schwere Organschädigungen, inklusive o Lebererkrankung mit ALT > 3 x ULN oder Vorgeschichte einer Zirrhose o Herzerkrankung, Herzversagen klassifiziert auf Basis der New York Heart Association (NYHA) Klassifikation 3 oder höher, signifikante, behandlungsbedürftige Arrhythmie oder Herzinfarkt innerhalb von 6 Monaten vor Randomisierung o Lungenerkrankung, inklusive klinisch signifikanter pulmonarer Fibrose oder pulmonarer Hypertonie, z. B. ≥ Grad 3 NCI-CTCAE Version 4.0 o Kreatinin-Clearance < 60 mL/min (nach Cockroft-Gault-Formel) • Proteinurie ≥ Grad 3 nach NCI-CTCAE Version 4.0 • Dauerhafte Behandlung mit systemischen Glukokortikoiden ≤ 12 Wochen vor Randomisierung (physiologische Ersatztherapie bei adrenaler Insuffizienz war erlaubt), Behandlung mit Glukokortikoiden an einzelnen Tagen (z. B. zur Prävention oder Behandlung von Transfusionsreaktionen waren erlaubt)

<table>
<tr><th>Charakteris-tikum</th><th>Beschreibung</th></tr>
<tr><td></td><td>

Behandlung mit zytotoxischen Substanzen oder Immunsuppressiva innerhalb von 28 Tagen vor Randomisierung (z. B. Anti-Thymozyten-Globulin (ATG) oder Cyclosporin)
Vorgeschichte einer malignen Erkrankung mit Ausnahme von:

kurativ resezierten, nicht-melanomen Hauttumoren,
kurativ behandelten Zervixkarzinomen in situ,
anderen soliden Tumoren ohne aktive Erkrankung nach Maßgabe des ärztlichen Prüfpersonals

</td></tr>
<tr><td>Intervention und Zahl der Patientinnen und Patienten</td><td>
<table>
<tr><th>Luspatercept + BSC</th><th>Placebo + BSC</th></tr>
<tr><td>N = 224</td><td>N = 112</td></tr>
</table>
Insgesamt 447 Personen wurden innerhalb der Studie gescreent, 336 Personen wurden im Verhältnis 2:1 auf die beiden Studienarme randomisiert.
</td></tr>
<tr><td>Ort und Zeitraum der Durchführung</td><td>
Ort
65 Studienzentren in 15 Ländern (Australien, Malaysia, Taiwan, Thailand, Israel, Libanon, Tunesien, Türkei, Bulgarien, Kanada, Frankreich, Griechenland, Italien, Großbritannien und USA)
Die Rekrutierung, die 48-wöchige doppelblinde Behandlungsphase und (vermutlich auch) die doppelblinde Langzeit-Behandlungsphase der Studie sind abgeschlossen. Die Nachbeobachtung innerhalb der offenen Extensionsphase bzw. des Post-Treatment-Follow-up laufen noch.
Datenschnitt: 11. Mai 2018 (primäre Analyse)

Erste Patientin bzw. erster Patient, erste Visite: 02. Mai 2016
Letzte Patientin bzw. letzter Patient, erste Visite: 09. Juni 2017
Letzte Patientin bzw. letzter Patient, Visite zu Woche 48: 14. Mai 2018[1)]
Zeitpunkt der Entblindung: k. A.

Weitere Datenschnitte

07.01.2019 (Nachforderung EMA)
01.07.2019 (Nachforderung EMA)

</td></tr>
<tr><td>Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studien-protokoll</td><td>
Primärer Endpunkt

Anteil an Patientinnen und Patienten mit einer Reduktion der Erythrozytenkonzentrat-(EK)-Einheiten um ≥ 33 % und mit einer Reduktion von mindestens 2 EK-Einheiten innerhalb des Zeitraums von Woche 13 bis Woche 24 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung)

Sekundäre Endpunkte

Anteil an Patientinnen und Patienten mit einer Reduktion der EK-Einheiten um ≥ 33 % innerhalb des Zeitraums von Woche 37 bis Woche 48 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung)
Anteil an Patientinnen und Patienten mit einer Reduktion der EK-Einheiten um ≥ 50 % innerhalb des Zeitraums von Woche 13 bis Woche 24 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung)
Anteil an Patientinnen und Patienten mit einer Reduktion der EK-Einheiten um ≥ 50 % innerhalb des Zeitraums von Woche 37 bis Woche 48 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung)
Durchschnittliche Änderung der transfundierten EK-Einheiten von Baseline (Screeningphase) zu Woche 13 bis Woche 24

</td></tr>
</table>

Charakteristikum	Beschreibung
	Weitere Endpunkte • Durchschnittliche Änderung der Lebereisenkonzentration (gemessen mittels MRT) im Vergleich zu Baseline • Durchschnittliche Änderung der mittleren täglichen Dosis an Eisenchelattherapie in den letzten 12 Wochen der Behandlungsphase im Vergleich zu Baseline (Screeningphase) • Durchschnittliche Änderung des Serum-Ferritins innerhalb der letzten 12 Wochen der Behandlungsphase im Vergleich zu den letzten 4 Wochen vor Randomisierung • Durchschnittliche Änderung der Knochendichte (gemessen in Hüfte und Lendenwirbelsäule mittels DXA) • Durchschnittliche Änderung des myokardialen Eisens (gemessen mittels T2* MRT) • Durchschnittliche Änderung der gesundheitsbezogenen Lebensqualität (gemessen mittels TranQoL und SF-36) • Inanspruchnahme der Gesundheitsversorgung • Anteil an Patientinnen und Patienten, die ≥ 8 Wochen transfusionsfrei sind • Dauer der Reduzierung der Transfusionslast • Dauer der Transfusionsfreiheit • Zeit bis zur Reduktion der Transfusionslast • Zeit bis zur erythroiden Response • Häufigkeit der Transfusionen nach Baseline • Pharmakokinetik **Explorative Endpunkte** • GDF11 und weitere verwandte Biomarker: Änderung und Korrelation mit Wirksamkeit • Änderung im fetalen Hämoglobin **Sicherheitsendpunkte** • Unerwünschte Ereignisse • Häufigkeit von gegen Luspatercept gerichteten Antikörpern und Auswirkungen derer auf Wirksamkeit und Sicherheit
Subgruppenanalysen	Laut SAP waren für den primären Endpunkt, wichtige sekundäre Endpunkte und die 33 % bzw. 50 % erythroiden Responseendpunkte mit einem jeglichen (verschiebbaren) 12- bzw. 24-wöchigen Basisintervall folgende Subgruppenanalysen geplant: • Geographische Region (Nordamerika und Europa / Mittlerer Osten und Nordafrika / Asien-Pazifik) • Alter (≤ 32 Jahre / > 32 Jahre) • Splenektomie (ja/nein) • Geschlecht (männlich/weiblich) • β-Thalassämie-Genmutation (β_0/β_0 / nicht-β_0/β_0) • Transfusionslast zu Baseline (≤ 6 EK-Einheiten pro 12 Wochen / > 6 EK-Einheiten pro 12 Wochen) • Durchschnittlicher Prä-Transfusions-Hb-Wert zu Baseline (< 9 g/d / ≥ 9 g/dL) • Lebereisengehalt zu Baseline (≤ 3 mg/g Trockengewicht / > 3 bis ≤ 7 mg/g Trockengewicht / > 7 bis ≤ 15 mg/g Trockengewicht / > 15 mg/g Trockengewicht) • Transfusionslast zu Baseline für Endpunkte basierend auf einem jeglichen 24-wöchigen verschiebbaren Basisintervall (niedrige Transfusionslast (≤ 10 Einheiten/24 Wochen) / mittlere Transfusionslast (> 10 bis ≤ 15 Einheiten/24 Wochen) / hohe Transfusionslast (> 15 Einheiten/24 Wochen)) • Transfusionslast zu Baseline für Endpunkte basierend auf einem jeglichen 12-wöchigen verschiebbaren Basisintervall (niedrige Transfusionslast (≤ 5 Einheiten/12

Charakteris-tikum	Beschreibung
	Wochen) / mittlere Transfusionslast (> 5 bis ≤ 7 Einheiten/12 Wochen) / hohe Transfusionslast (> 7 Einheiten/12 Wochen)) Für die UE waren folgende Subgruppenanalysen im SAP geplant: • Alter (≤ 32 Jahre / > 32 Jahre und ≤ 50 Jahre / > 50 Jahre) • Splenektomie (ja/nein) • Geschlecht (männlich/weiblich) • ADA-Status (positiv vorbestehend / positiv therapiebedingt / negativ) Für die Lebensqualitätsendpunkte waren folgende Subgruppen im HRQoL-SAP geplant: • Responder vs. Nonresponder gemäß 4 unterschiedlicher Definitionen bezüglich der Transfusionslast • Alter (≤ 32 Jahre / > 32 Jahre) • Geographische Region (Nordamerika und Europa / Mittlerer Osten und Nordafrika / Asien-Pazifik) • Transfusionslast zu Baseline (≤ 6 EK-Einheiten pro 12 Wochen / > 6 EK-Einheiten pro 12 Wochen) Die Subgruppenanalysen im Dossier orientieren sich an den für den primären Endpunkt geplanten Analysen und wurden für alle im Dossier dargestellten Endpunkte durchgeführt. Für die Subgruppenanalyse nach geographischer Region wurden abweichend von der Präspezifizierung und Randomisierung nicht drei, sondern vier Kategorien analysiert: Europa / Nordamerika / Mittlerer Osten und Nordafrika / Asien-Pazifik.

[1)] Die letzte Visite des letzten Patienten oder der letzten Patientin (14. Mai 2018) erfolgte erst nach dem Datenschnitt (11. Mai 2018). Es ist unklar, inwieweit Daten dieses Patienten bzw. dieser Patientin zu Woche 48 in die Ergebnisse eingingen.

Abkürzungen: ADA: antidrug antibody; ALT: Alanin-Aminotransferase; EMA: Europäische Arzneimittel-Agentur; DXA: Dual-Röntgen-Absorptiometrie; ECOG: Eastern Cooperative Oncology Group; EK: Erythrozytenkonzentrat; GDF11: Growth differentiation factor 11; HBV: Hepatitis-B-Virus; HCV: Hepatitis-C-Virus; HIV: Humanes Immundefizienz-Virus; HRQoL-SAP: Statistischer Analyseplan für die Lebensqualität; k. A.: keine Angabe; MRT: Magnetresonanztomographie; RNA: Ribonukleinsäure; SAP: Statistischer Analyseplan; SF-36: Short-Form-36; TranQoL: Transfusion-dependent Quality of Life Questionnaire; ULN: Obere Grenze des Normalbereiches

Der pU legt einen Studienbericht für den Datenschnitt der primären Analyse vom 11.05.2018 vor. Für die weiteren Datenschnitte, welche laut Dossier zur Beantwortung von Nachfragen der EMA angefertigt wurden, wurden keine Studienberichte mit dem Dossier eingereicht. Der Zeitpunkt des Datenschnittes vom 11.05.2018 entspricht in etwa dem Abschluss der 48-wöchigen Behandlungsphase des letzten Studienteilnehmenden. Im Anschluss an die Langzeitbehandlungsphase erfolgte die Entblindung der Studie. Studienteilnehmende aus dem Luspatercept + BSC-Behandlungsarm konnten danach in der offenen Extensionsphase weiterbehandelt werden. Für Teilnehmende aus dem Placebo + BSC-Arm war ein Therapiebeginn mit Luspatercept in der offenen Extensionsphase möglich. Somit konnte ein Wechsel vom Placebo + BSC- auf den Interventionsarm erfolgen. Für die vorgelegten Zusatzanalysen zu den späteren Datenschnitten wurden jedoch keine Informationen zum weiteren Patientenfluss, zum Zeitpunkt der Entblindung, zur Exposition mit der Intervention sowie zur verabreichten Begleittherapie vorgelegt. Die Datenschnitte vom 07.01.2019 und 01.07.2019 werden daher nicht für die Nutzenbewertung herangezogen.

Das Studienprotokoll wurde bisher zwei Mal geändert. Die erste Änderung erfolgte am 21. April 2017 und damit am Ende der Rekrutierungsphase (letzter Patient / letzte Patientin, erste Visite am 9. Juni 2017). Folgende Änderungen wurden vorgenommen:

- Hinzufügung weiterer explorativer Untersuchungen (z. B. weitere Biomarker neben GDF11)
- Überarbeitung des Studiendesigns, um aktive Behandlungen während der offenen Extensionsphase unabhängig von der initialen Behandlungszuweisung zu erlauben und diesen offenen Studienteil in die Studie mit einzugliedern (keine separate Studie)
- Verlängerung des Post-Treatment-Follow-up von neun auf 156 Wochen
- Hinzufügung von Auflagen zum Berichten von Malignitäten und prämalignen Erkrankungen, die in der Post-Treatment-Phase auftreten
- Einführung eines Patiententagebuchs zur Erfassung von Transfusionen, die außerhalb des gewöhnlichen Transfusionszentrums durchgeführt wurden
- Konkretisierung des Behandlungszeitpunkts mit Dosis eins an Tag eins mit Luspatercept hinsichtlich Transfusionen, die am selben Tag verabreicht wurden: ein Abstand von mindestens 1 Stunde vor bzw. 2 Stunden nach der EK-Transfusion mit der Behandlung mit Luspatercept soll eingehalten werden
- Ein- und Ausschlusskriterien wurden hinsichtlich der folgenden Kriterien angepasst:
  - Einschluss von β-Thalassämie-Patientinnen und -Patienten mit Mutation und/oder Multiplikation des α-Globins
  - Spezifikation der Liste der anerkannten kontrazeptiven Methoden für Frauen im gebärfähigen Alter
  - Überarbeitung des Ausschlusskriteriums von Patientinnen und Patienten mit aktiver HCV- oder HBV-Infektion durch Festlegung einer Definition für das Vorhandensein einer Infektion
  - Überarbeitung des Ausschlusskriteriums von Personen mit Diabetes mellitus: Personen mit schlecht kontrollierten Diabetes mellitus in den 24 Wochen vor Studienbeginn sollten ausgeschlossen werden
  - Konkretisierung der Ausschlusskriterien bezüglich Lungenerkrankungen und pulmonaler Fibrose
  - Überarbeitung des Ausschlusskriteriums bezüglich Nebenniereninsuffizienz: Personen mit systemischen Glukokortikoiden zur Behandlung einer Nebenniereninsuffizienz waren erlaubt
  - Konkretisierung des Ausschlusskriteriums hinsichtlich der Behandlung mit Antikoagulantien; niedrig-molekulares Heparin und dauerhafte Therapie mit Acetylsalicylsäure wurden erlaubt
  - Einschluss von Patientinnen und Patienten mit geringer Leberzirrhose
  - Ausschluss von Personen mit Malignitäten aufgrund von neuen, nicht-klinischen toxikologischen Daten
- Weglassen der ECOG-Messung während der Behandlung
- Bestimmung der Retikulozyten und Erythroblasten durch das lokale Labor
- Hinzufügung des MRT als optionale Methode zur Bestimmung der linksventrikulären Ejektionsfraktion
- Überarbeitung der Beobachtungszeiträume für MRT- und DXA-Messungen während der Langzeitbehandlungsphase
- Aufnahme von Ultraschall als Messinstrument für Milzuntersuchungen

- Löschung der Knochendichtemessung (DXA) zu Woche 24
- Erhebung der Eisenchelattherapie, Folgetherapien, Transfusionen, Nutzung von Gesundheitseinrichtungen und Unerwünschten Ereignisse wurde auf einen Zeitraum von bis zu neun Wochen nach letzter Dosis Studienmedikation geändert
- Konkretisierungen zur Behandlung mit der Studienmedikation (Rekonstitution von Luspatercept, gesamte Maximaldosis von 120 mg, maximales Injektionsvolumen von 1,2 ml, Dosistitration/-reduktion, Verfügbarkeit des Placebos, etc.)
- Ergänzung Anagrelid als nicht-erlaubte Begleitmedikation
- Konkretisierung der Kriterien für Dosisverzögerungen, Therapieabbruch und verbotene Begleitmedikationen
- Konkretisierung der Erhebung von UE nach Abbruch der Behandlung
- Hinzufügen, dass die Behandlung beim Neuauftreten von Malignitäten abgebrochen werden soll
- Überarbeitung der statistischen Methoden
- Entfernung der ursprünglich geplanten Interim-Analyse und Konkretisierung der Zielsetzung des Studienberichtes
- Ergänzung weiterer Informationen zu potentiellen Risiken basierend auf neuen Erkenntnissen aus klinischen und nicht-klinischen Studien

Eine zweite Änderung des Protokolls erfolgte am 11. Dezember 2018 und damit nach Durchführung des zur Nutzenbewertung herangezogenen Datenschnittes vom 11. Mai 2018. Die wesentlichen Änderungen lagen in der Anpassung der Dosierung der Studienmedikation gemäß einer Anforderung der FDA:

- Verzögerung der Dosierung im Fall von Anämie und Abbruch der Verabreichung der Studienmedikation bei Bestätigung einer hämatologischen Malignität
- Verzögerung der Dosierung bei Vorliegen einer Grad ≥ 3 Leukopenie, Neutropenie und/oder Thrombozytopenie und Abbruch der Verabreichung der Studienmedikation bei Bestätigung einer hämatologischen Malignität

Tabelle 3: Charakterisierung der Intervention Luspatercept während der (Langzeit-)Behandlungsphase der Studie ACE-536-B-THAL-001 (BELIEVE)

Intervention	**Kontrolle**
Luspatercept Bei Behandlungsbeginn sollte die Dosis an Luspatercept 1 mg/kg Körpergewicht betragen. Die Gesamtdosis sollte 120 mg pro Dosierung nicht übersteigen.	**Placebo** Als Placebo wurde 0,9 % Natriumchloridlösung eingesetzt. Prüfpersonal und Patient/in waren verblindet, sodass die Regeln zur Dosisreduktion, Verschiebung oder Abbruch (s. u.) analog der Interventionsgruppe erfolgten.
Die subkutane Verabreichung der Studienmediaktion erfolgte alle 21 Tage unter Berücksichtigung folgender Bedingungen: • Hb-Wert < 11,5 g/dL und Veränderung im Hb-Wert im Vergleich zum Ausgangswert vor der vorangegangenen Dosierung ≤ 2g/dL • Auf die Behandlung zurückzuführende Unerwünschte Ereignisse haben einen NCI-CTCAE-Schweregrad < Grad 2 • Leukozytenzahl < dreifach Baseline-Leukozytenzahl (= höchster Wert während der Screeningphase)	

Intervention	Kontrolle
Schrittweise Dosisreduktionen auf 0,8 mg/kg Körpergewicht, 0,6 mg/kg Körpergewicht oder 0,45 mg/kg Körpergewicht sowie eine Dosierung zu einem späteren Zeitpunkt waren nach festgelegten Kriterien vorgesehen, wenn die zuvor beschriebenen Bedingungen nicht erfüllt wurden. Eine Verschiebung der Dosierung von 15 Wochen oder länger führte zum Abbruch der Behandlung. Ein Abbruch der Behandlung erfolgte weiterhin nach mehr als 2 Dosisreduktionen aufgrund von UE, die auf die Behandlung zurückzuführen waren, und Grad 3 Leukozytose. Dosistitration zu 1,25 mg/kg Körpergewicht war während der (Langzeit-)Behandlungsphase möglich, solange keine Dosisverzögerung oder -verringerung erforderlich war. Eine Dosiserhöhung konnte unter folgenden Bedingungen erfolgen: • Verminderung der Transfusionslast um < 33 % über mindestens 2 Dosiszyklen (~ 6 Wochen) im Vergleich zu Baseline oder • Verminderung der Transfusionslast um ≥ 33 %, aber < 50 % über mindestens 2 Dosiszyklen (~ 6 Wochen) im Vergleich zu Baseline im Ermessen des Prüfpersonals	
Best Supportive Care (BSC) Best Supportive Care beinhaltete Transfusionen mit EK, Eisenchelattherapie sowie Antibiotikatherapie, antivirale und antimykotische Therapie und/oder Ernährungsunterstützung nach Bedarf.	
Nicht erlaubte Begleitmedikation • Hydroxyurea • Anagrelid • Hämatopoetische Wachstumsfaktoren • Antikoagulantien zur Behandlung von UE Sofern eine Behandlung während der Behandlungsphase notwendig war, wurde die Behandlung mit der Studienmedikation abgebrochen und die Person ging in die Post-Treatment-Follow-up-Phase über. **Erlaubte Begleitmedikation** • Eisenchelattherapie • Erythrozytenkonzentrate nach Ermessen des Prüfpersonals bei niedrigen Hb-Werten, mit Anämie assoziierten Symptomen (z. B. behandlungsbedürftigen, hämodynamischen oder pulmonaren Beeinträchtigungen) oder Komorbiditäten. Für jede Person wird individuell ein Prä-Transfusions-Grenzwert auf Basis der Transfusionslast 24 Wochen vor Randomisierung festgelegt. Überschreitet der Hb-Wert vor einer geplanten Transfusion den Prä-Transfusions-Grenzwert um 1 g/dL oder mehr, so soll die Transfusion um mindestens 7 Tage verschoben werden und/oder das Transfusionsvolumen um eine Einheit oder mehr reduziert werden. • Stabile Dosis an Medikamenten zur Behandlung chronischer Erkrankungen, die nicht unter die Ausschlusskriterien fielen • Antikoagulantien, Acetylsalicylsäure und niedrig molekulares Heparin zur Prophylaxe	

Abkürzungen: BSC: Best supportive care; NCI-CTCAE: National Cancer Institute - Common Terminology Criteria for Adverse Events; UE: Unerwünschte Ereignisse

2.3 Endpunkte

In diesem Kapitel werden die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Es werden nur Endpunkte aufgeführt, die für den Datenschnitt vom 11.05.2018 analysiert wurden. Dazu wurden das Herstellerdossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. Tabelle 4 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 4: Zusammenfassung der Endpunktbewertung

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Gesamtüberleben	Mortalität	Ja	Ja
Transfusionslast[1)]	Morbidität	Ja	Nein
Transfusionsfreiheit[2)]	Morbidität	Ja	Nein
Hospitalisierung	Morbidität	Ja	Ja
Reduktion der Eisenüberladung[3)]	Morbidität	Ergänzend	Nein
SF-36	Lebensqualität	Ja	Ja
TranQoL	Lebensqualität	Ja	Ja
Unerwünschte Ereignisse	Sicherheit	Ja	Ja

[1)] Unter dem Endpunkt Transfusionslast werden mehrere Endpunkte, welche die Reduktion der transfundierten EK-Einheiten über einen bestimmten Zeitraum erfassen, zusammengefasst. Der primäre Endpunkt ist operationalisiert als Anteil an Patientinnen und Patienten mit einer Reduktion der Erythrozytenkonzentrat-(EK)-Einheiten um ≥ 33 % und mit einer Reduktion von mindestens 2 EK-Einheiten innerhalb des Zeitraums von Woche 13 bis Woche 24 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung).
[2)] Die Transfusionsfreiheit bezeichnet den Anteil an Patientinnen und Patienten ohne Transfusion über ≥ 8 (bzw. ≥ 6 oder ≥ 12) Wochen.
[3)] Unter dem Endpunkt Reduktion der Eisenüberladung werden Analysen zur Lebereisenkonzentration, zur myokardialen Eisenkonzentration, zum Serum-Ferritin und zur Dosis der Eisenchelattherapie zusammengefasst.

