

# 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: Crizanlizumab**

Datum der Veröffentlichung: 1. März 2021

## Inhaltsverzeichnis

Tabellenverzeichnis.....	3
Abbildungsverzeichnis.....	4
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 .....	11
2.3 Endpunkte.....	18
2.3.1 Mortalität.....	19
2.3.2 Morbidität.....	19
2.3.3 Lebensqualität .....	32
2.3.4 Sicherheit.....	34
2.3.5 Übersicht der Erhebungszeitpunkte.....	37
2.4 Statistische Methoden.....	38
2.5 Verzerrungspotential auf Studien- und Endpunktebene.....	42
3 Ergebnisse der eingeschlossenen Studie.....	45
3.1 Studiencharakteristika und Studienmedikation .....	45
3.2 Mortalität .....	51
3.3 Morbidität .....	51
3.4 Lebensqualität.....	59
3.5 Sicherheit.....	59
4 Diskussion der Methodik und Ergebnisse.....	63
4.1 Zulassungsstatus und Zulassungspopulation von Crizanlizumab.....	63
4.2 Design und Methodik der Studie .....	64
4.3 Mortalität .....	67
4.4 Morbidität .....	67
4.5 Lebensqualität.....	72
4.6 Sicherheit.....	72
5 Anforderung an eine qualitätsgesicherte Anwendung.....	74
6 Zusammenfassung der Nutzenbewertung .....	75
Referenzen.....	77
Anhang .....	80

## Tabellenverzeichnis

Tabelle 1:	Übersicht über die Studienbasis .....	9
Tabelle 2:	Charakterisierung der Studie SUSTAIN .....	11
Tabelle 3:	Charakterisierung der Intervention in der Studie SUSTAIN.....	15
Tabelle 5:	Zusammenfassung der Endpunktbewertung .....	18
Tabelle 6:	Subklassifizierungen von vasookklusiven Schmerzkrisenereignissen .....	21
Tabelle 7:	Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie SUSTAIN .....	37
Tabelle 8:	Verzerrungspotential der Studie SUSTAIN .....	42
Tabelle 9:	Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie SUSTAIN.....	43
Tabelle 10:	Allgemeine Angaben der Studie SUSTAIN .....	45
Tabelle 11:	Charakterisierung der Studienpopulation der Studie SUSTAIN – ITT-Population ....	47
Tabelle 12:	Medizinische Vorgeschichte und Komorbiditäten mit $\geq 10\%$ Ereignisse nach MedDRA-Systemorganklassen und ausgewählten Preferred Terms zu Baseline der Studie SUSTAIN – ITT-Population .....	49
Tabelle 13:	Begleitmedikation in der Studie SUSTAIN – ITT-Population .....	50
Tabelle 14:	Nach Behandlungsbeginn aufgetretene VOC – Sicherheitspopulation.....	52
Tabelle 15:	Primäre Analyse des Endpunkts „Jährliche Rate VOC“ und die berücksichtigte Anzahl und Subklassifikation der VOC gemäß CRC – ITT-Population .....	53
Tabelle 16:	Sensitivitätsanalysen zu VOC basierend auf Negativ-Binomial- Regression – ITT-Population .....	54
Tabelle 17:	Analysen zu den Endpunkten „Zeit bis zur ersten VOC“ und „Zeit bis zur zweiten VOC“ – ITT-Population .....	55
Tabelle 18:	Primäre Analyse und Sensitivitätsanalyse der Subklassifikation „Jährliche Rate unkomplizierte VOC“ – ITT-Population .....	57
Tabelle 19:	Primäre Analyse der Subklassifikation „Jährliche Rate akute Thoraxsyndrome“ – ITT-Population .....	58
Tabelle 20:	Rücklaufquoten des BPI-LF in der Studie SUSTAIN – ITT-Population .....	58
Tabelle 21:	Gesamtraten unerwünschter Ereignisse – Sicherheitspopulation .....	59
Tabelle 22:	UE mit Inzidenz $\geq 10\%$ und schwere UE mit Inzidenz $\geq 5\%$ nach MedDRA-Systemorganklassen und Preferred Terms in der Studie SUSTAIN – Sicherheitspopulation .....	60
Tabelle 23:	SUE mit Inzidenz $\geq 5\%$ nach MedDRA-Systemorganklassen und Preferred Terms in der Studie SUSTAIN – Sicherheitspopulation .....	61
Tabelle 24:	UE von besonderem Interesse .....	62
Tabelle 25:	Zusammenfassende Darstellung der Ergebnisse der Studie SUSTAIN .....	75
Tabelle 26:	Übersicht zu Sensitivitätsanalysen des primären Endpunkts .....	80
Tabelle 27:	Rücklaufquoten des SF-36 in der Studie SUSTAIN – ITT-Population .....	81
Tabelle 28:	Ergebnisse zu den Summenscores des SF-36 in der Studie SUSTAIN – ITT-Population .....	82
Tabelle 29:	Ergebnisse zu den Subskalen des SF-36 in der Studie SUSTAIN – ITT-Population .....	83

## **Abbildungsverzeichnis**

Abbildung 1: Studienablauf .....	11
Abbildung 2: Kaplan Meier-Kurve für die Zeit bis zur ersten VOC (basierend auf primärer Analyse) .....	56

## Abkürzungsverzeichnis

ACS	Akutes Thoraxsyndrom
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ATC	Anatomisch-therapeutisch-chemisches Klassifikationssystem
BPI-LF	Brief Pain Inventory – Long Form
CMA	Conditional Marketing Authorisation
CRF	Case Report Form
CRC	Crisis Review Committee
DMC	Data Monitoring Committee
eCRF	electronic Case Report Form
EMA	European Medicines Agency
EPAR	European Public Assessment Report
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
GKV	Gesetzliche Krankenversicherung
HL	Hodges-Lehmann
HR	Hazard Ratio
HU	Hydroxyurea/Hydroxycarbamid
ITT	Intention-to-Treat
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IXRS	Interactive Voice/Web Response System
KI	Konfidenzintervall
LF	Long Form
M	Imputationsmodell x
MCS	Psychischer Summenscore
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modifizierte ITT-Population
MMRM	Gemischtes Modell mit wiederholten Messungen
NSAID	nichtsteroidale Antirheumatika
PCS	Körperlicher Summenscore
PD	Pharmakodynamik
PhS	Post-hoc-Sensitivitätsanalyse
PP	Per Protocol
PS	Präspezifizierte Sensitivitätsanalyse
pU	pharmazeutischer Unternehmer
RR	Relatives Risiko

SAP	Statistischer Analyseplan
SF-36	36-item Short Form Health Survey
SGB	Sozialgesetzbuch
SUE	Schwerwiegendes unerwünschtes Ereignis
TEAE	Behandlungsbezogenes unerwünschtes Ereignis
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des G-BA
VOC	Vasookklusive Schmerzkrise

## Hintergrund

Crizanlizumab 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 Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Crizanlizumab zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers (pU) heran und bewertet die darin enthaltenen Daten nach Maßgabe der 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 Crizanlizumab in seiner Sitzung am 23. Februar 2021 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 27. November 2020 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

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

Die Nutzenbewertung wird am 1. März 2021 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

Crizanlizumab (Adakveo®) 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. Die Bewertung bezieht sich gemäß Zulassungsstatus auf folgende Indikation [9]:

Prävention wiederkehrender vasookklusiver Krisen bei Patienten ab 16 Jahren mit Sichelzellerkrankheit.

Die Zulassung durch die European Medicines Agency (EMA) besteht seit dem 28. Oktober 2020 [3]. Crizanlizumab kann als Zusatztherapie zu Hydroxyurea/Hydroxycarbamid (HU) gegeben werden oder als Monotherapie bei Personen, bei denen die Anwendung von HU nicht geeignet oder unzureichend ist. Die empfohlene Dosis beträgt 5,0 mg/kg Körpergewicht, verabreicht über einen Zeitraum von 30 Minuten als intravenöse Infusion in Woche 0, Woche 2 und danach alle 4 Wochen. Es wird empfohlen im Falle einer schweren infusionsbedingten Reaktion Crizanlizumab abzusetzen.

Crizanlizumab wurde ursprünglich von Selexys Pharmaceuticals Corporation (bzw. Reprixys Pharmaceuticals) unter dem Namen „SelG1“ entwickelt und wird nach Übernahme durch Novartis in veränderter Form unter dem Namen „SEG101“ vertrieben [10].