Abkürzungen: SF-36: Short-Form-36; TranQoL: Transfusion-dependent Quality of Life Questionnaire

2.3.1 Mortalität

Gesamtüberleben

Der Endpunkt Gesamtüberleben wird in der Nutzenbewertung berücksichtigt.

Operationalisierung:

Beschreibung

Todesfälle wurden im Rahmen der Sicherheit erhoben. Im Protokoll oder SAP wurden keine Analysen bezüglich des Gesamtüberlebens präspezifiziert. Laut Modul 4 wurde das Gesamtüberleben als Zeit zwischen der Randomisierung und dem Todeszeitpunkt operationalisiert. Verstorbene Patientinnen und Patienten, unabhängig von der Todesursache, wurden als Patientinnen und Patienten mit Ereignis berücksichtigt. Patientinnen und Patienten, die zum Zeitpunkt der Analyse noch lebten, die Studie abgebrochen haben oder während der Verlaufsbeobachtung verloren gegangen sind, wurden zu dem letzten Zeitpunkt zensiert, an dem sie bekanntermaßen noch am Leben waren.

- Erhebungszeitpunkte:
  Gesamtüberleben wurde im Rahmen der Sicherheit kontinuierlich erhoben.
- Auswertung:
  Die Auswertung erfolgte für Modul 4 auf Basis der ITT-Population unter Verwendung einer Ereigniszeitanalyse.

Bewertung

Es erfolgte keine Erhebung des Gesamtüberlebens über die Teilnahme an Studienvisiten hinaus.

Patientenrelevanz:

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Der Endpunkt Gesamtüberleben ist patientenrelevant.

Validität:

Gesamtüberleben wurde im Rahmen der Sicherheit erfasst. Personen, die nicht Teil der Sicherheitspopulation waren, wurden in diese Erfassung nicht mit einbezogen. Auswertungen sind daher nur bezogen auf die Sicherheitspopulation möglich. Aufgrund der geringen Ereigniszahl ist eine Ereigniszeitanalyse der Daten nicht sinnvoll.

2.3.2 Morbidität

Transfusionslast

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

Operationalisierung:

Beschreibung

Für den Endpunkt Transfusionslast existieren unterschiedliche Operationalisierungen. Im Studienprotokoll wurden folgende definiert:

- Anteil an Patientinnen und Patienten mit einer Reduktion der EK-Einheiten um ≥ 33 % und mit einer Reduktion von mindestens 2 EK-Einheiten innerhalb des Zeitraums von Woche 13 bis Woche 24 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung)
- Anteil an Patientinnen und Patienten mit einer Reduktion der EK-Einheiten um ≥ 33 % innerhalb des Zeitraums von Woche 37 bis Woche 48 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung)
- Anteil an Patientinnen und Patienten mit einer Reduktion der EK-Einheiten um ≥ 50 % innerhalb des Zeitraums von Woche 13 bis Woche 24 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung)
- Anteil an Patientinnen und Patienten mit einer Reduktion der EK-Einheiten um ≥ 50 % innerhalb des Zeitraums von Woche 37 bis Woche 48 im Vergleich zur Screeningphase (12 Wochen vor Randomisierung)
- Durchschnittliche Änderung der transfundierten EK-Einheiten von Baseline (Screeningphase) zu Woche 13 bis Woche 24
- Dauer der Reduzierung der Transfusionslast
- Häufigkeit der Transfusionen nach Baseline

Zusätzlich sind im SAP Analysen beschrieben, bei denen statt der oben genannten festgelegten Intervalle jegliche verschiebbare Intervalle über 12 bzw. 24 Wochen im Vergleich zu Baseline

angewendet werden. Eine Response in diesen Analysen ist definiert als eine Reduktion der Transfusionslast um 33 % bzw. 50 % über ein jegliches 12- bzw. 24-Wochen-Intervall.

Post-hoc wurden daneben die mittlere Änderung des Hb-Wertes und die maximale Streckung eines Transfusionsintervalls berechnet.

Die Gabe von EK erfolgte nach Ermessen des Prüfpersonals bei niedrigen Hb-Werten, mit Anämie assoziierten Symptomen (z. B. behandlungsbedürftigen, hämodynamischen oder pulmonaren Beeinträchtigungen) oder Komorbiditäten. Für jede Person wurde individuell ein Prä-Transfusions-Grenzwert auf Basis der Transfusionslast 24 Wochen vor Randomisierung festgelegt. Überschritt der Hb-Wert vor einer geplanten Transfusion den Prä-Transfusions-Grenzwert um 1 g/dL oder mehr, so sollte die Transfusion um mindestens 7 Tage verschoben werden oder das Transfusionsvolumen um eine Einheit oder mehr reduziert werden. Für jede Transfusion wurde der Hb-Wert vor der Transfusion, die Anzahl an infundierten Einheiten und das Datum der Transfusion im Studienzentrum erfasst. Für Transfusionen, die außerhalb des Studienzentrums erfolgten, wurde mit dem Protokoll-Amendment 1 ein Patiententagebuch zur Erfassung dieser Infusionen eingeführt.

- Erhebungszeitpunkte:
  Die Anzahl der verabreichten EK-Einheiten wurde kontinuierlich während der gesamten Studiendauer erfasst. Die Erfassung erfolgte bis zur Woche 48 oder bis zu 9 Wochen nach Beendigung der Therapie mit Luspatercept, je nachdem, welcher Zeitpunkt später war. Die Erhebung des Hb-Wertes erfolgte vor jeder geplanten Transfusion.
- Auswertung:
  Die Auswertung erfolgte auf Basis der ITT-Population für den Datenschnitt vom 11. Mai 2018.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz:

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Eine langfristige Vermeidung von Transfusionen (Transfusionsunabhängigkeit) unter Einhaltung eines definierten Mindestwertes an Hämoglobin stellt in dem vorliegenden Anwendungsgebiet einen patientenrelevanten Endpunkt dar, da dadurch das im vorliegenden Anwendungsgebiet vorrangige Therapieziel der Vermeidung der Anämie erreicht werden kann.

Der transfusionsabhängigen β-Thalassämie liegt eine durch die deutlich reduzierte Produktion von funktionsfähigem β-Globin verursachte Anämie zugrunde, die häufige und lebenslange Transfusionen mit EK erforderlich macht. Die erforderlichen Transfusionen können bei den Patientinnen und Patienten darüber hinaus trotz einer Eisenchelattherapie zu einer zunehmenden Eisenüberladung der Organe und daraus folgenden Langzeitkomplikationen führen.

Die Transfusionslast in den vorliegenden Operationalisierungen stellt keinen per se patientenrelevanten Endpunkt dar, da sie keine Aussagen über eine langfristige Vermeidung von Transfusionen im Sinne einer Transfusionsunabhängigkeit erlaubt. Beim Hb-Wert handelt es sich um einen Laborparameter, der als solcher nicht per se patientenrelevant ist. Studien zur

Validierung als Surrogatparameter wurden nicht vorgelegt. Zudem wird die Transfusionsbedürftigkeit nicht ausschließlich auf Basis des Hb-Wertes festgelegt.

Validität:

Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt.

Transfusionsfreiheit

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

Operationalisierung:

Beschreibung

Gemäß Studienprotokoll ist die Transfusionsfreiheit definiert als zusammenhängender Zeitraum ohne Transfusion von EK über einen Zeitraum von 8 Wochen. Laut SAP werden zusätzlich Analyse über einen Zeitraum von 6 oder 12 Wochen durchgeführt. Im Dossier wird die Analyse über den Zeitraum von 8 Wochen als Hauptanalyse und Analysen über 6 bzw. 12 Wochen als Sensitivitätsanalysen berichtet.

Die Gabe von EK erfolgte nach Ermessen des Prüfpersonals bei niedrigen Hb-Werten, mit Anämie assoziierten Symptomen (z. B. behandlungsbedürftigen, hämodynamischen oder pulmonaren Beeinträchtigungen) oder Komorbiditäten. Für jede Person wurde individuell ein Prä-Transfusions-Grenzwert auf Basis der Transfusionslast 24 Wochen vor Randomisierung festgelegt. Überschritt der Hb-Wert vor einer geplanten Transfusion den Prä-Transfusions-Grenzwert um 1 g/dL oder mehr, so sollte die Transfusion um mindestens 7 Tage verschoben werden oder das Transfusionsvolumen um eine Einheit oder mehr reduziert werden. Für jede Transfusion wurde der Hb-Wert vor der Transfusion, die Anzahl an infundierten Einheiten und das Datum der Transfusion im Studienzentrum erfasst. Für Transfusionen, die außerhalb des Studienzentrums erfolgten, wurde mit dem Protokoll-Amendment 1 ein Patiententagebuch zur Erfassung dieser Infusionen eingeführt.

- Erhebungszeitpunkte:
  Die Anzahl der verabreichten EK-Einheiten wurde kontinuierlich während der gesamten Studiendauer erfasst. Die Erfassung erfolgte bis zur Woche 48 oder bis zu 9 Wochen nach Beendigung der Therapie mit Luspatercept, je nachdem, welcher Zeitpunkt später war.
- Auswertung:
  Die Auswertung erfolgte auf Basis der ITT-Population für den Datenschnitt vom 11. Mai 2018.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz:

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Eine langfristige Vermeidung von Transfusionen (Transfusionsunabhängigkeit) unter Einhaltung eines definierten Mindestwertes an

Hämoglobin stellt in dem vorliegenden Anwendungsgebiet einen patientenrelevanten Endpunkt dar, da dadurch das im vorliegenden Anwendungsgebiet vorrangige Therapieziel der Vermeidung der Anämie erreicht werden kann.

Der transfusionsabhängigen β-Thalassämie liegt eine durch die deutlich reduzierte Produktion von funktionsfähigem β-Globin verursachte Anämie zugrunde, die häufige und lebenslange Transfusionen mit EK erforderlich macht. Die erforderlichen Transfusionen können bei den Patientinnen und Patienten darüber hinaus trotz einer Eiseneliminationstherapie zu einer zunehmenden Eisenüberladung der Organe und daraus folgenden Langzeitkomplikationen führen.

Die Transfusionsfreiheit in den vorliegenden Operationalisierungen über 6, 8 oder 12 Wochen stellt keinen per se patientenrelevanten Endpunkt dar, da damit keine langfristige Freiheit von Transfusionen beschrieben wird. Die Intervalle der Transfusionsfreiheit werden durch die Gabe von EK begrenzt. Eine dauerhafte Unabhängigkeit von Transfusionen mit EK wird mit der vorliegenden Operationalisierung nicht beschrieben.

Validität:

Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt.

Hospitalisierung

Der Endpunkt Hospitalisierung wird in der Nutzenbewertung berücksichtigt.

Operationalisierung:

Beschreibung

Der Endpunkt Hospitalisierung ist definiert als die Anzahl an Personen mit Hospitalisierungen nach Unterzeichnung der Einverständniserklärung. Neben der Hospitalisierung an sich wurden die Hospitalisierungsgründe erfasst. Diese wurden in folgende Kategorien eingeteilt:

- Unerwünschte Ereignisse
- durch das Protokoll verursachte Prozeduren (z. B. Knochenmarksentnahme)
- nicht durch das Protokoll verursachte Prozeduren (z. B. Ultraschall)
- Transfusion (z. B. EK, Thrombozyten)
- Prozeduren, die bereits vor Unterzeichnung der Einverständniserklärung geplant waren (z. B. koronares Arteriogramm)
- geplante Prozeduren für eine bereits bestehende Erkrankung (z. B. Behebung von Hernien)
- soziale, technische oder praktische Gründe in Abwesenheit eines Unerwünschten Ereignisses (Entfernung zum Studienzentrum, die ansonsten eine Studienteilnahme verhindern würde)

- Erhebungszeitpunkte:
  Der Endpunkt wurde kontinuierlich ab Einschluss in die Studie bis zur Woche 48 oder bis zu 9 Wochen nach Beendigung der Therapie mit Luspatercept, je nachdem, welcher Zeitpunkt später war, erhoben.

- Auswertung:
  Die Auswertung erfolgte auf Basis der ITT-Population. Für die Hospitalisierung waren im SAP lediglich deskriptive Analysen geplant. Für das Dossier wurden vom pU zusätzlich

Unterschiede zwischen den Behandlungsarmen für die Gesamtrate (Hospitalisierung jeglicher Ursache) und Hospitalisierung aufgrund von UE / Transfusionen mittels CMH-Methode stratifiziert nach Region (Stratifizierungsvariable der Randomisierung) auf Basis von Odds Ratios (OR) berechnet. Relative Risiken im Sinne von Risk Ratios (RR) und Risikodifferenzen (RD) wurden unstratifiziert mittels Normal-Approximation berechnet. Für die Auswertungen im Dossier wurde lediglich die Behandlungsphase (Woche 1–48) herangezogen. Die deskriptiven Analysen im Studienbericht berücksichtigen dagegen die gesamte verblindete Studiendauer bestehend aus Behandlungsphase und Langzeitbehandlungsphase.

Bewertung

Die Operationalisierung ist nur eingeschränkt nachvollziehbar. Es ist unklar, wann die Aufnahme in ein Krankenhaus als Hospitalisierung gewertet wurde, z. B. ob dafür eine Aufnahme über Nacht oder einen bestimmten Zeitraum notwendig war oder ob eine ambulante oder teilstationäre Aufnahme ausreichte. Weiterhin ist nicht bei allen erfassten Hospitalisierungsgründen davon auszugehen, dass sie mit der Studienmedikation in Zusammenhang stehen.

Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode aufgrund der längeren Behandlungs- und Beobachtungszeit bevorzugt. Die Effektschätzer auf Basis der Behandlungsphase werden daher nicht dargestellt.

Patientenrelevanz:

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Der Endpunkt Hospitalisierung wird als patientenrelevant erachtet.

Validität:

Da es sich um eine multizentrische Studie handelt, bleibt unklar, ob regionale Unterschiede bestehen, die zu einer Verzerrung der Anzahl an Hospitalisierungen führen können. Die Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext ist unsicher. Der pU beschreibt im Modul 4 eine gute Übertragbarkeit auf den deutschen Versorgungskontext, legt aber keine Begründung dafür dar. Die externe Validität des Endpunktes kann somit nicht abschließend beurteilt werden. Auch die interne Validität kann durch regionale Unterschiede eingeschränkt sein, da die Stratifizierung der Randomisierung nicht auf Ebene der Studienzentren oder Länder erfolgte, sondern auf Ebene geographischer Regionen. Eine Ungleichverteilung der Testpersonen über Länder mit verschiedenen Kriterien der Notwendigkeit einer Hospitalisierung kann zu einer Verzerrung der Ergebnisse führen.

Für die Hospitalisierung aufgrund eines Unerwünschten Ereignisses erfolgte eine Doppelerfassung der Ereignisse im Endpunkt Hospitalisierung und im Endpunkt Unerwünschte Ereignisse.

Reduktion der Eisenüberladung

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

Operationalisierung:

Beschreibung

Unter dem Endpunkt Reduktion der Eisenüberladung werden Analysen zur Lebereisenkonzentration, zur myokardialen Eisenkonzentration, zum Serum-Ferritin und zur Dosis der Eisenchelattherapie zusammengefasst.

Die Messung der Lebereisenkonzentration erfolgte mittels MRT (T2* oder R2, gleiche Methode pro Person zu jedem Zeitpunkt).

Die Messung der myokardialen Eisenkonzentration erfolgte mittels T2* MRT.

Für jede Person wird der durchschnittliche Serumferritin-Baselinewert während der 12 Wochen vor der erstmaligen Behandlung (Screeningphase) mit dem Prüfpräparat berechnet. Der mittlere Serumferritin-Wert für die Behandlungsphase wird aus den Messwerten der letzten 12 Wochen der 48-wöchigen doppelblinden Behandlungsperiode oder während der letzten 12 Wochen der Studienbehandlung bei vorzeitigem Studienabbruch berechnet.

Bezüglich der Eisenchelattherapie wurde der Wirkstoff, Zeitraum und Dosisänderungen erfasst. Der Bedarf an Eisenchelatoren war definiert als durchschnittliche tägliche Dosis eines Eisenchelators über einen Zeitraum von 12 Wochen. Als Baseline-Werte wurden die Werte in den 12 Wochen vor Behandlungsbeginn herangezogen. Der Bedarf während der Behandlungsphase war definiert als mittlere tägliche Dosis eines Eisenchelators während der letzten 12 Wochen der 48-Wochen-Behandlungsphase oder während der letzten 12 Wochen der Studienbehandlung bei vorzeitigem Studienabbruch. Personen mit < 12 Wochen Behandlungsphase wurden als fehlende Werte berücksichtigt. Nur Personen, die mit dem gleichen Eisenchelator in Monotherapie über den gesamten Analysezeitraum hinweg behandelt wurden, ohne diesen zu wechseln, wurden in den Analysen berücksichtigt.

- Erhebungszeitpunkte:
  Der Endpunkt Lebereisenkonzentration wurde zu Baseline, Woche 24, Woche 48 und Woche 96 erhoben. Der Endpunkt myokardiale Eisenkonzentration wurde zu Baseline, Woche 48 und Woche 96 erhoben. Für beide Endpunkte erfolgte keine Erhebung in Woche 96, wenn die Studienteilnehmenden nicht in die Langzeitnachbeobachtungsphase übergingen. Bei vorzeitiger Beendigung der Studie konnten die Endpunkte auch zum Zeitpunkt des Ausscheidens aus der Studie erhoben werden.

  Der Endpunkt Serum-Ferritin wurde am Tag 1 jeder Dosierung mit der Studienmedikation erhoben. Ein Screeningwert wurde ca. 12 Wochen vor der Randomisierung erhoben. Weitere Zeitpunkte der Erhebung während der Screeningphase sind unklar. Die Messung erfolgte in der Regel in den Studienzentren. Zusätzlich wurde ein Serum-Ferritinwert von einem zentralen Labor ca. 4 Wochen vor Beginn der Dosierung mit der Studienmedikation erhoben. Nach Woche 48 wurde in der Langzeitbehandlungsphase der Serum-Ferritinwert nicht mehr bei jeder Dosierung, sondern nur noch alle 4 Dosen erhoben. In der Open-Label-Phase erfolgte keine regelhafte Erhebung. Bei Personen, die die Behandlung abbrachen, wurde der Serum-Ferritinwert 22 Tage und 9 Wochen nach der letzten Dosierung gemessen.
- Auswertung:
  Die Auswertung erfolgte auf Basis *der ITT-Population*. Es wird die mittlere Änderung zu Baseline mittels ANCOVA mit Kovariablen für die Randomisierungs-Stratifizierungsvariable Region und den jeweiligen Baselinewert berechnet.