## 2 Auswahl und Methodik der relevanten Studien

### 2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studiename	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Gründe, warum Studie nicht relevant für die Nutzenbewertung
<b>Studien zum Wirkstoff</b>				
SelG1-00005/ CSEG101A2201 (SUSTAIN) <sup>1)</sup>	Ja	Ja	Ja	
SelG1-00003/ CSEG101A2101 <sup>2)</sup>	Ja	Nein	Nein	Andere Population (gesunde Personen)
CSEG101A2102 <sup>3)</sup>	Ja	Nein	Nein	Andere Population (gesunde Personen)
CSEG101A2202 (SOLACE-adults) <sup>4)</sup>	Ja <sup>5)</sup>	Nein	Nein	<ul style="list-style-type: none"> <li>Unkontrollierte Studie</li> <li>Geplantes Ende: 2023</li> </ul>
CSEG101A2203 (STEADFAST) <sup>6)</sup>	Nein	Nein	Nein	<ul style="list-style-type: none"> <li>Andere Indikation bzw. Komorbidität</li> <li>Geplantes Ende: 2023</li> </ul>
CSEG101A2301 (STAND) <sup>7)</sup>	Ja <sup>8)</sup>	Nein	Nein	<ul style="list-style-type: none"> <li>Noch laufend und keine Ergebnisse verfügbar.</li> <li>Geplantes Ende: 2027</li> </ul>
CSEG101AUS05 (SPARTAN) <sup>9)</sup>	Nein	Nein	Nein	<ul style="list-style-type: none"> <li>Einarmige Studie</li> <li>Subpopulation (männliche Patienten mit Priapismus)</li> <li>Geplantes Ende: 2022</li> </ul>
CSEG101B2201 (SOLACE-kids) <sup>10)</sup>	Ja <sup>8)</sup>	Nein	Nein	Andere Population (< 18 Jahren)

<sup>1)</sup> Zulassungsrelevante Studie gemäß EPAR [26].

<sup>2)</sup> Abgeschlossene monozentrische, doppelblinde, placebokontrollierte Phase-I-Studie zur Untersuchung der Pharmakokinetik und Pharmakodynamik, Immunogenität und Sicherheit von ansteigenden Dosen SelG1 (Crizanlizumab) bei gesunden Personen. [27]

<sup>3)</sup> Abgeschlossene monozentrische, offene Phase-I-Studie mit Einzeldosisgabe und Parallelgruppendesign zur Untersuchung der Vergleichbarkeit von SEG101 (Crizanlizumab) und SelG1 (Crizanlizumab) hinsichtlich Pharmakokinetik und Pharmakodynamik bei gesunden Personen. [15]

<sup>4)</sup> Laufende multizentrische, offene, unkontrollierte Phase-II-Studie zur Untersuchung der Pharmakokinetik und Pharmakodynamik (sowie Wirksamkeit und Sicherheit) von zwei Dosen (5,0 oder 7,5 mg/kg KG) SEG101 (Crizanlizumab) mit oder ohne HU/HC bei Patientinnen und Patienten von 16 bis 70 Jahren mit Sichelzellerkrankung und historischen vasookklusiven Schmerzkrisen. [13]

<sup>5)</sup> Zu dieser Studie liegen erste Ergebnisse in Form von Interimsanalysen vor.

<sup>6)</sup> Laufende, multizentrische, offene, Standardtherapie-kontrollierte Phase-II-Studie zur Untersuchung der Wirksamkeit und Sicherheit von SEG101 (Crizanlizumab) mit Standardtherapie bei Patientinnen und Patienten ab 16 Jahren mit chronischer Niereninsuffizienz aufgrund von Sichelzellerkrankung. [17]

<sup>7)</sup> Laufende multizentrische, doppelblinde, placebokontrollierte Phase-III-Studie zur Untersuchung und dem Vergleich von zwei Dosen (5,0 oder 7,5 mg/kg KG) SEG101 (Crizanlizumab) mit oder ohne HU/HC hinsichtlich Wirksamkeit und Sicherheit bei Patientinnen und Patienten ab 12 Jahren mit Sichelzellerkrankung und historischen vasookklusiven Schmerzkrisen. [19,21]

<sup>8)</sup> Zu dieser Studie liegt bisher ausschließlich das Studienprotokoll vor.

<sup>9)</sup> Laufende multizentrische, offene, einarmige Phase-II-Studie zur Untersuchung der Wirksamkeit und Sicherheit von SEG101 (Crizanlizumab) bei männlichen Patienten ab 16 Jahren mit Sichelzellerkrankung und Priapismus. [20]

<sup>10)</sup> Laufende, multizentrische, offene, unkontrollierte Phase-II-Studie zur Dosisfindung und Untersuchung der Sicherheit von SEG101 (Crizanlizumab) mit oder ohne HU/HC bei pädiatrischen Patientinnen und Patienten bis 18 Jahren mit Sichelzellerkrankung und historischen vasookklusiven Schmerzkrisen. [14]

Abkürzungen: EPAR: European Public Assessment Report; HC: Hydroxycarbamid; HU: Hydroxyurea; KG: Körpergewicht; pU: pharmazeutischer Unternehmer.

Die pivotale Studie SelG1-00005/CSEG101A2201 (SUSTAIN) untersucht die Wirksamkeit und Sicherheit von SelG1 (Crizanlizumab) bei Patientinnen und Patienten mit Sichelzellerkrankung und mindestens zwei vorausgegangenen vasookklusiven Schmerzkrisen (VOC) während der vergangenen 12 Monate. Bei SelG1 handelt es sich um den ursprünglich von Selexys Pharmaceuticals Corporation (bzw. Reprixys Pharmaceuticals) entwickelten Wirkstoff, der anschließend von Novartis verändert wurde und nun unter der Bezeichnung SEG101 vertrieben wird. Der neue Wirkstoff wird im vorliegenden Anwendungsgebiet zurzeit in einer laufenden unkontrollierten Studie, CSEG101A2202 (SOLACE-adults) sowie einer laufenden Phase-III-Studie, CSEG101A2301 (STAND), untersucht. Während zu SOLACE-adults mit einem geplanten Studienende im Jahr 2023 vorläufige Ergebnisse anhand von Interimsanalysen vorliegen, standen für die Nutzenbewertung noch keine Ergebnisse zur placebokontrollierten Studie STAND zur Verfügung. Als Teil der Conditional Marketing Authorisation (CMA) von Crizanlizumab erwartet die EMA vom pU abschließende Ergebnisse beider Studien, d. h. die primäre Analyse der STAND-Studie und den Studienbericht der Studie SOLACE-adults, bis zum Dezember 2025.

Zur Einschätzung der Pharmakokinetik und Pharmakodynamik (PD) von Crizanlizumab hat die EMA [3] Ergebnisse von vier Studien herangezogen, anhand derer eine Pharmakokinetik-Populationsanalyse durchgeführt wurde: Studie SelG1-00003/CSEG101A2101 mit gesunden Personen (SelG1), die pivotale Studie SUSTAIN bei Patientinnen und Patienten mit Sichelzellerkrankung (SelG1), die Studie CSEG101A2102 zum Vergleich von SelG1 und SEG101 bei gesunden Personen sowie SOLACE-adults bei Patientinnen und Patienten mit Sichelzellerkrankung (SEG101). Bezüglich der Untersuchung der Bioäquivalenz wurden zwischen den beiden Crizanlizumab-Versionen SelG1 und SEG101 Unterschiede hinsichtlich Pharmakokinetik und Pharmakodynamik, insbesondere hinsichtlich einer höheren Exposition und höheren prozentualen Inhibition von SEG101 gegenüber dem ursprünglichen Wirkstoff SelG1, beobachtet. Es wurden daher durch die EMA Reanalysen von verfügbaren Laborproben veranlasst. Unter Berücksichtigung der Ergebnisse dieser Reanalysen sowie allen der EMA zur Verfügung gestellten Daten zu Pharmakologie, Wirksamkeit und Sicherheit geht diese davon aus, dass die Differenz der prozentualen P-Selektin-inhibition zwischen den beiden Wirkstoffen höchstwahrscheinlich auf die diagnostische Nutzung unterschiedlicher PD-Assays sowie der Verschiebung der Standardtitrationskurve zurückgeht. Die EMA schätzt daher die Pharmakodynamik der beiden Wirkstoffe als vergleichbar ein und sieht von einer weiteren Untersuchung der Unterschiede in Pharmakokinetik ab [3].