Bewertung

Die Operationalisierung der Endpunkte Lebereisenkonzentration, myokardiale Eisenkonzentration und mittlere Dosis der Eisenchelattherapie ist nachvollziehbar.

Die Erhebung der Serum-Ferritinwerte während der Screeningphase ist teilweise unklar. Der Baselinewert soll als Mittelwert über die 12-wöchige Screeningperiode berechnet werden. Die erste Erhebung erfolgte zu Beginn der Screeningperiode (Messung im lokalen Labor) und eine zweite Erhebung 4 Wochen vor der Randomisierung (Messung im zentralen Labor). Inwieweit weitere Messungen erfolgten und ob der Mittelwert aus Messergebnissen unterschiedlicher Labore gebildet wurde, ist unklar.

Patientenrelevanz:

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Der transfusionsabhängigen β-Thalassämie liegt eine durch die deutlich reduzierte Produktion von funktionsfähigem β-Globin verursachte Anämie zugrunde, die häufige und lebenslange Transfusionen mit EK erforderlich macht. Die erforderlichen Transfusionen können bei den Patientinnen und Patienten trotz einer Eisenchelattherapie zu einer zunehmenden Eisenüberladung der Organe und daraus folgenden Langzeitkomplikationen führen. Mit der Lebereisenkonzentration kann die Eisenüberladung der Leber gemessen werden. Sie wird als Marker der gesamten Eisenüberladung des Körpers gesehen. Die Messung erfolgt mittels MRT. Folgen und Symptome einer Eisenüberladung der Leber werden dadurch nicht erfasst. Der Endpunkt ist daher nicht unmittelbar patientenrelevant. Der pU legt keine Studien zur Surrogatvalidierung der Lebereisenkonzentration für patientenrelevante Endpunkte vor. Mit der myokardialen Eisenkonzentration wird die Eiseneinlagerung im Herzen mittels MRT gemessen. Diese stellt keinen unmittelbar patientenrelevanten Endpunkt dar. Die Eisenüberladung im Herzen kann langfristig zu Kardiomyopathien und dadurch bedingten Tod führen. Der pU legt im Dossier keine Daten vor, inwieweit die myokardiale Eisenkonzentration ein validiertes Surrogat für patientenrelevante Endpunkte darstellt.

Serum-Ferritin ist ein Laborparameter und als solcher nicht per se patientenrelevant. Der pU legt im Dossier keine Daten vor, inwieweit Serum-Ferritin ein validiertes Surrogat für patientenrelevante Endpunkte darstellt.

Die Veränderung der Häufigkeit oder Dosis einer Medikamenteneinnahme, wie Eisenchelate, ist nicht patientenrelevant. Die damit einhergehende mögliche Veränderung des Auftretens von Unerwünschten Ereignissen wird in der Endpunktkategorie Sicherheit erfasst.

Validität:

Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt.

2.3.3 Lebensqualität

TranQoL

Der Endpunkt TranQoL wird in der Nutzenbewertung berücksichtigt.

Operationalisierung:

Beschreibung

Beim Transfusion-dependent Quality of Life Questionnaire (TranQoL) handelt es sich um einen krankheitsspezifischen Fragebogen zur Messung der Lebensqualität von Personen mit transfusionsabhängigen Thalassämien. Er liegt in 4 verschiedenen Versionen vor. In der BELIEVE-Studie wurde die selbstauszufüllende Fragebogenversion für Erwachsene verwendet. Sie enthält 36 Fragen, die in 5 Domänen eingeteilt werden:

- körperliche Gesundheit (Fragen 1–10)
- emotionale Gesundheit (Fragen 11–24)
- sexuelle Aktivität (Frage 25)
- familiäre Situation (Fragen 26–30)
- Schule und Arbeit (Fragen 31–36)

Der Bezugszeitraum des Fragebogens beträgt eine Woche. Es wird ein Gesamtscore über alle Fragen hinweg berechnet. Das Scoring für den TranQoL ist im HRQoL-SAP dargelegt. Bei bis zu 25 % fehlender Werte innerhalb einer Domäne wird der Score berechnet und entsprechend der Anzahl an beantworteten Fragen gewichtet. Der Fragebogen galt insgesamt als auswertbar, wenn ≥ 75 % (≥ 27 der 36 Fragen) der Items beantwortet wurden. Ein höherer Score beschreibt eine höhere Lebensqualität. Zur Auswertung der Daten zur Lebensqualität wird eine „HRQoL evaluable population“ definiert. Diese enthält alle Personen der ITT-Population mit einem auswertbaren Fragebogen zu Baseline und mindestens einem Zeitpunkt nach Baseline.

Im HRQoL-SAP wurde eine Änderung von 4–6 Punkten als Responderdefinition für den TranQoL-Gesamtscore festgelegt.

- Erhebungszeitpunkte:
  Der Endpunkt wurde zu folgenden Zeitpunkten erhoben:
  - Screeningphase: innerhalb von 4 Wochen vor der ersten Dosierung
  - Behandlungsphase: Tag 1 der Studienwoche 12, 24, 36 und 48.
  - Alle 12 Wochen während der Langzeitbehandlungsphase
  - Keine Erhebung nach Behandlungsabbruch sowie in der offenen Extensionsphase
- Auswertung:
  Die Auswertung der Daten zum TranQoL erfolgt auf Basis der „HRQoL evaluable population“. Diese enthält alle Personen der ITT-Population mit einem auswertbaren Fragebogen zu Baseline und mindestens einem Zeitpunkt nach Baseline. Es wurden Mittelwertsdifferenzen zu Woche 24, Woche 48 und über die gesamte Studienperiode hinweg mittels ANCOVA bzw. MMRM-Modellen berechnet (siehe Abschnitt 2.4). Daneben wurden auch Responderanalysen zu Woche 24 und 48 geplant (siehe Abschnitt 2.4). Als primäre Domäne/Skala von Interesse wurde der Gesamtwert des TranQoL und die physische Gesundheitsdomäne präspezifiziert. Alle anderen Domänen/Skalen werden explorativ analysiert.

Bewertung

Die Operationalisierung ist nachvollziehbar. Es erfolgte keine Erhebung des TranQoL im Fall eines Behandlungsabbruches. Da anzunehmen ist, dass Personen, die die Behandlung abbrechen, tendenziell eine schlechtere Lebensqualität haben als Personen, die die Behandlung weiterführen, führt die fehlende Erhebung des TranQoL im Fall eines Behandlungsabbruches zur Verzerrung.

Patientenrelevanz:

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Der TranQoL misst die krankheitsspezifische, gesundheitsbezogene Lebensqualität und ist daher als patientenrelevant anzusehen.

Validität:

Der TranQoL wurde speziell für Patientinnen und Patienten mit Thalassämie major mit regelmäßigem Transfusionsbedarf entwickelt. In die Entwicklung waren insgesamt 120 Teilnehmende involviert: 31 Kinder, 30 Erwachsene, 43 Eltern, 16 Mitarbeitende des Gesundheitswesens [11]. Der TranQoL wurde anschließend an zwei Stichproben (Kinder, Erwachsene mit transfusionsabhängiger β-Thalassämie) in sechs Gesundheitseinrichtungen der USA und Kanada getestet [12]. Daneben liegen Validierungsstudien für die griechische [13] und persische [3] Übersetzung bei Personen mit transfusionsabhängiger β-Thalassämie vor. Die Studien zeigen insgesamt eine gute Test-Retest-Reliabilität für den Gesamtscore sowie für alle Domänen (ICC = 0,74–0,94 [12,13]) und akzeptable bis hohe interne Konsistenz für den Gesamtscore und die Einzeldomänen (Cronbach’s alpha = 0,64 (familiäre Situation) – 0,96 (Gesamtscore) [3,12,13]). In der Validierungsstudie der persischen Übersetzung wurde eine explorative Faktorenanalyse durchgeführt. Diese zeigte, dass 5 Items keine Korrelation zu den erhaltenen Faktoren zeigte und dementsprechend entfernt wurden [3]. Keine andere Validierungsstudie führte eine Faktorenanalyse durch. Im Rahmen der Untersuchungen zur konvergenten Validität des Gesamtscores wurden moderate Korrelationen zur PCS und MCS des SF-36 beobachtet [12,13]. Ausführliche Ergebnisse über Korrelationen der Einzelskalen wurden in den Publikationen nicht dargestellt. Die Analyse der Known-Groups-Validität des Gesamtscores zeigte signifikante Unterschiede zwischen Patientinnen und Patienten mit Komorbiditäten in einer der beiden Studien [12,13]. In der zweiten Studie wurden lediglich signifikante Unterschiede zwischen Personen mit bzw. ohne Hypothyreose gezeigt und keine signifikanten Unterschiede zwischen Personen mit bzw. ohne jegliche Komorbiditäten festgestellt [13]. In beiden Studien konnten keine signifikanten Unterschiede im Gesamtscore zwischen Patientinnen und Patienten mit unterschiedlichen Arten der Eisenchelattherapie (oral vs. subkutan) gezeigt werden [12,13]. Es liegen Hinweise auf Änderungssensitivität des Gesamtscores für Verbesserungen bei einer kleinen Anzahl an Patientinnen und Patienten vor [12,13]. Studien mit dem Ziel der Validierung einer MID wurden nicht durchgeführt. Bei Untersuchungen zur Änderungssensitivität wurde bei Personen, die angaben, sich eine Woche nach der Infusion ein wenig besser als vor der Infusion zu fühlen, eine mittlere Änderung des Gesamtscores von +6 (SD = 9,5, N = 18 [12]) bzw. +4,2 Punkten (SD = 5,1, N = 16 [13]) beobachtet. Der pU leitet aus diesen Ergebnissen eine Responderschwelle von 4–6 ab. Diesem Vorgehen wird nicht gefolgt, da (1) das Ziel der Untersuchung der Nachweis der Änderungssensitivität und nicht die Validierung einer MID war, (2) der Stichprobenumfang sehr

klein ist, (3) der Zeitraum, über den sich die Änderung vollzog, mit nur 1 Woche sehr kurz war und (4) bei dem Vorgehen anzunehmen ist, dass es sich bei den beobachteten Änderungen im TranQoL-Score um periodische Schwankungen, die innerhalb des Krankheitsbildes wiederkehrend auftreten, und nicht um eine klinisch relevante langfristige Veränderung handelt. Zudem sind die gewählten Responderschwellen kleiner als eine halbe Standardabweichung der Baselinewerte. Insgesamt wird der Gesamtscore des TranQoL als ausreichend validiert angesehen, um für die Nutzenbewertung herangezogen zu werden. Die Responderanalysen können aufgrund des Fehlens einer validierten MID allerdings nicht für die Nutzenbewertung berücksichtigt werden. Daher werden Mittelwertdifferenzen basierend auf den im SAP präspezifizierten Modellen (siehe Abschnitt 2.4) für die Nutzenbewertung herangezogen.

SF-36

Der Endpunkt SF-36 wird in der Nutzenbewertung berücksichtigt.

Operationalisierung:

Beschreibung

Es wurde die Version 2 des SF-36 verwendet. Der SF-36-Fragebogen beinhaltet 36 Items, die 8 Dimensionen zugeordnet werden: körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperlicher Schmerz, allgemeine Gesundheitswahrnehmung, Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden sowie zwei Summenskalen (Physical Component Summary (PCS) und Mental Component Summary (MCS)).

Items aller Domänen-Skalen gehen in die beiden Summenskalen ein. Die PCS schließt vor allem Items der Domänen-Skalen für körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperliche Schmerzen und allgemeine Gesundheitswahrnehmung ein. In die MCS gehen vor allem Items der Domänen-Skalen für Vitalität, soziale Funktionsfähigkeit, emotionale Rollenfunktion und psychisches Wohlbefinden ein.

Das Scoring erfolgte mittels QualityMetric's Health Outcomes Scoring Software 5.1 auf Basis der 2009 US-Allgemeinbevölkerungsdaten. Für die berechneten T-Scores spiegeln höhere Werte einen besseren Gesundheitszustand wider. Der Fragebogen galt insgesamt als auswertbar, wenn ≥ 50 % (≥ 18 der 36 Fragen) der Items beantwortet wurden. Der Bezugszeitraum der Standardversion des SF-36 beträgt 4 Wochen.

Im HRQoL-SAP wurden folgende Responderschwellen auf Basis des Manuals des SF-36 präspezifiziert:

- körperliche Funktionsfähigkeit: 4,3 Punkte
- körperliche Rollenfunktion: 4,0 Punkte
- körperlicher Schmerz: 5,5 Punkte
- allgemeine Gesundheitswahrnehmung: 7,0 Punkte
- Vitalität: 6,7 Punkte
- soziale Funktionsfähigkeit: 6,2 Punkte
- emotionale Rollenfunktion: 4,6 Punkte
- psychisches Wohlbefinden: 6,7 Punkte
- PCS: 3,8 Punkte
- MCS: 4,6 Punkte

Im Dossier werden für PCS/MCS Responderanalysen für Verbesserungen mit MID von 5 Punkten bzw. Sensitivitätsanalysen mit 2 (PCS) bzw. 3 (MCS) Punkten durchgeführt. Für die Einzeldomänen werden im Dossier MID von 3 Punkten für die Domänen körperliche Funktionsfähigkeit, körperliche Rollenfunktion, körperliche Schmerzen, soziale Funktionsfähigkeit sowie eine MID von 2 Punkten für die allgemeine Gesundheitswahrnehmung und 4 für die emotionale Rollenfunktion angewendet.

- Erhebungszeitpunkte:
  Der Endpunkt wurde zu folgenden Zeitpunkten erhoben:
  - Screeningphase: innerhalb von 4 Wochen vor der ersten Dosierung
  - Behandlungsphase: Tag 1 der Studienwoche 12, 24, 36 und 48
  - Alle 12 Wochen während der Langzeitbehandlungsphase
  - Keine Erhebung nach Behandlungsabbruch sowie in der offenen Extensionsphase
- Auswertung:
  Zur Auswertung der Daten des SF-36 wird eine „HRQoL evaluable population“ definiert. Diese enthält alle Personen der ITT-Population mit einem auswertbaren Fragebogen zu Baseline und mindestens einem Zeitpunkt nach Baseline. Der Fragebogen galt insgesamt als auswertbar, wenn ≥ 50 % (≥ 18 der 36 Fragen) der Items beantwortet wurden. Es wurden Mittelwertsdifferenzen, inkl. Hedges' g zu Woche 24, Woche 48 und über die gesamte Studienperiode hinweg mittels ANCOVA bzw. MMRM-Modellen berechnet (siehe Abschnitt 2.4). Daneben wurden auch Responderanalysen zu Woche 24 und 48 geplant (siehe Abschnitt 2.4). Als primäre Domäne/Skala von Interesse wurden die Domänen körperliche Funktionsfähigkeit und allgemeine Gesundheitswahrnehmung des SF-36 präspezifiziert. Alle anderen Domänen/Skalen werden explorativ analysiert.

Bewertung

Die Operationalisierung ist nachvollziehbar. Es erfolgte keine Erhebung des SF-36 im Fall eines Behandlungsabbruches. Da anzunehmen ist, dass Personen, die die Behandlung abbrechen, tendenziell eine schlechtere Lebensqualität haben als Personen, die die Behandlung weiterführen, führt die fehlende Erhebung des SF-36 im Fall eines Behandlungsabbruches zur Verzerrung. Responderanalysen wurden nur für Verbesserungen, nicht aber für Verschlechterungen der Lebensqualität durchgeführt.

Patientenrelevanz:

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Der SF-36 ist ein generischer Fragebogen zur Messung der Lebensqualität und wird daher als patientenrelevant angesehen.

Validität:

Der SF-36 ist ein häufig genutzter, valider, generischer Fragebogen zur gesundheitsbezogenen Lebensqualität. Es konnte lediglich eine sehr limitierte indikationsspezifische Validierungsstudie bei Personen mit β-Thalassämie einer persischen Übersetzung des SF-36 identifiziert werden [10]. Diese zeigte eine gute interne Konsistenz und strukturelle Validität [10]. Weitere indikationsspezifische Validierungsstudien sowie Studien zu einer indikationsspezifischen MID liegen nicht vor.

Der pU legt im Dossier und in einem separaten Bericht zur Lebensqualität in der BELIEVE-Studie Responderanalysen mit unterschiedlichen MID (siehe oben) basierend auf dem Manual des SF-36 vor. Die Herleitung der MID innerhalb des Manuals des SF-36 [14] beruht auf einem verteilungsbasierten Verfahren. Die Entwicklergruppe des SF-36v2 schlägt ein Änderungskriterium basierend auf statistischer Signifikanz (p-Wert 0,05) vor, das Ähnlichkeiten mit dem von Jacobson und Truax (1991) vorgeschlagenen „reliable change index“ (RCI) aufweist.

Zur Ermittlung einer Relevanzschwelle sollten jedoch mehrere Methoden verwendet werden, z. B. anker- und verteilungsbasierte Methoden. Im Allgemeinen kann eine Relevanzschwelle von ≥ 5 als eine Annäherung an eine geeignete MID für die Summenskalen des SF-36 angesehen werden. Die gewählte MID sollte ergänzend mit Auswertungen zu verteilungsbasierten Analysen der Baselinewerte (hier 0,5 Standardabweichung) oder mit ankerbasierten Verfahren in der untersuchten Indikation überprüft werden. Die im Studienbericht vorliegenden Baseline-Daten unterstützen die Wahl einer MID von 5 für PCS und MCS.

2.3.4 Sicherheit

Unerwünschte Ereignisse

Der Endpunkt Unerwünschte Ereignisse wird in der Nutzenbewertung berücksichtigt.

Operationalisierung:

Beschreibung

Art, Häufigkeit und Schweregrad aller Unerwünschten Ereignisse unabhängig vom Zusammenhang mit der Studienmedikation wurden mittels MedDRA Version 20.0 kodiert. Die Erhebung des Schweregrades der Unerwünschten Ereignisse erfolgte mittels NCI-CTCAE Version 4.0.

Als Schwerwiegendes UE (SUE) wurden UE unter folgende Bedingungen definiert: Ereignisse, die

- tödlich verliefen
- lebensbedrohlich waren
- eine Hospitalisierung erforderlich machten oder zur Verlängerung eines bestehenden stationären Aufenthalts führten
- in bestehenden oder bedeutsamen Behinderungen/Unfähigkeiten resultierten
- angeborene Anomalien/Geburtsdefekte waren oder
- wichtige medizinische Ereignisse waren, die nicht unmittelbar tödlich oder lebensbedrohlich waren oder eine Hospitalisierung oder Behinderung verursachten, aber die Patientinnen und Patienten diesen Risiken aussetzen oder eine medizinische oder chirurgische Behandlung zur Prävention dieser Ereignisse nötig machten.

Als UE von besonderem Interesse wurden im SAP Malignitäten und prämaligne Erkrankungen definiert. Im Protokoll wurde die Erhebung von Malignitäten und prämalignen Erkrankungen als „event of interest“ definiert. Als weitere UE, welche eine Analyse der Sicherheit erfordern, wurden die SMQ „Embolische und thrombotische Ereignisse“ und UE mit dem PT „Knochenschmerzen“ definiert.

- Erhebungszeitpunkte:
  UE wurden ab Unterzeichnung der Einverständniserklärung bis zu 9 Wochen nach der letzten Dosis kontinuierlich erhoben. Nach Woche 9 wurden nur mit der Studienmedikation in Zusammenhang stehende Unerwünschte Ereignisse erhoben.

- Auswertung:
  Die Auswertung erfolgte auf Basis der Sicherheits-Population. Für die Auswertung wurde der Zeitraum ab der ersten Dosis berücksichtigt. Im SAP wurden lediglich deskriptive Analysen zur Sicherheit geplant. Für das Dossier wurden OR, RR und RD dargestellt, allerdings ohne die Stratifizierungsvariable der Randomisierung (Region) bei der Berechnung zu berücksichtigen.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz:

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Der Endpunkt Unerwünschte Ereignisse wird als patientenrelevant erachtet. Die Patientenrelevanz von Laborparametern ist unklar.