Basierend auf den Einschätzungen der EMA zur Vergleichbarkeit der beiden Wirkstoffe SelG1 und SEG101 sowie dem Vorliegen nur vorläufiger Evidenz (Interimsanalysen) aus der unkontrollierten Studie SOLACE-adults zum neuen Wirkstoff SEG101 wird diese nicht zur Nutzenbewertung herangezogen. Die abgeschlossene multizentrische, doppelblinde, placebokontrollierte Phase-II-Studie SUSTAIN, welche höherwertige Evidenz zur Beurteilung des Zusatznutzens von Crizanlizumab aufweist, ist folglich ausschließlich Gegenstand der vorliegenden Nutzenbewertung.

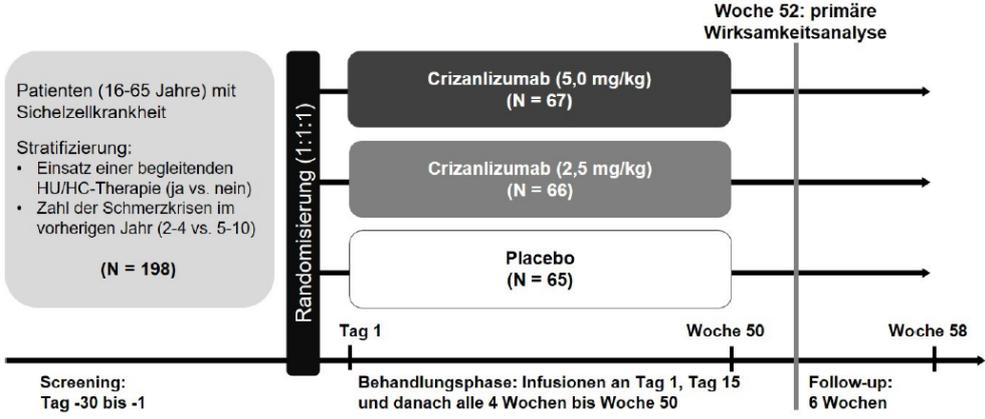
#### **Zur Nutzenbewertung für Adakveo® wurden folgende Studien und Daten herangezogen:**

- Herstellerdossier zu Crizanlizumab [10]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [3,9,11,12,16,18,19,25]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie SelG1-00005 (SUSTAIN) [26]

## 2.2 Aufbau und Design der eingeschlossenen Studie

Die Unterlagen zur Nutzenbewertung für Crizanlizumab basieren auf der Zulassungsstudie SelG1-00005 (SUSTAIN) [26]. Der in SUSTAIN genutzte Wirkstoff SelG1 wird im Folgenden mit Crizanlizumab bezeichnet. Die Studie und die Intervention werden in Tabelle 2 charakterisiert.

Tabelle 2: Charakterisierung der Studie SUSTAIN

Charakteristikum	Beschreibung
Design	<p>Bei der Studie SUSTAIN handelt es sich eine randomisierte, multizentrische, doppelblinde, placebokontrollierte Phase-II-Studie zur Bewertung der Wirksamkeit und Sicherheit zweier Dosierungen des Wirkstoffs Crizanlizumab mit Parallelgruppendesign. Im Zuge der Studie wurde Crizanlizumab mit oder ohne HU-Behandlung (gegenüber Placebo mit oder ohne HU-Behandlung bei Patientinnen und Patienten von 16 bis 65 Jahren mit Sichelzellerkrankheit und mit zwei bis zehn VOC während der vergangenen 12 Monate verglichen. Das primäre Studienziel war die Untersuchung der Sicherheit und Verträglichkeit von jeweils einer Hoch- und Niedrigdosierung Crizanlizumab sowie deren Wirksamkeit gegenüber Placebo nach 52 Wochen. Der Studienablauf gliederte sich in drei Phasen mit Visiten (Screening, Behandlung und Follow-up<sup>1)</sup>), welche in Abbildung 1 dargestellt sind. Ab Beginn der Behandlungsphase wiesen alle Visiten-terminierungen eine Toleranz von <math>\pm 3</math> Tagen auf. Die Behandlungsphase umfasste 50 Wochen mit Behandlungen und eine Behandlungsabschlussvisite in Woche 52.</p>  <p>Abbildung 1: Studienablauf [10]<sup>1)</sup></p> <p>Die untersuchten Dosierungen umfassen 5,0 mg/kg und 2,5 mg/kg, welche in zwei parallelen Behandlungsarmen angewendet wurden. Die Zuteilung der Patientinnen und Patienten zu den zwei Dosierungen sowie dem Placebo-Arm erfolgte gemäß Studienbericht i. d. R. nach Abschluss der Screeningvisite/n an Tag 1 der Behandlungsphase in einer 1:1:1-Randomisierung (siehe Kapitel 2.2 – Protokolländerungen). Diese war stratifiziert nach Begleitbehandlung mit HU (ja vs. nein) und der Anzahl an VOC während der vergangenen 12 Monate (dichotomisiert zu 2–4 bzw. 5–10 VOC). Es wurden 329 Personen für den Studieneinschluss evaluiert und 198 Personen in die Studie eingeschlossen, wovon 133 auf die für die Bewertung relevanten Arme randomisiert wurden. Es wurden 118 Personen aufgrund der Ein- und Ausschlusskriterien sowie 13 Personen wegen Teilnahmeablehnung nicht in die Studie eingeschlossen.</p> <p>Die Behandlung erfolgte bis zum Auftreten von UE, die zum Abbruch führten, von inakzeptablen QTc-Intervallerhöhungen, die zum Abbruch führten, von Schwangerschaft, bis zum Widerruf der Einverständniserklärung, bis zum Tod, bis zu einem anderweitig begründeten Studienabbruch oder bis zur Beendigung der 50-wöchigen Behandlung. Der Anteil der Personen mit HU wurde bei Einschluss auf 65 % limitiert. Wenn möglich sollte die Behandlungsabschlussvisite (i. d. R. in Woche 52) bei Behandlungsabbruch</p>