Validität:

Die Erhebung erfolgte vom Einschluss in die Studie bis 9 Wochen nach der letzten Dosis vollständig. Nach Woche 9 nach Beendigung der Studienmedikation erfolgte lediglich eine selektive Erfassung. SUE, bei denen eine Erfassung als SUE aufgrund einer Hospitalisierung erfolgte, werden ebenfalls im Endpunkt Hospitalisierung erfasst. Hierfür liegt somit eine Doppelerfassung vor. Der pU legt eine Sensitivitätsanalyse zu den bis Woche 48 aufgetretenen Gesamtraten von UE ohne Berücksichtigung von Hospitalisierungen vor. Diese Analysen werden nicht als adäquat erachtet, da UE, die zu einer Hospitalisierung führen, ein integraler Bestandteil der Auswertungen zur Sicherheit sind. Es wurden keine Sensitivitätsanalysen zu UE ohne Berücksichtigung von Ereignissen vorgelegt, die auf die Grunderkrankung zurückzuführen sind und im Rahmen der UE miterfasst wurden.

2.3.5 Übersicht der Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 5.

Tabelle 5: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie ACE-536-B-THAL-001 (BELIEVE)

Studienvisite (Woche) Endpunkt	**Base line**	**Studien-woche 12**	**Studien-woche 24**	**Studien-woche 36**	**Studien-woche 48**	**Langzeit-nachbeo-bachtung**	**End of Treatment /Post Treatment Follow-up**
Gesamtüberleben	Erfassung im Rahmen der Sicherheit (siehe UE)						
Hospitalisierung	kontinuierliche Erfassung bis 9 Wochen nach letzter Dosis						
TranQoL	x	x	x	x	x	x[1]	-
SF-36	x	x	x	x	x	x1	-

Studienvisite (Woche) Endpunkt	Base line	Studien-woche 12	Studien-woche 24	Studien-woche 36	Studien-woche 48	Langzeit-nachbeo-bachtung	End of Treatment /Post Treatment Follow-up
Unerwünschte Ereignisse	kontinuierliche, vollständige Erfassung bis 9 Wochen nach letzter Dosis, danach nur unvollständige Erfassung						

[1)] Alle 12 Wochen

Abkürzungen: SF-36: Short-Form-36; TranQoL: Transfusion-dependent Quality of Life Questionnaire; UE: Unerwünschte Ereignisse

2.4 Statistische Methoden

Für die Studie liegen zwei SAP vor, ein genereller vom 19. Juni 2018 und ein spezifischer für die Analysen zur Lebensqualität vom 12. Juni 2018. Beide SAP wurden erst nach dem Datenschnitt am 11. Mai 2018 finalisiert. Bei der BELIEVE-Studie handelt es sich um eine Überlegenheitsstudie. Das Signifikanzniveau wurde auf 0,05 für alle 2-seitigen Tests festgelegt.

Analysepopulationen

Die ITT-Population beinhaltet alle randomisierten Individuen. Personen werden der Gruppe zugeordnet, zu der sie randomisiert wurden (Luspatercept + BSC: N = 224; Placebo + BSC: N = 112).

Die Sicherheitspopulation beinhaltet alle Personen, die randomisiert wurden und mindestens eine Dosis der Studienmedikation erhalten haben. Personen werden der Gruppe zugeordnet, deren Studienmedikation sie erhielten (Luspatercept + BSC: N = 223; Placebo + BSC: N = 109).

Die „HRQoL evaluable Population“ enthält alle Personen der ITT-Population mit einem auswertbaren Fragebogen zu Baseline und mindestens einem Zeitpunkt nach Baseline. Sie wurde für jeden Fragebogen separat festgelegt: Der TranQoL-Fragebogen galt als auswertbar, wenn mindestens 75 % aller Items beantwortet wurden (≥ 27 der 36 Fragen) (Luspatercept + BSC: N = 211; Placebo + BSC: N = 103). Der SF-36 galt als auswertbar, wenn mindestens 50 % aller Items ausgefüllt wurden (≥ 18 der 36 Fragen) (Luspatercept + BSC: N = 210; Placebo + BSC: N = 103). Alle Analysen zur Wirksamkeit mit Ausnahme derjenigen zur Lebensqualität werden auf Basis der ITT-Population durchgeführt. Die Analysen zur Sicherheit werden in der Sicherheitspopulation und die zur Lebensqualität in der „HRQoL evaluable Population“ durchgeführt.

Statistische Analysen

Gesamtmortalität

Die Analyse der Gesamtmortalität erfolgte deskriptiv.

Hospitalisierung

Für die Auswertung der Daten zur Hospitalisierung waren im SAP lediglich deskriptive Analysen der Häufigkeiten geplant. Für das Dossier wurden vom pU zusätzlich Unterschiede zwischen den Behandlungsarmen mittels CMH-Methode stratifiziert nach Region (Stratifizierungsvariable der Randomisierung) auf Basis von OR berechnet. RR und RD wurden unstratifiziert mittels Normal-

Approximation berechnet. Unstratifizierte RR und RD in einer stratifiziert randomisierten Studie werden als nicht adäquat angesehen.

Lebensqualität

Laut SAP war geplant, für jeden Domänenscore sowie für die gebildeten Summenscores die Veränderung zu Baseline als Least-Square-Means für jede geplante Erhebung mittels ANCOVA-Modellen mit Adjustierung für den Baselinewert und der Stratifizierungsvariable der Randomisierung Region zu berechnen. Der primäre Vergleich soll zu Woche 24 erfolgen. Die Analysen zu Woche 48 werden als explorativ angesehen. Für das Dossier wurden post-hoc Veränderungen zu Baseline als Least-Square-Means für jede geplante Erhebung mittels ANOVA-Modellen berechnet. Angaben zu Adjustierung und Stratifizierung in diesen Modellen liegen nicht vor. Für die Nutzenbewertung wird auf die Ergebnisse der präspezifizierten Analysen mittels ANCOVA zurückgegriffen.

Zum Vergleich der Lebensqualität über die gesamte Studienperiode hinweg wurden Mixed-model repeated Measure (MMRM)-Modelle angewendet. Diese Modelle sind laut SAP folgendermaßen spezifiziert. Sie nutzen die restricted Maximum Likelihood (REML)-Methode unter der Annahme von „Missing at random“. Als abhängige Variable wird die Änderung des jeweiligen Scores zu Baseline verwendet. Die Modelle enthalten folgende Kovariablen: Behandlungsgruppe, Zeit (kontinuierlich als Anzahl an Wochen seit Einnahme der 1. Dosis), Stratifizierungsvariable Region, Baselinewert, Interaktionsterm Baselinewert x Studienvisite und Interaktionsterm Behandlungsgruppe x Studienvisite. Intercept und Zeit wurden als zufällige Effekte modelliert, alle anderen Kovariablen als feste Effekte. Die Modelle wurden genutzt, um LS-Means für den Gruppenunterschied zwischen Luspatercept und Placebo für jede Studienvisite einzeln und über die gesamte Behandlungszeit hinweg zu berechnen. Neben den präspezifizierten Analysen wurden post-hoc für das Dossier leicht abweichende MMRM-Modelle angewendet. In diesen fehlt die Kovariable Stratifizierungsvariable Region und der Interaktionsterm Behandlungsgruppe x Studienvisite. Eine Begründung für die Änderung der MMRM-Modelle wird nicht gegeben. Für die Nutzenbewertung wird auf die Ergebnisse der präspezifizierten Analysen zurückgegriffen.

Laut SAP waren Responderanalysen für eine Verbesserung um präspezifizierte MID (siehe Abschnitt 2.3.3) mittels Cochran-Mantel-Haenszel (CMH)-Test mit modifizierten Ridit-Scores stratifiziert für die Region (Stratifizierungsfaktor der Randomisierung) geplant. Die Homogenität der OR innerhalb der Strata sollte über den Breslow-Day-Test überprüft werden. Bei Verletzung der Homogenitätsannahme sollten stratifizierte Ergebnisse dargestellt werden. Für die Analysen war die Berechnung von Odds ratios (OR) mittels CMH-Test präspezifiziert. Für das Dossier wurden post-hoc OR, relative Risiken (RR) und Risikodifferenzen (RD) berechnet. Die Berechnung der OR erfolgte, anders als im SAP präspezifiziert, mittels logistischer Regression. RR und RD wurden mittels Normal-Approximation berechnet. Detailliertere Angaben zur Berechnung der Effektschätzer und p-Werte sind aus den Ergebnisoutputs der Post-hoc-Analysen nicht zu entnehmen. Ein SAP für die Post-hoc-Analysen wurde nicht vorgelegt. Angaben zur Analyse im Modul 4 sind teils widersprüchlich. So wird im Modul 4 angegeben, dass die OR mittels CMH-Test berechnet wurden, im Ergebnisoutput wird aber eine Berechnung mittels logistischer Regression aufgeführt. Insgesamt sind die post-hoc durchgeführten Responderanalysen damit nicht nachvollziehbar beschrieben. Es ist insbesondere unklar, ob die in einer stratifiziert, randomisierten Studie adäquate Stratifizierung der Analyse erfolgte. Die Effektschätzer und p-Werte der post-hoc durchgeführten Responderanalysen werden daher in der Nutzenbewertung nicht berücksichtigt.

Als Sensitivitätsanalysen waren für die Lebensqualität die Analyse von Dropouts und Imputation von fehlenden Werten mit Pattern-mixture-Modellen sowie die explorative Analyse zur Identifizierung von Determinanten der Lebensqualität vorgesehen.

Sicherheit

Für die Auswertung der Daten zur Sicherheit waren im SAP lediglich deskriptive Analysen der Häufigkeiten geplant. Für das Dossier wurden vom pU zusätzlich Unterschiede zwischen den Behandlungsarmen mittels CMH-Methode berechnet und OR, RR und RD inkl. 95%-KI dargestellt. Zusätzlich wird im Dossier angegeben, dass die OR, RR und RD unstratifiziert mittels Normal-Approximation berechnet wurden. Dies wird in einer stratifiziert, randomisierten Studie als inadäquat für die Nutzenbewertung angesehen.

Subgruppenanalysen

Siehe Tabelle 2

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 6: Verzerrungspotential der Studie ACE-536-B-THAL-001 (BELIEVE)

Studie	Adäquate Erzeugung der Randomisierungs-sequenz	Verdeckung der Gruppenzuteilung	Verblindung Patient/in	Verblindung Behandler	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
BELIEVE	Ja[1)]	Ja[1)]	Ja[2)]	Ja[2, 3)]	Unklar[4, 5)]	Nein	Niedrig

[1)] Die Randomisierung und Verdeckung der Gruppenzuteilung erfolgte durch ein Interactive Response Technology-System. Die Randomisierung erfolgte erst nach Abschluss der Screeningphase bei Erfüllung aller Einschlusskriterien. Die Randomisierung wurde nach Region stratifiziert und im Verhältnis 2:1 Intervention : Kontrolle durchgeführt. Das Verfahren zur Generierung der Randomisierungssequenz wurde nicht näher beschrieben. Die Charakteristika der Studienteilnehmenden zu Baseline sind weitestgehend gleichmäßig über die Behandlungsarme verteilt.

[2)] Studienteilnehmende, ärztliches Personal und Sponsor waren verblindet. Apothekerinnen und Apotheker in den Studienzentren, welche die Injektionen vorbereiteten, waren über die Intervention informiert. Spezifische Nebenwirkungen, die zur unbeabsichtigten Entblindung hätten führen können, wurden nicht beobachtet.

[3)] Es erfolgte keine Verblindung der Behandler gegenüber dem primären Endpunkt, der die Transfusionslast widerspiegelt. Zudem waren Dosiserhöhungen bei Nichterreichen einer 33 % Verminderung der Transfusionslast über mindestens 2 Dosiszyklen (~6 Wochen) im Vergleich zu Baseline geplant oder bei einer Verminderung der Transfusionslast um ≥ 33 %, aber < 50 % über mindestens 2 Dosiszyklen (~6 Wochen) im Vergleich zu Baseline im Ermessen des Prüfpersonals möglich. Aufgrund des genauen Monitorings der Transfusionslast durch den Behandler und der damit möglichen Schlussfolgerungen über die Wirksamkeit der Intervention ist es unklar, ob die Verblindung vollständig aufrechterhalten werden konnte.

[4)] Die SAP wurden am 19. Juni 2018 (SAP) bzw. 12. Juni 2018 (SAP HRQoL), also nach dem Datenschnitt am 11. Mai 2018 finalisiert.

[5)] Es konnte kein SAP für die Post-hoc-Analysen identifiziert werden. Es ist unklar, warum die Methodik der Post-hoc-Analysen von den primär geplanten Analysen bei der Berechnung der relativen Effektmaße hinsichtlich der Berücksichtigung der Stratifizierungsfaktoren, Adjustierungsvariablen und Analyseverfahren abweicht.

Abkürzungen: SAP: statistischer Analyseplan; SAP HRQoL: statistischer Analyseplan für die Analysen der Lebensqualität

Das Verzerrungspotential auf Studienebene wird als niedrig eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 7 dargestellt.

Tabelle 7: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie ACE-536-B-THAL-001 (BELIEVE) für den Datenschnitt vom 11.05.2018

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Gesamtmortalität	Ja	Ja[1)]	Nein	Nein	Niedrig
Hospitalisierung	Ja	Ja	Nein	Ja[2)]	Hoch
TranQoL	Ja	Nein[3)]	Nein	Ja[4)]	Hoch
SF-36	Ja	Nein[3)]	Unklar[5)]	Ja[4)]	Hoch
UE	Ja	Ja[6)]	Nein	Nein	Niedrig

[1)] Todesfälle wurden im Rahmen der Sicherheit innerhalb der Sicherheitspopulation (Luspatercept + BSC: n = 223, Placebo + BSC: n = 109) erhoben.
[2)] Es sind Einschränkungen der internen Validität aufgrund regionaler Unterschiede bei der Hospitalisierung möglich. Die Stratifizierung der Randomisierung erfolgte nicht auf Ebene der Studienzentren oder Länder, sondern auf Ebene geographischer Regionen. Eine Ungleichverteilung der Testpersonen über Länder mit verschiedenen Kriterien der Notwendigkeit einer Hospitalisierung kann zu einer Verzerrung der Ergebnisse führen.
[3)] Die Analyse beruht auf der HRQoL evaluable Population (Luspatercept + BSC: n = 211 (TranQoL) / n = 210 (SF-36), Placebo + BSC: n = 103)). Diese umfasst alle Personen der ITT-Population mit einem auswertbaren Fragebogen zu Baseline und mindestens einem Zeitpunkt nach Baseline und damit nicht die gesamte ITT-Population. Zudem erfolgte keine Erhebung des Fragebogens bei Beendigung der Studienmedikation.
[4)] Hohe Anzahl an fehlenden Werten. Zur Studienwoche 48 lag der Anteil an fehlenden Werten bei knapp über 20 % in beiden Studienarmen bezogen auf die ITT-Population.
[5)] Der pU verwendet nicht das im SAP präspezifizierte Verfahren, um für die Nutzenbewertung relevante, post-hoc Responderanalysen mit den beim G-BA etablierten Responderschwellen durchzuführen. Es ist unklar, warum davon abgewichen wurde.
[6)] Die Erhebung erfolgte in der Sicherheitspopulation (Luspatercept + BSC: n = 223, Placebo + BSC: n = 109).

Abkürzungen: BSC: Best Supportive Care; ITT: Intention to treat; SAP: statistischer Analyseplan; SF-36: Short-Form-36; TranQoL: Transfusion-dependent Quality of Life Questionnaire

3 Ergebnisse der eingeschlossenen Studie

3.1 Studiencharakteristika und Studienmedikation

Allgemeine Angaben zur Studie BELIEVE mit Datenschnitt vom 11.05.2018 können Tabelle 8 entnommen werden.

Tabelle 8: Allgemeine Angaben zur Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt: 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	**Luspatercept + BSC (N = 224) n (%)**	**Placebo + BSC (N = 112) n (%)**
Behandlung erhalten	223 (99,6)	109 (97,3)
Behandlung für 24 Wochen erhalten	210 (93,8)	102 (91,1)
Behandlung für 48 Wochen erhalten	200 (89,3)	96 (85,7)
HRQoL evaluable Population[1)] TranQoL SF-36	 211 (94,2) 210 (93,8)	 103 (92) 103 (92)
Behandlung laufend	181 (80,8)	85 (75,9)
Abbruch der Einnahme der Studienmedikation aufgrund von UE fehlender Wirksamkeitsnachweis Entscheidung des Arztes Lost to Follow-up Tod Protokollverletzung Entzug der Einverständniserklärung andere[2)]	42 (18,8) 10 (4,5) 2 (0,9) 0 0 0 1 (0,4) 26 (11,6) 3 (1,3)	24 (21,4) 1 (0,9) 8 (7,1) 0 0 0 0 12 (10,7) 3 (2,7)
Abbruch der Studienteilnahme Tod aufgrund von UE Entzug der Einverständniserklärung andere[3)]	26 (11,6) 1 (0,4) 4 (1,8) 13 (5,8) 2 (0,9)	14 (12,5) 1 (0,9) 0 6 (5,4) 1 (0,9)
Eintritt in Langzeitbehandlungsphase	k. A.	k. A.
Beendigung der Langzeitbehandlungsphase	k. A.	k. A.
Eintritt in offene Extensionsphase	k. A.	k. A.
Eintritt in Post-Treatment-Follow-up-Phase	k. A.	k. A.
Mediane Behandlungsdauer Wochen (min; max)	n = 223 64,1 (3; 97)	n = 109 64,0 (9; 92)
Mediane Beobachtungsdauer Wochen (min; max)	k. A.	k. A.

[1)] Die HRQoL evaluable Population enthält alle Personen der ITT-Population mit einem auswertbaren Fragebogen zu Baseline und mindestens einen Zeitpunkt nach Baseline.

[2)] Andere Gründe für den Abbruch der Behandlung im Luspatercept + BSC-Behandlungsarm waren: Interesse an anderen klinischen Studien, Planung einer Schwangerschaft, Wohnungswechsel und persönliche Gründe. Andere Gründe für den Abbruch der Behandlung im Placebo + BSC-Behandlungsarm waren: persönliche Gründe (n = 2) und der Wunsch, nicht in die Langzeitbehandlungsphase einzutreten.

[3)] Andere Gründe für den Studienabbruch im Luspatercept + BSC-Behandlungsarm waren: persönliche Gründe und Umzug in eine andere Stadt. Der „andere“ Grund für den Abbruch der Studie im Placebo + BSC-Behandlungsarm war: persönlicher Grund.

Insgesamt wurden 336 Personen innerhalb der Studie randomisiert: 224 in den Luspatercept + BSC-Behandlungsarm und 112 in den Placebo + BSC-Behandlungsarm. Davon erhielten 223 (Luspatercept + BSC) bzw. 109 (Placebo + BSC) die Intervention und sind somit Bestandteil der Sicherheitspopulation. Zum Zeitpunkt des Datenschnittes vom 11. Mai 2018 erhielten noch 80,8 % der Personen im Luspatercept + BSC-Arm und 75,9 % im Placebo + BSC-Arm die Intervention. Die mediane Behandlungsdauer betrug 64 Wochen in beiden Behandlungsarmen. Gründe für den Abbruch der Einnahme der Studienmedikation waren Entzug der Einverständniserklärung (Luspatercept + BSC: 11,6 %, Placebo + BSC: 10,7 %), UE (Luspatercept + BSC: 4,5 %, Placebo + BSC: 0,9 %), fehlender Wirksamkeitsnachweis (Luspatercept + BSC: 0,9 %, Placebo + BSC: 7,1 %), Protokollverletzungen (Luspatercept + BSC: 0,4 %, Placebo + BSC: 0) sowie andere Ursachen (Luspatercept + BSC: 1,3 %, Placebo + BSC: 2,7 %). Zum Zeitpunkt des Datenschnittes hatten 11,6 % (Luspatercept + BSC: n = 26) bzw. 12,5 % (Placebo + BSC: n = 14) der Teilnehmenden die Studie vorzeitig beendet: 5,8 % (Luspatercept + BSC: n = 13) bzw. 5,4 % (Placebo + BSC: n = 6) aufgrund des Entzugs der Einverständniserklärung, 1,8 % (Luspatercept + BSC: n = 4) bzw. 0 (Placebo + BSC) aufgrund von UE, 0,4 % (Luspatercept + BSC: n = 1) bzw. 0,9 % (Placebo + BSC: n = 1) aufgrund von Tod und jeweils 0,9 % pro Studienarm (Luspatercept + BSC: n = 2, Placebo + BSC: n = 1) aus anderen Gründen. Unklar ist, warum die Aufsummierung der Anzahl an Personen pro Abbruchgrund (Luspatercept + BSC: n = 20, Placebo + BSC: n = 8) kleiner ist als die Anzahl an Personen, welche die Studie abgebrochen haben (Luspatercept + BSC: n = 26, Placebo + BSC: n = 14).

Insgesamt 19 Personen (8,5 %) im Luspatercept + BSC-Arm und 5 Personen (4,5 %) im Placebo + BSC-Arm hatten mindestens eine Protokollverletzung. Diese betraf das Berichten der Sicherheit (Luspatercept n = 8, Placebo n = 0), Nichterfüllen der Einschlusskriterien (Luspatercept n = 4, Placebo n = 1), Labortest/Prozeduren (Luspatercept n = 2, Placebo n = 3), Aspekte des Prüfmedikaments, z. B. Dosierungsfehler (Luspatercept n = 3, Placebo n = 1), Zeitplan der Studienvisiten (Luspatercept n = 1, Placebo n = 1), Begleitmedikation und/oder Prozeduren (Luspatercept n = 1, Placebo n = 0) und andere (Luspatercept n = 2, Placebo n = 0).

Die Charakteristika der Studienpopulation der Studie BELIEVE sind in Tabelle 9 zu finden. Das Alter der Studienteilnehmenden lag in beiden Behandlungsarmen im Median bei 30 Jahren. Lediglich 7,6 % (Luspatercept + BSC) bzw. 4,5 % (Placebo + BSC) der Studienteilnehmenden waren älter als 50 Jahre. Der Anteil an Frauen an der Studienpopulation betrug 58,9 % (Luspatercept + BSC) bzw. 56,3 % (Placebo + BSC). Ca. 30 % der Studienteilnehmenden (Luspatercept + BSC: 30,4 %, Placebo + BSC: 31,3 %) hatten einen β_0/β_0-Genotyp, ca. 70 % (Luspatercept + BSC: 69,2 %, Placebo + BSC: 68,8 %) einen Nicht-β_0/β_0-Genotyp. Bei über 70 % (Luspatercept + BSC: 77,7 %, Placebo + BSC: 74,1 %) der Studienpopulation lag eine reine β-Thalassämie vor, 13,8 % (Luspatercept + BSC) bzw.18,8 % (Placebo + BSC) wiesen eine Hämoglobin E / β-Thalassämie auf und 8,0 % (Luspatercept + BSC) bzw. 7,1 % (Placebo + BSC) eine β-Thalassämie kombiniert mit einer α-Thalassämie. Die Charakteristika der Studienpopulation (siehe Tabelle 9) sind weitestgehend gleichmäßig über beide Behandlungsarme verteilt.

Tabelle 9: Charakterisierung der Studienpopulation ACE-536-B-THAL-001 (BELIEVE) zu Baseline (Datenschnitt 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	**Luspatercept + BSC N = 224**	**Placebo + BSC N = 112**
Alter, Jahre		
Mittelwert (SD)	32,2 (10,7)	31,9 (9,9)
Median (min; max)	30 (18; 66)	30 (18; 59)
Altersgruppe, n (%)		
≤ 32 Jahre	129 (57,6)	63 (56,3)
> 32 bis ≤ 50 Jahre	78 (34,8)	44 (39,3)
> 50 Jahre	17 (7,6)	5 (4,5)
Geschlecht, n (%)		
Männlich	92 (41,1)	49 (43,8)
Weiblich	132 (58,9)	63 (56,3)
Abstammung, n (%)		
Europäisch	122 (54,5)	60 (53,6)
Asiatisch	81 (36,2)	36 (32,1)
Afrikanisch	1 (0,4)	0
Andere	15 (6,7)	11 (9,8)
Fehlend	5 (2,2)	5 (4,5)
Region, n (%)		
Europa und Nordamerika	100 (44,6)	51 (45,5)
Mittlerer Osten und Nordafrika	52 (23,2)	26 (23,2)
Asien/Pazifik	72 (32,1)	35 (31,3)
β-Thalassämie-Diagnose, n (%)		
β-Thalassämie	174 (77,7)	83 (74,1)
Hämoglobin E / β-Thalassämie	31 (13,8)	21 (18,8)
β-Thalassämie kombiniert mit α-Thalassämie	18 (8,0)	8 (7,1)
Fehlend	1 (0,4)	0
Alter, an dem mit regelmäßigen Transfusionen begonnen wurde, Jahre	n = 169	n = 85
Median (min; max)	2 (0; 52)	2 (0; 51)
β-Thalassämie-Genmutation, n (%)		
β_0/β_0	68 (30,4)	35 (31,3)
Nicht-β_0/β_0	155 (69,2)	77 (68,8)
Fehlend	1 (0,4)	0
ECOG-Status, n (%)		
0	176 (78,6)	91 (81,3)
1	48 (21,4)	20 (17,9)
Fehlend	0	1 (0,9)
Splenektomie, n (%)		
Ja	129 (57,6)	65 (58,0)
Transfusionslast zu Baseline, EK-Einheiten / 24 Wochen[1)]		
Median (min, max)	14,0 (6,0; 24,0)	15,0 (6,0; 26,0)
Hb-Wert vor der Transfusion, g/dL[2)]		
Median (min, max)	9,31 (4,5; 11,4)	9,15 (5,8; 11,7)

ACE-536-B-THAL-001 (BELIEVE)	Luspatercept + BSC N = 224	Placebo + BSC N = 112
Lebereisengehalt, mg/g Trockengewicht		
Median (min, max)	6,14 (0,8; 125,0)	5,05 (0,2; 53,2)
Myokardialer Eisengehalt, mg/g Trockengewicht		
Median (min, max)	0,6 (0; 12)	0,6 (0; 5)
Eisenchelattherapie zu Studienbeginn, n (%)[3]	n = 223	n = 109
Ja	217 (97,3)	106 (97,2)
β-Thalassämie-Komorbiditäten, n (%)[3,4]	n = 223	n = 109
Eisenüberladung[5]	209 (93,7)	104 (95,4)
Osteopenie	123 (55,2)	68 (62,4)
Osteoporose	106 (47,5)	45 (41,3)
Knochenschmerzen	36 (16,1)	12 (11,0)
Splenektomie	128 (57,4)	62 (56,9)
Cholezystektomie	53 (23,8)	25 (22,9)
Splenomegalie	102 (45,7)	44 (40,4)
Hypogonadismus	74 (33,2)	36 (33,0)
Hypothyroidismus	40 (17,9)	17 (15,6)
Cholelithiasis	45 (20,2)	23 (21,1)
Hepatitis C	38 (17,0)	14 (12,8)
Frakturen	27 (12,1)	10 (9,2)
Auditive Störung	23 (10,3)	9 (8,3)
Nephrolithiasis	23 (10,3)	12 (11,0)

[1] Berechnet aus Daten von 12 Wochen vor Studienbeginn (retrospektive Erfassung) und den 12 Wochen der Screening/Run-in-Phase.
[2] Mittelwert aller in den 24 Wochen vor der ersten Dosis dokumentierten Hb-Werte, die vor den jeweiligen Transfusionen bestimmt wurden.
[3] Angaben liegen nur für die Sicherheitspopulation vor.
[4] Häufigkeit > 10 % in einem der beiden Arme
[5] Definiert als Serum-Ferritin > 1000 mg/L oder Lebereisenkonzentration > 7mg/g Trockengewicht

Abkürzungen: BSC: Best Supportive Care; ECOG: Eastern Cooperative Oncology Group; EK: Erythrozytenkonzentrat; SD: Standardabweichung

Exposition mit der Studienmedikation

Die wichtigsten Angaben zur Exposition mit der Studienmedikation innerhalb der Studie BELIEVE sind in Tabelle 10 zusammengefasst.

Tabelle 10: Angaben zur Exposition mit der Studienmedikation innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt: 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	Luspatercept + BSC N = 223	Placebo + BSC N = 109
Dauer der Exposition, Wochen[1]		
Mittelwert (SD)	62,0 (16,83)	60,1 (16,83)
Median (min; max)	64,1 (3; 97)	64,0 (9; 92)
Anzahl an erhaltenen Dosen		
Mittelwert (SD)	19,8 (5,51)	19,6 (5,58)
Median (min; max)	21 (1; 32)	21 (3; 30)

ACE-536-B-THAL-001 (BELIEVE)	Luspatercept + BSC N = 223	Placebo + BSC N = 109
Dosisanpassungen		
Unterbrechung, n (%)	137 (61,4)	45 (41,3)
Erhöhung, n (%)[2]	107 (48,0)	74 (67,9)
Reduktion, n (%)[3]	29 (13,0)	4 (3,7)

[1)] Definiert als (Datum Behandlungsende - Datum der ersten Dosis)/7. Das Datum des Behandlungsendes entspricht dem Datum der letzten Dosis + 20 Tage oder dem Sterbedatum, falls dieses weniger als 20 Tage nach der letzten Dosis lag.
[2)] Eine Erhöhung der Dosis auf 1,25 mg/kg erfolgte bei 103 Personen (46,2 %) im Luspatercept + BSC-Arm und bei 72 Personen (66,1 %) im Placebo + BSC-Arm.
[3)] Bei 25 Personen (11,2 %) im Luspatercept + BSC-Arm erfolgte eine Dosisreduktion auf 0,80 mg/kg, bei 4 Personen (1,8 %) auf 0,60 mg/kg. Bei 3 Personen (2,8 %) im Placebo + BSC-Arm erfolgte eine Dosisreduktion auf 0,80 mg/kg, bei einer Person (0,9 %) auf 0,60 mg/kg. Keine Person erhielt eine Dosis von 0,45 mg/kg.

Abkürzungen: BSC: Best Supportive Care; SD: Standardabweichung

Eine bedarfsgerechte Transfusionstherapie mit EK in Kombination mit einer Chelattherapie ist die Basis der Therapie der transfusionsabhängigen β-Thalassämie. Angaben zur Therapie mit EK und Eisenchelattherapie sind in Tabelle 11 zusammengefasst. Angaben zur Anzahl an transfundierten EK-Einheiten/Zeiteinheit für den gesamten Zeitraum bis zum Datenschnitt und für alle Studienteilnehmenden konnten in den Studienunterlagen nicht identifiziert werden.

Tabelle 11: Angaben zur Therapie mit Erythrozytenkonzentraten und Eisenchelatoren innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt: 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	Luspatercept + BSC N = 223	Placebo + BSC N = 109
Anzahl EK-Einheiten, Woche 1–48	n = 200	nN = 96
Mittelwert (SD)	24,1 (8,5)	30,2 (7,4)
Median (min; max)	24,0 (0,0; 58,0)	31,5 (8,0; 45,0)
Anzahl EK-Einheiten/Zeiteinheit, gesamte Zeit bis zum Datenschnitt		
Mittelwert (SD)	k. A.	k. A.
Median (min; max)	k. A.	k. A.
Patientinnen und Patienten mit mind. 1 Eisenchelattherapie, n (%)	222 (99,6)	109 (100,0)
Deferasirox	145 (65,0)	71 (65,1)
Deferiprone	97 (43,5)	39 (35,8)
Deferoxamin Mesilat / Deferoxamin	72 (32,3)	40 (36,7)

Abkürzungen: BSC: Best Supportive Care; EK: Erythrozytenkonzentrat; SD: Standardabweichung

Begleitmedikation

Angaben zur verabreichten Begleitmedikation sind der Tabelle 12 zu entnehmen.

Tabelle 12: Begleitmedikation mit einer Häufigkeit von ≥ 10 % in der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE) **ATC2 Level** Preferred Term	**Luspatercept + BSC** **N = 223** **n (%)**	**Placebo + BSC** **N = 109** **n (%)**
Patientinnen und Patienten mit mind. einer Begleitmedikation[1)]	*213 (95,5)*	*105 (96,3)*
Anilide	127 (57,0)	62 (56,9)
Paracetamol	123 (55,2)	59 (54,1)
Vitamin D und Vitamin-D-Analoga	85 (38,1)	46 (42,2)
Colecalciferol	40 (17,9)	26 (23,9)
Vitamin D nicht spezifiziert	35 (15,7)	10 (9,2)
Propionsäurederivate	62 (27,8)	30 (27,5)
Ibuprofen	40 (17,9)	15 (13,8)
Kombinationen von Penicillinen, inkl. Beta-Lactamase-Inhibitoren	58 (26,0)	28 (25,7)
Augmentin	51 (22,9)	25 (22,9)
Thrombozytenaggregationshemmer	45 (20,2)	15 (13,8)
Acetylsalicylsäure	44 (19,7)	15 (13,8)
Folsäure und Derivate	43 (19,3)	22 (20,2)
Folsäure	43 (19,3)	22 (20,2)
Substituierte Alkylamine	40 (17,9)	14 (12,8)
Chlorphenamin	21 (9,4)	11 (10,1)
Chlorphenaminmaleat	20 (9,0)	3 (2,8)
Glukokortikoide	39 (17,5)	11 (10,1)
Andere Antihistaminika zur systemischen Anwendung	38 (17,0)	19 (17,4)
Loratadin	27 (12,1)	15 (13,8)
Calcium	34 (15,2)	18 (16,5)
Calciumcarbonat	21 (9,4)	15 (13,8)
Protonenpumpenhemmer	34 (15,2)	17 (15,6)
Penicilline mit erweitertem Wirkungsspektrum	32 (14,3)	11 (10,1)
Amoxicillin	26 (11,7)	9 (8,3)
Makrolide	27 (12,1)	7 (6,4)
Sulfonamide	26 (11,7)	14 (12,8)
Furosemid	26 (11,7)	14 (12,8)
Essigsäure-Derivate und verwandte Substanzen	23 (10,3)	9 (8,3)
Ascorbinsäure (Vitamin C), rein	20 (9,0)	18 (16,5)
Ascorbinsäure	20 (9,0)	17 (15,6)

[1)] Im Folgenden sind ATC-Klassen und Einzelsubstanzen mit einer Häufigkeit von ≥ 10 % in einem Arm aufgeführt.

Abkürzungen: ATC2: Anatomisch-Therapeutisch-Chemisch – Klassifikation; BSC: Best Supportive Care

3.2 Mortalität

Gesamtüberleben

Innerhalb der BELIEVE-Studie traten bis zum Datenschnitt vom 11.05.2018 zwei Todesfälle auf, einer im Luspatercept-Arm, einer im Placebo-Arm (Tabelle 13).

Tabelle 13: Gesamtüberleben in der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	Luspatercept + BSC N = 223	Placebo + BSC N = 109
Tod, n (%)	1 (0,4)	1 (0,9)

Abkürzungen: BSC: Best Supportive Care

Aufgrund der geringen Anzahl an Events sind Subgruppenanalysen nicht möglich.

3.3 Morbidität

Die Ergebnisse zum Morbiditätsendpunkt Hospitalisierungen sind in Tabelle 14 dargestellt.

Hospitalisierung

Tabelle 14: Anzahl Personen mit Hospitalisierungen während der verblindeten Behandlungs- und Langzeitbehandlungsphase in der ITT-Population der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	Luspatercept + BSC N = 224 n (%)	Placebo + BSC N = 112 n (%)	RR [95%-KI][1)] p-Wert
Anzahl an Personen mit Hospitalisierung	42 (18,8)	4 (3,6)	k. A.
Hospitalisierungsgründe[2)]			
Unerwünschte Ereignisse	35 (15,6)	4 (3,6)	-
Transfusion	11 (4,9)	1 (0,9)	-
Nicht protokollbedingte Behandlung oder Untersuchung	9 (4,0)	0	-
Behandlung, die vor Unterzeichnung der Einverständniserklärung geplant war	1 (0,4)	0	-
Elektive Behandlung einer vorbestehenden Erkrankung	9 (4,0)	2 (1,8)	-
Soziale, technische oder praktische Gründe in Abwesenheit eines UE	3 (1,3)	0	-
Anzahl an Personen mit Hospitalisierung aufgrund von UE oder Transfusion[1)]	k. A.	k. A.	k. A.

[1)] Die vom pU im Dossier angegebenen Auswertungen berücksichtigen nur die Behandlungsphase von Woche 1–48 und damit nicht die gesamte verblindete Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase. Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode aufgrund der längeren Behandlungs- und

Beobachtungszeit bevorzugt. Die Effektschätzer auf Basis der Behandlungsphase werden daher nicht dargestellt. Zudem wurden die RR unstratifiziert berechnet, d. h. ohne Berücksichtigung der Stratifizierungsvariable Region, welche für die Randomisierung verwendet wurde.

[2)] Die Aufsummierung der Hospitalisierungsgründe übersteigt die Anzahl an Personen mit Hospitalisierung. Es ist unklar, ob dies darauf zurückzuführen ist, dass eine Person zu einem Zeitpunkt wegen mehrerer Gründe hospitalisiert werden konnte oder ob Mehrfachhospitalisierungen der gleichen Person gezählt wurden oder ob beides möglich war.

Abkürzungen: BSC: Best Supportive Care; k. A.: keine Angabe; KI: Konfidenzintervall; RR: Relatives Risiko / Risk Ratio; UE: Unerwünschtes Ereignis

Der pU legte keine Subgruppenanalysen bezogen auf die gesamte verblindete Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase vor.

3.4 Lebensqualität

TranQoL

Tabelle 15: Rücklaufquote TranQoL bezogen auf die ITT-Population der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	Luspatercept + BSC N = 224		Placebo + BSC N = 112	
Zeitpunkt	**Personen mit auswertbarem Fragebogen**	**Rücklauf-quote (%)[1)]**	**Personen mit auswertbarem Fragebogen**	**Rücklauf-quote (%)[1)]**
Baseline[2)]	222	99,1	112	100
Woche 12	197	87,9	94	83,9
Woche 24	200	89,3	94	83,9
Woche 36	188	83,9	90	80,4
Woche 48	179	79,9	88	78,6
Woche 60	114	50,1	50	44,6
Woche 72	35	15,6	18	16,1
Woche 84	8	3,5	3	2,7
Woche 96	3	1,3	2	1,8
Woche 108	2	0,9	2	1,8
Woche 120	1	0,4	2	1,8
Woche 132	1	0,4	2	1,8
Woche 144	1	0,4	2	1,8

[1)] Die Rücklaufquote wurde bezogen auf die ITT-Population berechnet.
[2)] Die Erhebung zu Baseline erfolgte innerhalb von 4 Wochen vor der ersten Dosierung der Studienmedikation.

Abkürzungen: BSC: Best Supportive Care

Zu allen Studienvisiten bis Woche 48 lag die Rücklaufquote des TranQoL bezogen auf die ITT-Population über 70 % (Tabelle 15). Zu Woche 48 lag die Rücklaufquote in beiden Studienarmen bei knapp unter 80 %, anschließend fiel die Rücklaufquote auf deutlich unter 70 %. Es wurden keine signifikanten Unterschiede in den mittleren Veränderungen des TranQoL-Gesamtscores bis Woche 48 beobachtet (Tabelle 16). Gleiches gilt für die Scores der einzelnen Domänen (Ergebnisse nicht dargestellt).

Tabelle 16: Mittlere Veränderung des TranQoL in der HRQoL evaluable Population der Studie ACE-536-B-THAL-001 (BELIEVE) zu Woche 48 (Datenschnitt: 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	**Luspatercept + BSC N = 211**	**Placebo + BSC N = 103**
TranQoL-Gesamtscore[1)]		
Baseline Mittelwert (SD)	n = 211 66,9 (15,2)	n = 103 70,4 (14,0)
Mittlere Veränderung zwischen Baseline und Woche 48[2)] LS Mean [95%-KI]	n = 179 -0,2 [-2,1; 1,7]	n = 88 0,4 [-2,3; 3,1]
Mittelwertdifferenz zu Woche 48[1)] [95%-KI], p-Wert	-0,6 [-3,8; 2,7]; 0,73	
Mittlere Veränderung von Baseline zu Woche 48[3)] LS Mean [95%-KI]	n = 179 k. A.	n = 88 k. A.
Mittelwertdifferenz über den Studienverlauf bis Woche 48[3)] [95%-KI], p-Wert	k. A., n. s.	

[1)] Der TranQoL-Gesamtscore kann Werte zwischen 0 und 100 annehmen. Ein höherer Wert repräsentiert eine höhere Lebensqualität.
[2)] Analyse basierend auf dem präspezifizierten ANCOVA-(analysis of covariance)-Modell mit Adjustierung für den Baselinewert und die Stratifizierungsvariable Region.
[3)] Die Tabellen mit den Ergebnissen für das präspezifizierte MMRM-Modell fehlen im Bericht der HRQoL-Analysen. Es wird lediglich im Text angegeben, dass keine statistisch signifikanten Veränderungen beobachtet wurden.

Abkürzungen: BSC: Best Supportive Care; k. A.: keine Angabe; KI: Konfidenzintervall; MMRM: mixed-model repeated measure; n. s.: nicht signifikant; SD: Standardabweichung; TranQoL: Transfusion-dependent Quality of Life Questionaire

Der pU legte keine Interaktionstests für Subgruppenanalysen gemäß dem präspezifizierten Analysemodell bezogen auf den Endpunkt TranQoL vor. Die post-hoc definierten Modelle weichen von den präspezifizierten Modellen ab. Gründe für die Änderungen der Analysemodelle werden nicht genannt. Die Subgruppenanalysen werden daher nicht dargestellt.