Charakteristikum	Beschreibung
	so zeitnah wie möglich durchgeführt werden. Zudem sollte die Follow-up-Visite (i. d. R. in Woche 58) ungefähr 8 Wochen nach der letzten Dosisgabe erfolgen.
<b>Population</b>	<p><b>Wesentliche Einschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Männer und Frauen im Alter zwischen 16 und 65 Jahren (die untere Altersgrenze wurde im Verlauf der Studie geändert, siehe Kapitel 2.2 – Protokolländerungen).</li> <li>• Eine bestätigte Diagnose einer Sichelzellerkrankung (einschließlich der Genotypen HbSS, HbSC, HbSβ0 und HbSβ+).</li> <li>• Negativer Serum-Schwangerschaftstest in der Screeningphase sowie negativer Urin-Schwangerschaftstest vor der Randomisierung und der ersten Studienmedikationsgabe.</li> <li>• Bei einer gleichzeitigen Therapie mit HU musste diese für mindestens 6 Monate vor der ersten Studienmedikationsgabe verschrieben worden sein sowie mit einer stabilen Dosis für mindestens 3 Monate vor der ersten Studienmedikationsgabe vorliegen. Etwaige Dosisänderungen und/oder ein Behandlungsabbruch von HU waren (außer aus Sicherheitsgründen) nicht vorgesehen.<sup>2)</sup></li> <li>• Bei einer gleichzeitigen Therapie mit Erythropoetin musste diese für mindestens 6 Monate vor der ersten Studienmedikationsgabe verschrieben worden sein sowie mit einer stabilen Dosis für mindestens 3 Monate vor der ersten Studienmedikationsgabe vorliegen.<sup>2)</sup></li> <li>• Auftreten von 2 bis 10 VOC in den vorangegangenen 12 Monaten gemäß der medizinischen Vorgeschichte oder der Erinnerungen der Patientin / des Patienten. Die Schmerzkrisen sollten den Kriterien gemäß dem primären Endpunkt entsprechen.<sup>2)</sup></li> <li>• Thoraxröntgenbild ohne akut pathologische Befunde, wobei das letzte Thoraxröntgenbild am Tag der ersten Studienmedikationsgabe nicht älter als 3 Monate sein durfte.</li> <li>• Klinisch akzeptable/s<sup>3)</sup> <ul style="list-style-type: none"> <li>○ 12-Kanal-EKG beim Screening,</li> <li>○ medizinische Vorgeschichte,</li> <li>○ körperliche Untersuchung,</li> <li>○ Vitalparameter und</li> <li>○ Labortests.</li> </ul> </li> </ul> <p><b>Wesentliche Ausschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Teilnahme an einem chronischen Transfusionsprogramm, definiert als eine Reihe von geplanten (einfachen oder Austausch-) Transfusionen als Prophylaxe-Maßnahme, oder Hämoglobin A &gt; 20 % des Gesamthämoglobins<sup>2)</sup>.</li> <li>• Geplante Austauschtransfusion während des Studienzeitraums<sup>2)</sup>.</li> <li>• Hämoglobin &lt; 4,0 g/dl.</li> <li>• Geplanter großer chirurgischer Eingriff während des Studienzeitraums;</li> <li>• Patient/in plant, eine Therapie mit HU während der Studienphase zu beginnen oder zu beenden bzw. die Dosis zu verändern (Dosisanpassung aus Sicherheitsgründen erlaubt)<sup>2)</sup>.</li> <li>• Chronische Therapie mit Antikoagulantien außer Aspirin (z. B. Warfarin oder Heparin)<sup>2)</sup>.</li> <li>• Bestehen von signifikanten aktiven und schlecht kontrollierten kardiovaskulären, neurologischen, endokrinen, hepatischen oder renalen Erkrankungen (einschließlich atriale oder ventrikuläre Arrhythmie), die eindeutig nicht im Zusammenhang mit der Sichelzellerkrankung stehen. Folgende Laborbefunde führten zum Ausschluss: <ul style="list-style-type: none"> <li>○ Kreatinin ≥ 1,2 mg/dl;</li> <li>○ Direktes Bilirubin ≥ 2,0 mg/dl;</li> <li>○ ALT ≥ 3 x ULN.</li> </ul> </li> <li>• Diagnose einer Krebserkrankung (außer nichtmelanozytärer Hautkrebs und Zervixkarzinom in situ) innerhalb der letzten 5 Jahre.</li> <li>• Schlaganfall innerhalb der letzten 2 Jahre.</li> </ul>