SF-36

Tabelle 17: Rücklaufquote des SF-36 bezogen auf die ITT-Population der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	**Luspatercept + BSC N = 224**		**Placebo + BSC N = 112**	
Zeitpunkt	**Personen mit auswertbarem Fragebogen**	**Rücklauf-quote (%)[1)]**	**Personen mit auswertbarem Fragebogen**	**Rücklauf-quote (%)[1)]**
Baseline	222	99,1	112	100
Woche 12	194	86,6	90	80,4
Woche 24	197	87,9	91	81,3
Woche 36	186	83,0	89	79,5
Woche 48	176	78,6	88	78,6
Woche 60	113	50,4	50	44,6
Woche 72	34	15,2	18	16,1
Woche 84	8	3,5	3	2,7

ACE-536-B-THAL-001 (BELIEVE)	Luspatercept + BSC N = 224		Placebo + BSC N = 112	
Woche 96	3	1,3	2	1,8
Woche 108	2	0,9	2	1,8
Woche 120	1	0,4	2	1,8
Woche 132	1	0,4	2	1,8
Woche 144	1	0,4	2	1,8

[1)] Die Rücklaufquote wurde bezogen auf die ITT-Population berechnet.

Abkürzungen: BSC: Best Supportive Care

Zu allen Studienvisiten bis Woche 48 lag die Rücklaufquote des SF-36 bezogen auf die ITT-Population über 70 % (Tabelle 17). Zu Woche 48 lag die Rücklaufquote in beiden Studienarmen bei knapp unter 80 %, anschließend fiel die Rücklaufquote auf deutlich unter 70 %. Der Anteil an Respondern in PCS und MCS war im Luspatercept + BSC-Arm numerisch höher (PCS: 17,6 %, MCS: 19,0 %) als im Placebo + BSC-Arm (PCS: 10,7 %, MCS: 14,6 %, Tabelle 18). Die statistische Signifikanz der Unterschiede kann nicht beurteilt werden, da aufgrund von Unklarheiten in der Beschreibung der durchgeführten Analysen die Validität der Analyse nicht beurteilt werden kann (siehe Abschnitt 2.4).

Tabelle 18: Responderanalysen für Verbesserungen des PCS und MCS des SF-36 in der HRQoL evaluable Population der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

ACE-536-B-THAL-001 (BELIEVE)	Anzahl an Respondern zu Woche 48		RR [95%-KI][2)] p-Wert
SF-36 Score	Luspatercept + BSC N = 210 n (%)	Placebo + BSC N = 103 n (%)	
Körperlicher Summenscore (PCS)[1)]	37 (17,6)	11 (10,7)	k. A.
Psychischer Summenscore (MCS)[1)]	40 (19,0)	15 (14,6)	k. A.

[1)] Es werden Responderanalysen für Verbesserung der T-Scores um einen Wert von 5 dargestellt. Ein höherer T-Score repräsentiert eine höhere Lebensqualität.

[2)] Auswertungen zum RR unter Berücksichtigung der Stratifizierungsvariable Region liegen nicht vor.

Abkürzungen: BSC: Best Supportive Care; k. A.: keine Angabe; KI: Konfidenzintervall; MCS: Psychischer Summenscore; PCS: Körperlicher Summenscore; RR: Relatives Risiko / Risk Ratio

3.5 Sicherheit

Ergebnisse zur Sicherheit sind in Tabelle 19 zusammengefasst. Die Zusammenfassung erfolgt lediglich deskriptiv, da für die gesamte verblindete Behandlungs- und Langzeitbehandlungsphase keine Effektschätzer und p-Werte vorgelegt wurden.

Tabelle 19: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE während der verblindeten Behandlungs- und Langzeitbehandlungsphase innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

Patientinnen und Patienten mit mindestens einem …	**Luspatercept + BSC N = 223 n (%)**	**Placebo + BSC N = 109 n (%)**	**RR [95%-KI][1)] p-Wert**
UE	214 (96,0)	101 (92,7)	-
UE des NCI-CTCAE-Grades ≥ 3	65 (29,1)	17 (15,6)	k. A.
SUE	34 (15,2)	6 (5,5)	k. A.
UE, das zum Abbruch der Studienmedikation führte[2)]	12 (5,4)	1 (0,9)	k. A.

[1)] Die vom pU im Dossier angegebenen Effektschätzer und p-Werte berücksichtigen nur die Behandlungsphase von Woche 1–48 und damit nicht die gesamte verblindete Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase. Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode aufgrund der längeren Behandlungs- und Beobachtungszeit bevorzugt. Die Effektschätzer auf Basis der Behandlungsphase werden daher nicht dargestellt. Zudem wurden die RR unstratifiziert berechnet, d. h. ohne Berücksichtigung der Stratifizierungsvariable Region, welche für die Randomisierung verwendet wurde.

[2)] Die Studienteilnehmenden erhielten die Studienmedikation bis zum Auftreten inakzeptabler UE, fehlender Wirksamkeit oder der Rücknahme der Einverständniserklärung, je nachdem, was früher auftrat. Mögliche Therapieabbruchgründe, die vor einem potentiellen Abbruch aufgrund von UE auftreten können, stellen ein konkurrierendes Ereignis dar, weshalb die Ergebnissicherheit und Interpretierbarkeit eingeschränkt ist.

Abkürzungen: BSC: Best Supportive Care; k. A.: keine Angabe; KI: Konfidenzintervall; NCI-CTCAE: National Cancer Institute - Common Terminology Criteria for Adverse Events; RR: Relatives Risiko / Risk Ratio; (S)UE: (Schwerwiegende/s) Unerwünschte/s Ereignis/se

Unerwünschte Ereignisse des NCI-CTCAE-Grades ≥ 3 (Luspatercept + BSC: 29,1 %, Placebo + BSC: 15,6 %), SUE (Luspatercept + BSC: 15,2 %, Placebo + BSC: 5,5 %) und UE, die zum Abbruch der Studienmedikation führten (Luspatercept + BSC: 5,4 %, Placebo + BSC: 0,9 %) traten innerhalb der Behandlungs- und Langzeitbehandlungsphase häufiger im Luspatercept + BSC-Behandlungsarm auf als im Placebo + BSC-Behandlungsarm.

Unerwünschte Ereignisse

Unerwünschte Ereignisse nach SOC und PT sind in Tabelle 20 dargestellt. Unterschiede von mehr als 10 Prozentpunkten zwischen den Behandlungsarmen wurden für UE jeglichen Schweregrads auf SOC-Ebene für Erkrankungen des Nervensystems (Luspatercept + BSC: 40,4 %, Placebo + BSC: 29,4 %) und auf PT-Ebene für Knochenschmerzen (Luspatercept + BSC: 19,7 %, Placebo + BSC: 8,3 %) beobachtet. Unterschiede von 5 Prozentpunkten oder mehr zwischen den Behandlungsarmen mit einem häufigeren Auftreten innerhalb des Luspatercept + BSC-Behandlungsarms wurden für folgende SOC beobachtet: Infektionen und parasitäre Erkrankungen (Luspatercept + BSC: 63,2 %, Placebo + BSC: 57,8 %), Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen (Luspatercept + BSC: 61,4 %, Placebo + BSC: 56,0 %), Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (Luspatercept + BSC: 47,1 %, Placebo + BSC: 41,3 %), Erkrankungen der Atemwege, des Brustraums und des Mediastinums (Luspatercept + BSC: 31,8 %, Placebo + BSC: 26,6 %), Erkrankungen der Haut und des Unterhautzellgewebes (Luspatercept + BSC: 21,5 %, Placebo + BSC: 15,6 %), Untersuchungen (Luspatercept + BSC: 19,7 %, Placebo + BSC: 11,9 %), Stoffwechsel- und Ernährungsstörungen (Luspatercept + BSC: 15,2 %, Placebo + BSC: 6,4 %), Erkrankungen des Bluts und des Lymphsystems (Luspatercept + BSC: 14,8 %, Placebo + BSC: 6,4 %) und Gefäßerkrankungen (Luspatercept + BSC: 11,2 %, Placebo + BSC: 5,5 %). Auf Ebene der PT traten Arthralgie (Luspatercept + BSC: 19,3 %, Placebo + BSC: 11,9 %), Schwindel (Luspatercept + BSC: 11,2 %, Placebo + BSC: 4,6 %), Hyperurikämie

(Luspatercept + BSC: 7,2 %, Placebo + BSC: 0 %) und Hypertension (Luspatercept + BSC: 8,1 %, Placebo + BSC: 2,8 %) um mindestens 5 Prozentpunkte häufiger im Luspatercept + BSC-Arm auf als im Placebo + BSC-Arm. Im Placebo + BSC-Arm traten Infektionen der oberen Atemwege (Luspatercept + BSC: 26,5 %, Placebo + BSC: 33,0 %) und Pyrexie (Luspatercept + BSC: 16,1 %, Placebo + BSC: 21,1 %) um mindestens 5 Prozentpunkte häufiger als im Luspatercept + BSC-Arm auf.

Tabelle 20: UE mit Inzidenz ≥ 10 % bzw. ≥ 10 Ereignisse und einer Inzidenz von ≥ 1 % nach MedDRA-Systemorganklasse und Preferred Term innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

MedDRA Systemorganklasse, Preferred Term[1)]	**Luspatercept + BSC N = 223 n (%)**	**Placebo + BSC N = 109 n (%)**	**RR [95%-KI][2)] p-Wert**
Infektionen und parasitäre Erkrankungen	**141 (63,2)**	**63 (57,8)**	**k. A.**
Infektion der oberen Atemwege	59 (26,5)	36 (33,0)	k. A.
Pharyngitis	20 (9,0)	13 (11,9)	k. A.
Influenza	19 (8,5)	6 (5,5)	k. A.
Virale Infektion der oberen Atemwege	14 (6,3)	2 (1,8)	k. A.
Gastroenteritis	12 (5,4)	8 (7,3)	k. A.
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	**137 (61,4)**	**61 (56,0)**	**k. A.**
Rückenschmerzen	61 (27,4)	32 (29,4)	k. A.
Knochenschmerzen	44 (19,7)	9 (8,3)	k. A.
Arthralgie	43 (19,3)	13 (11,9)	k. A.
Myalgie	22 (9,9)	11 (10,1)	k. A.
Schmerzen in den Extremitäten	21 (9,4)	9 (8,3)	k. A.
Schmerzen des Muskel- und Skelettsystems	14 (6,3)	9 (8,3)	k. A.
Nackenschmerzen	10 (4,5)	8 (7,3)	k. A.
Wirbelsäulenschmerzen	10 (4,5)	4 (3,7)	k. A.
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	**105 (47,1)**	**45 (41,3)**	**k. A.**
Pyrexie	36 (16,1)	23 (21,1)	k. A.
Fatigue	30 (13,5)	14 (12,8)	k. A.
Asthenie	22 (9,9)	11 (10,1)	k. A.
Influenza-artige Erkrankung	17 (7,6)	8 (7,3)	k. A.
Schmerzen	13 (5,8)	4 (3,7)	k. A.
Erkrankungen des Nervensystems	**90 (40,4)**	**32 (29,4)**	**k. A.**
Kopfschmerz	58 (26,0)	26 (23,9)	k. A.
Schwindel	25 (11,2)	5 (4,6)	k. A.
Lethargie	10 (4,5)	4 (3,7)	k. A.

MedDRA Systemorganklasse, Preferred Term[1)]	Luspatercept + BSC N = 223 n (%)	Placebo + BSC N = 109 n (%)	RR [95%-KI][2)] p-Wert
Erkrankungen des Gastrointestinaltrakts	**80 (35,9)**	**36 (33,0)**	**k. A.**
Diarrhö	27 (12,1)	11 (10,1)	k. A.
Nausea	20 (9,0)	6 (5,5)	k. A.
Bauchschmerzen	18 (8,1)	7 (6,4)	k. A.
Erbrechen	18 (8,1)	8 (7,3)	k. A.
Oberbauchschmerzen	15 (6,7)	7 (6,4)	k. A.
Dyspepsie	11 (4,9)	1 (0,9)	k. A.
Zahnschmerzen	11 (4,9)	4 (3,7)	k. A.
Erkrankungen der Atemwege, des Brustraums und des Mediastinums	**71 (31,8)**	**29 (26,6)**	**k. A.**
Husten	32 (14,3)	12 (11,0)	k. A.
Schmerzen im Oropharynx	28 (12,6)	12 (11,0)	k. A.
Kongestion der Nase	12 (5,4)	5 (4,6)	k. A.
Verletzungen, Vergiftungen und durch Eingriffe bedingte Komplikationen	**48 (21,5)**	**24 (22,0)**	**k. A.**
Erkrankungen der Haut und des Unterhautzellgewebes	**48 (21,5)**	**17 (15,6)**	**k. A.**
Urtikaria	11 (4,9)	4 (3,7)	k. A.
Untersuchungen	**44 (19,7)**	**13 (11,9)**	**k. A.**
Lebereisenkonzentration erhöht	12 (5,4)	2 (1,8)	k. A.
Alanin-Aminotransferase erhöht	10 (4,5)	4 (3,7)	k. A.
Stoffwechsel- und Ernährungsstörungen	**34 (15,2)**	**7 (6,4)**	**k. A.**
Hyperurikämie	16 (7,2)	0	k. A.
Erkrankungen des Bluts und des Lymphsystems	**33 (14,8)**	**7 (6,4)**	**k. A.**
Anämie	10 (4,5)	1 (0,9)	k. A.
Herzerkrankungen	**26 (11,7)**	**11 (10,1)**	**k. A.**
Gefäßerkrankungen	**25 (11,2)**	**6 (5,5)**	**k. A.**
Hypertension	18 (8,1)	3 (2,8)	k. A.

[1)] Bei mehreren Ereignissen einer/eines Studienteilnehmenden in einer bestimmten Systemorganklasse wurde dies als einmaliges Ereignis in der Systemorganklasse gezählt.

[2)] Die vom pU im Dossier angegebenen Effektschätzer und p-Werte berücksichtigen nur die Behandlungsphase von Woche 1–48 und damit nicht die gesamte verblindete Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase. Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode aufgrund der längeren Behandlungs- und Beobachtungszeit bevorzugt. Die Effektschätzer auf Basis der Behandlungsphase werden daher nicht dargestellt. Zudem wurden die RR unstratifiziert berechnet, d. h. ohne Berücksichtigung der Stratifizierungsvariable Region, welche für die Randomisierung verwendet wurde.

Abkürzungen: BSC: Best Supportive Care; k. A.: keine Angabe; KI: Konfidenzintervall; RR: Relatives Risiko / Risk Ratio

Schwerwiegende unerwünschte Ereignisse

Schwerwiegende unerwünschte Ereignisse nach SOC und PT sind in Tabelle 21 dargestellt.

Tabelle 21: SUE mit Inzidenz ≥ 5 % nach MedDRA-Systemorganklasse und Preferred Term innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

MedDRA Systemorganklasse, Preferred Term[1)]	Luspatercept + BSC N = 223 n (%)	Placebo + BSC N = 109 n (%)	RR [95%-KI][2)] p-Wert
Infektionen und parasitäre Erkrankungen	**13 (5,8)**	**3 (2,8)**	**k. A.**

[1)] Bei mehreren Ereignissen einer/eines Studienteilnehmenden in einer bestimmten Systemorganklasse wurde dies als einmaliges Ereignis in der Systemorganklasse gezählt.
[2)] Die vom pU im Dossier angegebenen Effektschätzer und p-Werte berücksichtigen nur die Behandlungsphase von Woche 1–48 und damit nicht die gesamte verblindete Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase. Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode aufgrund der längeren Behandlungs- und Beobachtungszeit bevorzugt. Die Effektschätzer auf Basis der Behandlungsphase werden daher nicht dargestellt. Zudem wurden die RR unstratifiziert berechnet, d. h. ohne Berücksichtigung der Stratifizierungsvariable Region, welche für die Randomisierung verwendet wurde.

Abkürzungen: BSC: Best Supportive Care; k. A.: keine Angabe; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; RR: Relatives Risiko / Risk Ratio; UE: Unerwünschte/s Ereignis/se

Unerwünschte Ereignisse des NCI-CTCAE-Grades ≥ 3

Unerwünschte Ereignisse des NCI-CTCAE-Grades ≥ 3 nach SOC und PT sind in Tabelle 22 dargestellt.

Tabelle 22: UE des NCI-CTCAE-Grades ≥ 3 mit Inzidenz ≥ 5 % nach MedDRA-Systemorganklasse und Preferred Term innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

MedDRA Systemorganklasse, Preferred Term[1)]	Luspatercept + BSC N = 223 n (%)	Placebo + BSC N = 109 n (%)	RR [95%-KI][2)] p-Wert
Infektionen und parasitäre Erkrankungen	**15 (6,7)**	**6 (5,5)**	**k. A.**
Untersuchungen	**12 (5,4)**	**7 (6,4)**	**k. A.**

[1)] Bei mehreren Ereignissen einer/eines Studienteilnehmenden in einer bestimmten Systemorganklasse wurde dies als einmaliges Ereignis in der Systemorganklasse gezählt.
[2)] Die vom pU im Dossier angegebenen Effektschätzer und p-Werte berücksichtigen nur die Behandlungsphase von Woche 1–48 und damit nicht die gesamte verblindete Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase. Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode aufgrund der längeren Behandlungs- und Beobachtungszeit bevorzugt. Die Effektschätzer auf Basis der Behandlungsphase werden daher nicht dargestellt. Zudem wurden die RR unstratifiziert berechnet, d. h. ohne Berücksichtigung der Stratifizierungsvariable Region, welche für die Randomisierung verwendet wurde.

Abkürzungen: BSC: Best Supportive Care; k. A.: keine Angabe; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; NCI-CTCAE: National Cancer Institute – Common Terminology Criteria for Adverse Events; RR: Relatives Risiko / Risk Ratio; UE: Unerwünschte/s Ereignis/se

Unerwünschte Ereignisse von besonderem Interesse und weitere relevante Sicherheitsereignisse

Tabelle 23: Anzahl an Patientinnen und Patienten mit UE von besonderem Interesse und weiteren relevanten Sicherheitsereignissen innerhalb der Sicherheitspopulation der Studie ACE-536-B-THAL-001 (BELIEVE) (Datenschnitt 11.05.2018)

Ereignis[1)]	**Luspatercept + BSC N = 223 n (%)**	**Placebo + BSC N = 109 n (%)**	**RR [95%-KI][2)] p-Wert**
UE von besonderem Interesse[3)]	1 (0,4)	0	-
Thromboembolische oder thrombophlebitische Ereignisse[4)]	9 (4,0)	1 (0,9)	k. A.
Knochenschmerzen[5)]	44 (19,7)	9 (8,3)	k. A.

[1)] Bei mehreren Ereignissen einer/eines Studienteilnehmenden wurde dies als einmaliges Ereignis gezählt.
[2)] Die vom pU im Dossier angegebenen Effektschätzer und p-Werte berücksichtigen nur die Behandlungsphase von Woche 1–48 und damit nicht die gesamte verblindete Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase. Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode aufgrund der längeren Behandlungs- und Beobachtungszeit bevorzugt. Die Effektschätzer auf Basis der Behandlungsphase werden daher nicht dargestellt. Zudem wurden die RR unstratifiziert berechnet, d. h. ohne Berücksichtigung der Stratifizierungsvariable Region, welche für die Randomisierung verwendet wurde.
[3)] Als UE von besonderem Interesse waren Malignitäten und prämaligne Erkrankungen definiert. Das eine aufgetretene Ereignis wird als erhöhter humaner Choriongonadotropin-(HCG)-Wert beschrieben, der im Rahmen eines falsch-positiven Schwangerschaftstests festgestellt wurde.
[4)] Als weiteres relevantes Sicherheitsereignis präspezifizierter SMQ.
[5)] Als weiteres relevantes Sicherheitsereignis präspezifizierter Preferred Term.

Abkürzungen: BSC: Best Supportive Care; k. A.: keine Angabe; KI: Konfidenzintervall; RR: Relatives Risiko / Risk Ratio; SMQ: Standardised MedDRA Query; UE: Unerwünschte/s Ereignis/se

Subgruppenanalysen

Der pU legte keine Interaktionstests für Subgruppenanalysen bezogen auf den Endpunkt UE in der gesamten verblindeten Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase vor.