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> <li>• Positiver Urintest beim Screening für Kokain, Phencyclidin oder Amphetamine.</li> <li>• Schwerwiegende physische oder mentale Erkrankung (einschließlich Psychose), die nach Einschätzung des ärztlichen Prüfpersonals die Teilnahme an der Studie beeinträchtigen würde (z. B. eingeschränkte mentale Leistungsfähigkeit, Alkoholismus).</li> <li>• Bestehen einer anderen Erkrankung, die nach Einschätzung des ärztlichen Prüfpersonals eine sichere Teilnahme an der Studie beeinträchtigen bzw. mit der Interpretation der Studienergebnisse interferieren könnte.</li> </ul>
<b>Intervention und Zahl der Patientinnen und Patienten</b>	<p>Die Hochdosis-Interventionsgruppe mit 5,0 mg/kg entspricht der Dosierung von Crizanlizumab gemäß Fachinformation.</p> <ul style="list-style-type: none"> <li>• ITT-Population<sup>4)</sup> Hochdosis-Interventionsgruppe mit 5,0 mg/kg Crizanlizumab: N = 67 Niedrigdosis-Interventionsgruppe mit 2,5 mg/kg Crizanlizumab: N = 66 Placebo: N = 65</li> <li>• Sicherheitspopulation<sup>5)</sup> Hochdosis-Interventionsgruppe mit 5,0 mg/kg Crizanlizumab: N = 66 Niedrigdosis-Interventionsgruppe mit 2,5 mg/kg Crizanlizumab: N = 64 Placebo: N = 62</li> </ul>
<b>Ort und Zeitraum der Durchführung</b>	<p><b>Ort</b> 60 Studienzentren in 3 Ländern: USA (n = 51), Brasilien (n = 8), Jamaika (n = 1).</p> <p><b>Studienzeitraum</b> Die Einverständniserklärung der/des ersten Patientin/Patienten lag zum 23.07.2013 vor, während die/der letzte Patientin/Patient die Studie am 23.03.2016 beendet hat. Im Studienprotokoll wurde a priori eine Gesamtstudiendauer von 25 Monaten mit 12 Monaten Rekrutierungs- sowie 13 Monaten Behandlungs- und Follow-up-Phase geplant. Die tatsächliche Gesamtstudiendauer vom ersten Studieneinschluss bis zum Studienabschluss belief sich auf etwa 32 Monate:</p> <ul style="list-style-type: none"> <li>• Studieneinschluss erste/r Patient/in: 24.07.2013</li> <li>• Studienabschluss letzte/r Patient/in: 23.03.2016</li> </ul> <p>Nach Angaben des pU gegenüber der EMA wurde die Datenbank am 13.05.2016 gesperrt [3]. Es liegen die Auswertungen der Studienergebnisse vom 20.06.2016 im Rahmen des Studienberichtes vom 14.10.2016 vor. Zudem wurden sowohl Zusatzanalysen vom 27.07.2020, 31.07.2020 und 04.08.2020 als auch ein Post-hoc-Studienbericht vom 28.10.2020 eingereicht. Die Zusatzanalysen beinhalten u. a. Post-hoc-Analysen zu „Jährliche Rate VOC“, „Jährliche Rate unkomplizierte VOC“, „Zeit bis zur zweiten VOC“, „Anteil der Personen ohne VOC während der Behandlungsphase“ sowie zu Subgruppenanalysen und UE (jeglicher Art).</p>
<b>Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll</b>	<p>Die Endpunkte werden hier gemäß der letzten Protokollversion (Protokoll 4) aufgeführt, welches nach Studienende Gültigkeit erlangte (siehe Kapitel 2.2 – Protokolländerungen).</p> <p><b>Primärer Wirksamkeitsendpunkt</b> Jährliche Rate VOC</p> <p><b>Sekundäre Wirksamkeitsendpunkte</b></p> <ul style="list-style-type: none"> <li>• Jährliche Rate der Krankenhaustage</li> <li>• Zeit bis zur ersten VOC</li> <li>• Zeit bis zur zweiten VOC</li> <li>• Jährliche Rate unkomplizierte VOC</li> <li>• Jährliche Rate akute Thoraxsyndrome</li> <li>• BPI-LF</li> </ul>