4 Diskussion der Methodik und Ergebnisse

4.1 Zulassungsstatus und Zulassungspopulation von Luspatercept

Luspatercept ist zugelassen für die Behandlung von erwachsenen Patientinnen und Patienten mit transfusionsabhängiger Anämie, die mit einer β-Thalassämie verbunden ist. In die Studie sollten gemäß Einschlusskriterien Patientinnen und Patienten mit β-Thalassämie eingeschlossen werden, die 6–20 EK-Einheiten innerhalb von 24 Wochen vor Beginn der Behandlung mit Luspatercept benötigten. Außerdem durfte kein transfusionsfreier Zeitraum von 35 Tagen oder länger vorliegen. Die untere Begrenzung der Transfusionshäufigkeit und die Festlegung des maximalen transfusionsfreien Zeitraums definiert damit den Begriff der Transfusionsabhängigkeit für die Studie. In der Fachinformation wird eine solche Definition nicht vorgenommen. Auch die Begrenzung auf ein Maximum von 20 EK-Einheiten pro 24 Wochen spiegelt sich in der Zulassung nicht wider und stellt einen Ausschluss von Personen mit einer sehr hohen Transfusionslast dar. Aus den Baselinecharakteristika geht jedoch hervor, dass zum Teil auch Patientinnen und Patienten mit einer größeren Anzahl an EK-Einheiten pro 24 Wochen eingeschlossen wurden (Luspatercept + BSC: max = 24 EK-Einheiten / 24 Wochen; Placebo + BSC: max = 26 EK-Einheiten / 24 Wochen). In die Studienpopulation wurden weiterhin nur Personen mit einem ECOG-Status von 0 oder 1 eingeschlossen, was der in klinischen Studien üblichen Begrenzung auf Menschen mit geringen Einschränkungen entspricht. Von der Studienteilnahme ausgeschlossen waren Menschen mit Hämoglobin S / β-Thalassämie. Der pU beschreibt hierzu im Dossier, dass es sich bei der gemischt heterozygoten Sichelzellkrankheit (Hämoglobin S / β-Thalassämie) um eine anomale Strukturvariante des β-Globin-Gens ohne β-Thalassämie-Phänotyp handele, welche separat zu betrachten und nicht vom Anwendungsgebiet von Luspatercept umfasst sei.

Innerhalb der Studie waren Dosisreduktionen auf geringere Dosen als in der Fachinformation angegeben möglich. Diese wurden nur bei 3 Personen angewendet. Die weitere Dosierung erfolgte innerhalb der Studie gemäß Fachinformation.

Die Zulassung erfolgte ohne Besonderheiten bzw. weitere Bedingungen. Im EPAR werden weitere Langzeituntersuchungen zur Sicherheit einschließlich embolischer und thromboembolischer Ereignisse (nur bei Patientinnen und Patienten mit β-Thalassämie und Splenektomie) und Malignitäten bzw. prämalignen Erkrankungen von Studienteilnehmenden aller durch den pU gesponserten klinischen Studien mit Luspatercept im Rahmen der Pharmakovigilanz gefordert. Diese sollen jährlich für die ersten 5 Jahre und final im 2. Quartal 2029 vorgelegt werden.

4.2 Design und Methodik der Studie

Die Nutzenbewertung basiert auf der multizentrischen, doppelblinden, randomisiert kontrollierten Phase-III-Studie ACE-536-B-THAL-001 (BELIEVE) zur Untersuchung der Sicherheit und Wirksamkeit von Luspatercept bei Patientinnen und Patienten mit transfusionsabhängiger β-Thalassämie. In die Phase-III-Studie ACE-536-B-THAL-001 (BELIEVE) wurden 336 Patientinnen und Patienten eingeschlossen, welche in insgesamt 65 Studienzentren (davon keines in Deutschland und 23 in Europa) rekrutiert und untersucht wurden. Mehr als die Hälfte der Studienpopulation (55 %) wurde im Mittleren Osten und Nordafrika sowie in Asien und Australien behandelt. Dies kann die Übertragung der Studienergebnisse auf den deutschen Versorgungskontext erschweren.

Die Studie beinhaltete eine 12-wöchige Screening-Phase, eine 48-wöchige Behandlungsphase, eine Langzeit-Behandlungsphase und eine Open-Label-Phase sowie eine Post-Treatment-Follow-up-Phase. In der 12-wöchigen Screeningphase wurden die Ein- bzw. Ausschlusskriterien überprüft und der Bedarf an Bluttransfusionen monitoriert. Bei Erfüllung der Ein- und Ausschlusskriterien erfolgte die Randomisierung im Verhältnis 2:1 stratifiziert nach Region (Nordamerika und Europa, Mittlerer Osten und Nordafrika, Asien-Pazifik) auf die beiden Studienarme: Luspatercept + BSC vs. Placebo + BSC. Die Behandlungsphase beginnt mit der Verabreichung der ersten Dosis und dauert 48 Wochen. Eine verblindete Fortsetzung der Behandlung war nach Beendigung der 48-wöchigen Behandlungsphase innerhalb der Langzeit-Behandlungsphase solange im Ermessen des ärztlichen Prüfpersonals möglich, bis der letzte Patient bzw. die letzte Patientin der Studie die 48-wöchige Behandlungsphase abgeschlossen hatte. Die Langzeit-Behandlungsphase endet daher, nachdem alle Patientinnen und Patienten die 48-wöchige Behandlungsphase abgeschlossen oder diese vorzeitig beendet hatten. Nach Abschluss der Langzeitbehandlungsphase erfolgte eine Entblindung. In der offenen Extensionsphase war eine Weiterbehandlung mit Luspatercept für Patientinnen und Patienten aus dem Luspatercept + BSC-Behandlungsarm möglich sowie ein Therapiebeginn mit Luspatercept für Studienteilnehmende aus dem Placebo + BSC-Arm. Eine Nachbeobachtung der Studienteilnehmenden in der offenen Extensionsphase ist für insgesamt 5 Jahre nach der ersten Dosis geplant. Studienteilnehmende beider Behandlungsarme, welche die Therapie abbrachen, werden unabhängig vom Zeitpunkt des Abbruchs für 3 Jahre (156 Wochen) innerhalb der Post-Treatment-Follow-up-Phase nachbeobachtet.

Der zur Nutzenbewertung herangezogene Datenschnitt der Studie BELIEVE erfolgte am 11. Mai 2018 zum Abschluss der Langzeitbehandlungsphase. Der pU reicht darüber hinaus Analysen zum Gesamtüberleben und weitere Ereigniszeitanalysen für Datenschnitte am 07.01.2019 und 01.07.2019 ein. Es wird davon ausgegangen, dass diese Datenschnitte nach der Entblindung der Studie stattfanden und ein großer Teil der Personen im Placebo + BSC-Arm eine Therapie mit Luspatercept begonnen hatte. Für die vorgelegten Zusatzanalysen zu den späteren Datenschnitten wurden jedoch keine Informationen zum weiteren Patientenfluss, zum Zeitpunkt der Entblindung, zur Exposition mit der Intervention sowie zur verabreichten Begleittherapie vorgelegt. Die Datenschnitte vom 07.01.2019 und 01.07.2019 werden daher nicht für die Nutzenbewertung herangezogen. Zudem ist der generelle Stellenwert von Analysen nach Therapiewechsel unklar. Für die Analysen zum Gesamtüberleben wird angegeben, dass Personen, die ursprünglich mit Placebo behandelt wurden und nach Entblindung zu Luspatercept wechselten, zum Zeitpunkt des Therapiewechsels zensiert wurden. Weiterhin wurden keine Daten zur Lebensqualität in der offenen Extensionsphase erhoben. Valide vergleichende Analysen erscheinen dadurch zeitlich nur in sehr begrenztem Umfang nach dem Datenschnitt vom 11.05.2018 prinzipiell möglich zu sein.

Insgesamt wurden 336 Personen innerhalb der Studie randomisiert: 224 in den Luspatercept + BSC-Behandlungsarm und 112 in den Placebo + BSC-Behandlungsarm. Davon erhielten 223 (Luspatercept + BSC) bzw. 109 (Placebo + BSC) die Intervention und sind somit Bestandteil der Sicherheitspopulation. Zum Zeitpunkt des Datenschnittes vom 11. Mai 2018 erhielten noch 80,8 % der Personen im Luspatercept + BSC-Arm und 75,9 % im Placebo + BSC-Arm die Intervention. Die mediane Behandlungsdauer betrug 64 Wochen in beiden Behandlungsarmen. Gründe für den Abbruch der Einnahme der Studienmedikation waren Entzug der Einverständniserklärung (Luspatercept + BSC: 11,6 %, Placebo + BSC: 10,7 %), UE (Luspatercept + BSC: 4,5 %, Placebo + BSC: 0,9 %), fehlender Wirksamkeitsnachweis

(Luspatercept + BSC: 0,9 %, Placebo + BSC: 7,1 %), Protokollverletzungen (Luspatercept + BSC: 0,4 %, Placebo + BSC: 0) sowie andere Ursachen (Luspatercept + BSC: 1,3 %, Placebo + BSC: 2,7 %). Zum Zeitpunkt des Datenschnittes hatten 11,6 % (Luspatercept + BSC: n = 26) bzw. 12,5 % (Placebo + BSC: n = 14) der Teilnehmenden die Studie vorzeitig beendet: 5,8 % (Luspatercept + BSC: n = 13) bzw. 5,4 % (Placebo + BSC: n = 6) aufgrund des Entzugs der Einverständniserklärung, 1,8 % (Luspatercept + BSC: n = 4) bzw. 0 (Placebo + BSC) aufgrund von UE, 0,4 % (Luspatercept + BSC: n = 1) bzw. 0,9 % (Placebo + BSC: n = 1) aufgrund von Tod und jeweils 0,9 % pro Studienarm (Luspatercept + BSC: n = 2, Placebo + BSC: n = 1) aus anderen Gründen. Die Aufsummierung der Anzahl an Personen pro Abbruchgrund (Luspatercept + BSC: n = 20, Placebo + BSC: n = 8) ist kleiner als die Anzahl an Personen, welche die Studie abgebrochen haben (Luspatercept + BSC: n = 26, Placebo + BSC: n = 14). Diese Diskrepanz wird auch in der durch den pU eingereichten Abbildung zum Patientenfluss im Dossier (Abbildung 4-24) deutlich. Eine Erklärung für diese Diskrepanz konnte in den Studienunterlagen nicht identifiziert werden.

Laut Entscheidungsbaum zum Management der Studienteilnehmenden sollte nach Abschluss der 48-wöchigen Behandlungsphase für jede Person entschieden werden, ob sie von der Behandlung klinisch profitiert, und nur Personen, welche profitieren, die Einnahme der Studienmedikation fortsetzen. Kriterien zur Beurteilung des klinischen Profitierens werden im Studienprotokoll nicht benannt. Es ist unklar, wie diese Regelung in der Praxis umgesetzt wurde, da die Mehrheit der Studienpopulation zum Zeitpunkt des Datenschnittes unter Behandlung stand und damit vermutlich auch in die Langzeitbehandlungsphase übergegangen war. Angaben zur Anzahl an Personen, welche in die Langzeitbehandlungsphase oder Post-Treatment-Follow-up-Phase eingetreten sind bzw. diese abgebrochen oder beendet haben, wurden nicht identifiziert. Insgesamt ist der Patientenfluss für den Datenschnitt vom 11.05.2018 nur eingeschränkt nachvollziehbar. Angaben zur medianen Beobachtungsdauer liegen nicht vor.

Das Alter der Studienteilnehmenden lag in beiden Behandlungsarmen im Median bei 30 Jahren. Lediglich 7,6 % (Luspatercept + BSC) bzw. 4,5 % (Placebo + BSC) der Studienteilnehmenden war älter als 50 Jahre. Der Anteil an Frauen an der Studienpopulation betrug 58,9 % (Luspatercept + BSC) bzw. 56,3 % (Placebo + BSC). Ca. 30 % der Studienteilnehmenden (Luspatercept + BSC: 30,4 %, Placebo + BSC: 31,3 %) hatte einen β_0/β_0-Genotyp, ca. 70 % (Luspatercept + BSC: 69,2 %, Placebo + BSC: 68,8 %) einen Nicht-β_0/β_0-Genotyp. Bei über 70 % (Luspatercept + BSC: 77,7 %, Placebo + BSC: 74,1 %) der Studienpopulation lag eine reine β-Thalassämie vor, 13,8 % (Luspatercept + BSC) bzw.18,8 % (Placebo + BSC) wiesen eine Hämoglobin E / β-Thalassämie auf und 8,0 % (Luspatercept + BSC) bzw. 7,1 % (Placebo + BSC) eine β-Thalassämie kombiniert mit einer α-Thalassämie. Die Charakteristika der Studienpopulation (siehe Tabelle 9) sind weitestgehend gleichmäßig über beide Behandlungsarme verteilt.

Für die Studie liegen zwei SAP vor, ein genereller vom 19. Juni 2018 und ein spezifischer für die Analysen zur Lebensqualität vom 12. Juni 2018. Beide SAP wurden erst nach dem Datenschnitt am 11. Mai 2018 finalisiert. Es ist daher unklar, ob die statistischen Analysen komplett unabhängig und in Unkenntnis der Studiendaten geplant wurden. Weiterhin wird im Studienbericht angegeben, dass die Visite zu Woche 48 des letzten Patienten bzw. der letzten Patientin am 14. Mai 2018 und damit 3 Tage nach dem Datenschnitt erfolgte. Daher ist unklar, wie mit den Daten des letzten Patienten bzw. der letzten Patientin umgegangen wurde und warum der Datenschnitt nicht wie geplant nach Abschluss der Visite zu Woche 48 des letzten Patienten bzw. der letzten Patientin erfolgte.

Für das Dossier wurden über den SAP hinausgehende Post-hoc-Analysen durchgeführt. Für diese Analysen wurde kein separater Analyseplan vorgelegt und die Analysen sind teilweise nicht gut dokumentiert, sodass die angewendeten Analyseverfahren zum Teil unklar bleiben. Es wird kritisch gesehen, dass die Post-hoc-Analysen, soweit nachvollziehbar, von den präspezifizierten Analysen hinsichtlich Adjustierungsvariablen, Stratifizierungsfaktoren und Modellauswahl abweichen. Für mehrere Endpunkte wird im SAP die Berechnung von OR mittels stratifizierten CMH-Tests geplant. Mit dem gleichen Verfahren können auch RR unter Berücksichtigung der Stratifizierungsfaktoren berechnet werden. Dennoch wurden vom pU RR ohne Berücksichtigung der Stratifizierungsfaktoren berechnet, was in einer stratifiziert randomisierten Studie als nicht adäquat angesehen wird. OR werden in einer klinischen Studie nicht als Effektmaß der Wahl angesehen, da diese nicht das Verhältnis der Risiken, sondern das der Chancen abbilden. Das OR kann insbesondere bei Ereignisraten, wie sie in der Studie auftreten, nicht als Approximation des RR angesehen werden.

Die Studienmedikation (Luspatercept oder Placebo) wurde alle 3 Wochen innerhalb der Studie verabreicht. In Abhängigkeit von Veränderungen im Hb-Wert, dem Auftreten von UE und Veränderungen der Leukozytenzahl (siehe Tabelle 3) waren Dosisreduktionen oder Dosisunterbrechungen vorgesehen. Dosisunterbrechungen wurden häufig durchgeführt und traten bei mehr Personen im Luspatercept + BSC-Behandlungsarm auf (n = 137, 61,4 %) als im Placebo + BSC-Behandlungsarm (n = 45, 41,3 %). Dosisreduktionen wurden dagegen seltener angewendet und traten ebenfalls bei mehr Personen im Luspatercept + BSC-Behandlungsarm auf (n = 29, 13,0 %) als im Placebo + BSC-Behandlungsarm (n = 4, 3,7 %). Innerhalb der Studie waren Dosisreduktionen auf geringere Dosen als in der Fachinformation angegeben möglich. Diese wurden nur bei 3 Personen angewendet. Dosiserhöhungen auf 1,25 mg/kg Körpergewicht waren bei Nichterreichen einer Reduktion der Anzahl der transfundierten EK-Einheiten um 33 % oder mehr über mindestens 2 Dosiszyklen möglich und lagen bei Reduktion der transfundierten EK-Einheiten um mehr als 33 % aber weniger als 50 % im Ermessen des Prüfpersonals. Eine Dosiserhöhung auf die maximal mögliche Dosis (1,25 mg/kg Körpergewicht) wurde bei 46,2 % der Personen im Luspatercept + BSC-Arm und bei 66,1 % im Placebo + BSC-Arm durchgeführt. Die Anpassung der Dosierung bei fehlender Wirksamkeit orientierte sich am primären Endpunkt und es ist nicht auszuschließen, dass das genaue Monitoring der transfundierten EK-Einheiten Vermutungen über den zugeteilten Behandlungsarm ermöglichte. Studienteilnehmende aus dem Luspatercept + BSC-Behandlungsarm erhielten im Mittel (SD) 24,1 (8,5) EK-Einheiten innerhalb der 48-wöchigen Behandlungsphase. Im gleichen Zeitraum erhielten Studienteilnehmende aus dem Placebo + BSC-Behandlungsarm im Mittel (SD) 30,2 (7,4) EK-Einheiten. Angaben über die Anzahl an EK-Einheiten pro Zeiteinheit über die gesamte Studienperiode hinweg unter Berücksichtigung aller Studienteilnehmenden liegen nicht vor.

Das Verzerrungspotential auf Studienebene wird trotz der zuvor beschriebenen Limitationen als niedrig eingeschätzt.

4.3 Mortalität

Innerhalb der BELIEVE-Studie traten bis zum Datenschnitt vom 11.05.2018 zwei Todesfälle auf, einer im Luspatercept + BSC-Arm und einer im Placebo + BSC-Arm. Aufgrund der geringen Ereigniszahl ist eine Ereigniszeitanalyse der Daten nicht sinnvoll. Um die für Ereigniszeitanalysen notwendige, deutlich größere Anzahl an Ereignissen zu erreichen, ist innerhalb des Krankheitsbildes eine wesentlich längere Beobachtungszeit notwendig. Es ist davon auszugehen,

dass diese innerhalb der BELIEVE-Studie nicht erreicht werden kann, da nach dem Datenschnitt eine Entblindung und ein Therapiebeginn mit Luspatercept für Personen aus dem Placebo + BSC-Arm möglich war. Patientinnen und Patienten aus dem Placebo + BSC-Arm wurden zum Zeitpunkt des Therapiewechsels zensiert, wodurch der Placebo + BSC-Arm stark ausgedünnt wird und eine vergleichende Analyse nicht mehr möglich ist.

Zusammenfassende Einschätzung zur Wirksamkeit

Insgesamt traten innerhalb der BELIEVE-Studie nur sehr wenige Todesfälle auf, einer im Luspatercept + BSC-Arm und einer im Placebo + BSC-Arm. Der Einfluss von Luspatercept auf die Gesamtmortalität kann auf Basis der vorgelegten Daten nicht bewertet werden.

4.4 Morbidität

Innerhalb der BELIEVE-Studie ist die Hospitalisierung der einzige patientenrelevante Endpunkt der Kategorie Morbidität. Aus der Operationalisierung des Endpunktes ist nicht ersichtlich, ob stets eine stationäre Aufnahme mit einer Übernachtung oder ob auch ambulante und teilstationäre Aufenthalte als Hospitalisierung gewertet wurden. Innerhalb der gesamten kontrollierten Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase traten bei 18,8 % der Personen im Luspatercept + BSC-Arm Hospitalisierungen auf. Im Placebo + BSC-Arm dagegen traten nur bei 3,6 % der Personen Hospitalisierungen auf. Angaben zur medianen Beobachtungsdauer für den Endpunkt Hospitalisierung liegen nicht vor. Ebenfalls liegt keine über die gesamte Studienperiode dauernde statistische Analyse der Daten vor. Der pU legt nur für die 48-wöchige Behandlungsphase statistische Analysen vor. Auch für diese Analyse liegen keine adäquaten Effektschätzer in Form von RR vor, welche unter Berücksichtigung der Stratifizierungsvariable der Randomisierung berechnet wurden.

Der Unterschied zwischen den Behandlungsarmen erscheint insgesamt deutlich zu sein. Hospitalisierungen traten aus allen der aufgelisteten Gründe häufiger im Luspatercept + BSC-Arm auf als im Placebo + BSC-Arm, unabhängig davon, ob der Grund der Hospitalisierung innerhalb der Studie liegt, z. B. Unerwünschte Ereignisse (Luspatercept + BSC: 15,6 %; Placebo + BSC: 3,6 %), oder außerhalb der Studie, z. B. elektive Behandlung einer vorbestehenden Erkrankung (Luspatercept + BSC: 4,0 %; Placebo + BSC: 1,8 %). Die Aufsummierung der Hospitalisierungsgründe übersteigt die Anzahl an Personen mit Hospitalisierung. Es ist unklar, ob dies darauf zurückzuführen ist, dass eine Person zu einem Zeitpunkt wegen mehrerer Gründe hospitalisiert werden konnte oder ob Mehrfachhospitalisierungen der gleichen Person gezählt wurden oder ob beides möglich war. Auffällig ist zudem, dass im Luspatercept + BSC-Arm auch häufiger Hospitalisierungen aufgrund von Transfusionen stattfanden (Luspatercept + BSC: 4,9 %; Placebo + BSC: 0,9 %), obwohl die Anzahl an transfundierten EK-Einheiten insgesamt im Luspatercept + BSC-Arm niedriger war als im Placebo + BSC-Arm (siehe Tabelle 11). Der häufigste Grund für Hospitalisierungen waren Unerwünschte Ereignisse. Für diese erfolgte eine Doppelerfassung der Ereignisse im Endpunkt Hospitalisierung und im Endpunkt Unerwünschte Ereignisse.

Das Verzerrungspotential für den Endpunkt Hospitalisierung wird insgesamt als hoch eingeschätzt, insbesondere da Verzerrungen aufgrund von regionalen Unterschieden bei der multizentrischen Studie nicht ausgeschlossen werden können. So kann die interne Validität durch regionale Unterschiede eingeschränkt sein, da die Stratifizierung der Randomisierung nicht auf Ebene der Studienzentren oder Länder erfolgte, sondern auf Ebene geographischer Regionen. Eine

Ungleichverteilung der Testpersonen über Länder mit verschiedenen Kriterien der Notwendigkeit einer Hospitalisierung kann zu einer Verzerrung der Ergebnisse führen. Weiterhin ist die Übertragbarkeit auf den deutschen Versorgungskontext unklar.

Zusammenfassende Einschätzung zur Lebensqualität

Der Einfluss von Luspatercept auf den Endpunkt Hospitalisierung kann aufgrund der vorgelegten Daten und statistischen Analysen nicht abschließend bewertet werden.

4.5 Lebensqualität

Daten zur gesundheitsbezogenen Lebensqualität wurden in der Studie BELIEVE mit Hilfe der selbst-ausfüllbaren Fragebögen TranQoL und SF-36 erhoben. Für beide Fragebögen wurden bis zur Woche 48 Rücklaufquoten über 70 % erreicht, sodass Daten aus diesem Zeitraum für die Nutzenbewertung herangezogen werden. Es ist bekannt, dass die Lebensqualität von Menschen mit transfusionsabhängiger β-Thalassämie periodisch schwankt. Kurz nach einer Transfusion weisen die Patientinnen und Patienten weniger Anämie-bedingte Symptome und eine höhere Lebensqualität auf als kurz vor einer erneuten Transfusion. Dies wurde auch in den Validierungsstudien für den TranQoL für einen Teil der Studienteilnehmenden berichtet. In der BELIEVE-Studie erfolgte die Erhebung der Lebensqualität zu festgelegten Studienvisiten zu Woche 12, 24, 36 und 48 und unabhängig vom Transfusionszyklus.

Für den TranQoL wurden keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen festgestellt (siehe Tabelle 16). Das Verzerrungspotential für den Endpunkt TranQoL wird aufgrund der hohen Anzahl an fehlenden Werten und der fehlenden Erhebung bei Abbruch der Einnahme der Studienmedikation als hoch eingeschätzt. Da anzunehmen ist, dass Personen, die die Behandlung abbrechen, tendenziell eine schlechtere Lebensqualität haben als Personen, die die Behandlung weiterführen, führt die fehlende Erhebung des Fragebogens im Fall eines Behandlungsabbruches zur Verzerrung.

Der Anteil an Respondern für Verbesserungen in PCS und MCS des SF-36 war im Luspatercept + BSC-Arm numerisch höher (PCS: 17,6 %, MCS: 19,0 %) als im Placebo + BSC-Arm (PCS: 10,7 %, MCS: 14,6 %). Responderanalysen für Verschlechterungen konnten in den Studienunterlagen nicht identifiziert werden. Die statistische Signifikanz der Unterschiede kann nicht beurteilt werden, da aufgrund von Unklarheiten in der Beschreibung der durchgeführten Analysen die Validität der Analyse nicht beurteilt werden kann (siehe Abschnitt 2.4). Auch das Verzerrungspotential der Analysen zum SF-36 wird aufgrund der hohen Anzahl an fehlenden Werten und der fehlenden Erhebung bei Abbruch der Einnahme der Studienmedikation als hoch eingeschätzt.

Zusammenfassende Einschätzung zur Lebensqualität

Bezüglich der gesundheitsbezogenen Lebensqualität gemessen mittels TranQoL wurde kein statistisch signifikanter Unterschied zwischen den Studienarmen festgestellt. Die mittels SF-36 gemessene Lebensqualität kann aufgrund von Unklarheiten in der Beschreibung der durchgeführten Post-hoc-Analysen nicht beurteilt werden. Das Verzerrungspotential wird für beide Lebensqualitätsendpunkte aufgrund fehlender Werte und fehlender Erhebung bei Abbruch der Einnahme der Studienmedikation als hoch eingeschätzt.

4.6 Sicherheit

Unerwünschte Ereignisse wurden bis 9 Wochen nach Abbruch der Einnahme der Studienmedikation vollständig erfasst. Angaben zur medianen Beobachtungsdauer für die Unerwünschten Ereignisse liegen nicht vor.

Grundsätzlich ist zu beachten, dass in den dargestellten UE auch Ereignisse der Grunderkrankung enthalten sein können. Es wurden keine Auswertungen von UE vorgelegt, für die Ereignisse, die der Grunderkrankung zugerechnet werden, begründet und präspezifiziert ausgeschlossen wurden. Hospitalisierungen, die aufgrund eines UE erfolgten, wurden im Endpunkt Hospitalisierungen und im Endpunkt Unerwünschte Ereignisse und damit doppelt erfasst.

Unerwünschte Ereignisse des NCI-CTCAE-Grades ≥ 3 (Luspatercept + BSC: 29,1 %, Placebo + BSC: 15,6 %), SUE (Luspatercept + BSC: 15,2 %, Placebo + BSC: 5,5 %) und UE, die zum Abbruch der Studienmedikation führten (Luspatercept + BSC: 5,4 %, Placebo + BSC: 0,9 %), traten innerhalb der Behandlungs- und Langzeitbehandlungsphase häufiger im Luspatercept + BSC-Behandlungsarm auf als im Placebo + BSC-Behandlungsarm. Im Interventionsarm wurden insgesamt weniger SUE (n = 34) berichtet als Hospitalisierungen aufgrund von UE (n = 35). Ob dies darauf zurückzuführen ist, dass Mehrfachhospitalisierungen der gleichen Person für die Erfassung der Hospitalisierungsgründe gezählt wurden oder ob hierfür andere Gründe vorliegen, ist unklar.

Für das Dossier berechnet der pU post-hoc Effektschätzer (OR, RR und RD) auf Basis der 48-wöchigen Behandlungsphase. Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode bestehend aus Behandlungs- und Langzeitbehandlungsphase aufgrund der längeren Behandlungs- und Beobachtungszeit bevorzugt. Zudem erfolgte die Berechnung der vorgelegten Effektschätzer gemäß Angaben im Modul 4 nicht unter Berücksichtigung des Stratifizierungsfaktors der Randomisierung (Region) und damit nicht adäquat. Die Effektschätzer auf Basis der 48-wöchigen Behandlungsphase werden daher nicht für die Nutzenbewertung herangezogen. Dies gilt ebenfalls für UE, SUE und UE des NCI-CTCAE-Grades ≥ 3 nach SOC und PT. Unterschiede von mehr als 10 Prozentpunkten zwischen den Behandlungsarmen wurden für UE jeglichen Schweregrades auf SOC-Ebene für Erkrankungen des Nervensystems (Luspatercept + BSC: 40,4 %, Placebo + BSC: 29,4 %) und auf PT-Ebene für Knochenschmerzen (Luspatercept + BSC: 19,7 %, Placebo + BSC: 8,3 %) beobachtet. Unterschiede von 5 Prozentpunkten oder mehr zwischen den Behandlungsarmen mit einem häufigeren Auftreten innerhalb des Luspatercept + BSC-Behandlungsarms wurden für folgende SOC beobachtet: Infektionen und parasitäre Erkrankungen (Luspatercept + BSC: 63,2 %, Placebo + BSC: 57,8 %), Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen (Luspatercept + BSC: 61,4 %, Placebo + BSC: 56,0 %), Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (Luspatercept + BSC: 47,1 %, Placebo + BSC: 41,3 %), Erkrankungen der Atemwege, des Brustraums und des Mediastinums (Luspatercept + BSC: 31,8 %, Placebo + BSC: 26,6 %), Erkrankungen der Haut und des Unterhautzellgewebes (Luspatercept + BSC: 21,5 %, Placebo + BSC: 15,6 %), Untersuchungen (Luspatercept + BSC: 19,7 %, Placebo + BSC: 11,9 %), Stoffwechsel- und Ernährungsstörungen (Luspatercept + BSC: 15,2 %, Placebo + BSC: 6,4 %), Erkrankungen des Bluts und des Lymphsystems (Luspatercept + BSC: 14,8 %, Placebo + BSC: 6,4 %) und Gefäßerkrankungen (Luspatercept + BSC: 11,2 %, Placebo + BSC: 5,5 %). Auf Ebene der PT traten Arthralgie (Luspatercept + BSC: 19,3 %, Placebo + BSC: 11,9 %), Schwindel (Luspatercept + BSC: 11,2 %, Placebo + BSC: 4,6 %), Hyperurikämie (Luspatercept + BSC: 7,2 %, Placebo + BSC: 0 %) und Hypertension (Luspatercept + BSC: 8,1 %, Placebo + BSC: 2,8 %) um

mindestens 5 Prozentpunkte häufiger im Luspatercept + BSC-Arm auf als im Placebo + BSC-Arm. Im Placebo + BSC-Arm traten Infektionen der oberen Atemwege (Luspatercept + BSC: 26,5 %, Placebo + BSC: 33,0 %) und Pyrexie (Luspatercept + BSC: 16,1 %, Placebo + BSC: 21,1 %) um mindestens 5 Prozentpunkte häufiger als im Luspatercept + BSC-Arm auf.

Zusammenfassende Einschätzung zur Sicherheit

Unerwünschte Ereignisse des NCI-CTCAE-Grades ≥ 3 (Luspatercept + BSC: 29,1 %, Placebo + BSC: 15,6 %), SUE (Luspatercept + BSC: 15,2 %, Placebo + BSC: 5,5 %) und UE, die zum Abbruch der Studienmedikation führten (Luspatercept + BSC: 5,4 %, Placebo + BSC: 0,9 %), traten innerhalb der Behandlungs- und Langzeitbehandlungsphase häufiger im Luspatercept + BSC-Behandlungsarm auf als im Placebo + BSC-Behandlungsarm. Über die gesamte Studienperiode hinweg liegen nur deskriptive Auswertungen zu Unerwünschten Ereignissen vor, weshalb der Effekt von Luspatercept auf das Auftreten von Unerwünschten Ereignissen nicht abschließend beurteilt werden kann. Das Verzerrungspotential für den Endpunkt Unerwünschte Ereignisse wird als niedrig eingeschätzt.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen. Die Therapie soll durch eine Ärztin oder einen Arzt mit Erfahrung in der Behandlung von hämatologischen Erkrankungen eingeleitet werden.

6 Zusammenfassung der Nutzenbewertung

Luspatercept ist zugelassen für die Behandlung von erwachsenen Patientinnen und Patienten mit transfusionsabhängiger Anämie, die mit einer β-Thalassämie verbunden ist. Die Nutzenbewertung von Luspatercept basiert auf der zulassungsbegründenden Studie ACE-536-B-THAL-001 (BELIEVE). Es handelt sich bei der Studie ACE-536-B-THAL-001 (BELIEVE) um eine multizentrische, doppelblinde, randomisiert kontrollierte Phase-III-Studie.

Die Ergebnisse der Studie werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Luspatercept ein statistisch signifikant positiver Effekt bei hohem oder unklarem (↑) oder niedrigem (↑↑) Verzerrungspotential bzw. ein statistisch signifikant negativer Effekt bei hohem oder unklarem (↓) oder niedrigem (↓↓) Verzerrungspotential bzw. kein Unterschied (↔) gezeigt werden konnte.

Tabelle 24: Zusammenfassende Darstellung der Ergebnisse der Studie ACE-536-B-THAL-001 (BELIEVE)

Endpunkt	Luspatercept + BSC N = 224	Placebo + BSC N = 112	Effekt
Mortalität			
Gesamtmortalität			n. b.[1)]
Todesfälle, n (%)	1 (0,4)	1 (0,9)	
Morbidität			
Hospitalisierungen			n. b.[2)]
Mindestens eine Hospitalisierung, n (%)	42 (18,8)	4 (3,6)	
Lebensqualität			
TranQoL-Gesamtscore[3)]			↔
Mittelwertdifferenz [95%-KI]; p-Wert[4)]	-0,6 [-3,8; 2,7]; p = 0,73		
SF-36			n. b.[6)]
PCS: Verbesserung um ≥ 5 Punkte, n (%)[5)]	37 (17,6)	11 (10,7)	
MCS: Verbesserung um ≥ 5 Punkte, n (%)[5)]	40 (19,0)	15 (14,6)	n. b.[6)]
Sicherheit			
UE des NCI-CTCAE-Grades ≥ 3, n (%)	65 (29,1)	17 (15,6)	n. b.[2)]
SUE, n (%)	34 (15,2)	6 (5,5)	n. b.[2)]
Therapieabbruch aufgrund von UE, n (%)[7)]	12 (5,4)	1 (0,9)	n. b.[2)]
UE von besonderem Interesse, n (%)[8)]	1 (0,4)	0	n. b.[1)]
Weitere relevante Sicherheitsereignisse			
Thromboembolische und thrombophlebitische Ereignisse, n (%)[9)]	9 (4,0)	1 (0,9)	n. b.[2)]
Knochenschmerzen, n (%)[10)]	44 (19,7)	9 (8,3)	n. b.[2)]

[1)] Die Ereigniszahl ist gering. Die Berechnung relativer Effektschätzer würde keine verlässliche Aussage zulassen.
[2)] Die vom pU im Dossier angegebenen Effektschätzer und p-Werte berücksichtigen nur die Behandlungsphase von Woche 1–48 und damit nicht die gesamte verblindete Studienperiode bestehend aus Behandlungsphase und Langzeitbehandlungsphase. Für die Nutzenbewertung wird die Darstellung der gesamten verblindeten Studienperiode

aufgrund der längeren Behandlungs- und Beobachtungszeit bevorzugt. Die Effektschätzer auf Basis der Behandlungsphase werden daher nicht dargestellt. Zudem wurden die RR unstratifiziert berechnet, d. h. ohne Berücksichtigung der Stratifizierungsvariable Region, welche für die Randomisierung verwendet wurde.

[3] Der TranQoL-Gesamtscore kann Werte zwischen 0 und 100 annehmen. Ein höherer Wert repräsentiert eine höhere Lebensqualität.

[4] Analyse basierend auf dem präspezifizierten ANCOVA (analysis of covariance) Modell mit Adjustierung für den Baselinewert und die Stratifizierungsvariable Region.

[5] Es werden Responderanalysen für Verbesserung der T-Scores um einen Wert von 5 dargestellt. Ein höherer T-Score repräsentiert eine höhere Lebensqualität.

[6] Es liegen keine Auswertungen zum RR unter Berücksichtigung der Stratifizierungsvariable Region vor. Es sind daher keine verlässlichen Aussagen bezüglich des Effektes möglich.

[7] Die Studienteilnehmenden erhielten die Studienmedikation bis zum Auftreten inakzeptabler UE, fehlender Wirksamkeit oder der Rücknahme der Einverständniserklärung, je nachdem, was früher auftrat. Mögliche Therapieabbruchgründe, die vor einem potentiellen Abbruch aufgrund von UE auftreten können, stellen ein konkurrierendes Ereignis dar, weshalb die Ergebnissicherheit und Interpretierbarkeit eingeschränkt ist.

[8] Als UE von besonderem Interesse waren Malignitäten und prämaligne Erkrankungen definiert. Das eine aufgetretene Ereignis wird als erhöhter humaners Choriongonadotropin-(HCG)-Wert beschrieben, der im Rahmen eines falsch-positiven Schwangerschaftstests festgestellt wurde.

[9] Als weiteres relevantes Sicherheitsereignis präspezifizierter SMQ.

[10] Als weiteres relevantes Sicherheitsereignis präspezifizierter Preferred Term.

Abkürzungen: BSC: Best Supportive Care; KI: Konfidenzintervall; MCS: Mental Component Summary; NCI-CTCAE: National Cancer Institute - Common Terminology Criteria for Adverse Events; n. b.: nicht bewertbar; PCS: Physical Component Summary; RR: Relatives Risiko / Risk Ratio; SD: Standardabweichung; SF-36: Short Form 36; (S)UE: (Schwer-wiegende/s) Unerwünschte/s Ereignis/se; TranQoL: Transfusion-dependent Quality of Life Questionare

Referenzen

1. **Acceleron Pharma.** Luspatercept (ACE-536): Study A536-04; a phase 2, open-label, ascending dose study to evaluate the effects of ACE-536 in patients with B-Thalassemia; clincial study report [unveröffentlicht]. 29.08.2018.
2. **Acceleron Pharma.** Luspatercept (ACE-536): Study A536-06; an open-label extension study to evaluate the long-term effects of ACE-536 in patients with B-Thalassemia previously enrolled in study A536-04; clinical study report [unveröffentlicht]. 29.08.2018.
3. **Adib-Hajbaghery M, Poormansouri S, Ahmadi M.** Psychometric Evaluation of the Persian Translation of the Thalassemia-Specific Quality of Life Questionnaire. J Nurs Meas 2017;25(3):202-217.
4. **Celgene.** Dossier zur Nutzenbewertung gemäß § 35a SGB V: Luspatercept (Reblozyl); Erwachsene Patienten mit transfusionsabhängiger Anämie, die mit einer Beta-Thalassämie verbunden ist; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamen Zusatznutzen [unveröffentlicht]. 27.07.2020.
5. **Celgene.** Luspatercept (ACE-536): ACE-536-B-THAL-001; a phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of Luspatercept (ACE-536) versus placebo in adults who require regular red blood cell transfusions due to Beta-Thalassemia; clinical study report [unveröffentlicht]. 05.02.2019.
6. **Celgene.** Luspatercept (ACE-536): ACE-536-B-THAL-001; a phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of Luspatercept (ACE-536) versus placebo in adults who require regular red blood cell transfusions due to Beta-Thalassemia; the "BELIEVE" trial; health-related quality of life report [unveröffentlicht]. 12.09.2019.
7. **Celgene.** Luspatercept (ACE-536): ACE-536-B-THAL-001; a phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of Luspatercept (ACE-536) versus placebo in adults who require regular red blood cell transfusions due to Beta-Thalassemia; the "BELIEVE" trial; post hoc analysis with cut off date of 11 May 2018 (health-related quality of life report) [unveröffentlicht].
8. **Celgene.** Luspatercept (ACE-536): ACE-536-B-THAL-001; a phase 3, double-blind, randomized, placebo-controlled, multicenter study to determine the efficacy and safety of Luspatercept (ACE-536) versus placebo in adults who require regular red blood cell transfusions due to Beta-Thalassemia; the "BELIEVE" trial; post hoc analysis with cutoff date of 11 May 2018 [unveröffentlicht].
9. **European Medicines Agency (EMA).** Reblozyl: European public assessment report EMEA/H/C/004444/0000 [online]. 30.04.2020. Amsterdam (NED): EMA. [Zugriff: 14.08.2020]. URL: https://www.ema.europa.eu/en/documents/assessment-report/reblozyl-epar-public-assessment-report_en.pdf.
10. **Jafari H, Lahsaeizadeh S, Jafari P, Karimi M.** Quality of life in thalassemia major: reliability and validity of the Persian version of the SF-36 questionnaire. J Postgrad Med 2008;54(4):273-275.

11. **Klaassen RJ.** Introducing the Tran Qol: A New Disease-Specific Quality of Life Measure for Children and Adults with Thalassemia Major. Journal of Blood Disorders & Transfusion 2013;04(04).

12. **Klaassen RJ, Barrowman N, Merelles-Pulcini M, Vichinsky EP, Sweeters N, Kirby-Allen M, et al.** Validation and reliability of a disease-specific quality of life measure (the TranQoL) in adults and children with thalassaemia major. Br J Haematol 2014;164(3):431-437.

13. **Klonizakis P, Klaassen R, Sousos N, Liakos A, Tsapas A, Vlachaki E.** Evaluation of the Greek TranQoL: a novel questionnaire for measuring quality of life in transfusion-dependent thalassemia patients. Ann Hematol 2017;96(11):1937-1944.

14. **Maruish ME.** User's manual for the SF-36v2 health survey. 3rd ed. Lincoln: QualityMetric; 2011.