Charakteristikum	Beschreibung
	<p><b>Explorative Endpunkte</b></p> <ul style="list-style-type: none"> <li>• Jährliche Rate der transfundierten Erythrozytenkonzentrate</li> <li>• Jährliche Rate der Schmerzkrisentage<sup>6)</sup></li> <li>• Veränderung der Retikulozyten-Zahl gegenüber Studienbeginn</li> <li>• Veränderung des Haptoglobin-Werts gegenüber Studienbeginn</li> <li>• Veränderung des indirekten Bilirubins gegenüber Studienbeginn</li> <li>• SF-36</li> <li>• Veränderung des Hämoglobin-Werts gegenüber Studienbeginn</li> <li>• Veränderung des LDH-Werts gegenüber Studienbeginn</li> </ul> <p><b>Weitere Endpunkte</b></p> <ul style="list-style-type: none"> <li>• Biomarker (z. B. lösliches P-Selektin)</li> <li>• Pharmakokinetik und Pharmakodynamik</li> </ul> <p><b>Sicherheitsendpunkte</b></p> <ul style="list-style-type: none"> <li>• UE</li> <li>• Klinische Laborparameter</li> <li>• Weitere Laboruntersuchungen</li> <li>• Körperliche Untersuchung und Vitalzeichen</li> <li>• 12-Kanal-EKG</li> <li>• Immunogenität (Test auf Anti-Crizanlizumab-Antikörper)</li> <li>• Schwangerschaftstest</li> </ul> <p><b>Post hoc im Dossier zusätzlich berichtete Endpunkte</b></p> <ul style="list-style-type: none"> <li>• Jährliche Rate der Krankenhaustage aufgrund VOC</li> <li>• Anteil der Personen ohne VOC während der Behandlungsphase</li> </ul>
<p><b>Subgruppenanalysen</b></p>	<p>Nach Abschluss der Studie wurden im finalen Protokoll 4 und im SAP folgende Subgruppenanalysen geplant, die im Studienbericht dargestellt wurden:</p> <ul style="list-style-type: none"> <li>• Jährliche Rate VOC <ul style="list-style-type: none"> <li>○ nach HU-Therapie (ja vs. nein)<sup>7)</sup></li> <li>○ nach kategorisierter Anzahl vorheriger VOC (2–4 vs. 5–10)</li> <li>○ nach Sichelzellerkrankungstyp (HbSS vs. alle Genotypen ausschließlich HbSS)</li> </ul> </li> <li>• Deskriptive Darstellung TEAE <ul style="list-style-type: none"> <li>○ nach HU-Therapie</li> <li>○ nach kategorisierter Anzahl vorheriger VOC</li> </ul> </li> </ul> <p>Zusätzlich wurde im Studienbericht folgende Subgruppe deskriptiv dargestellt<sup>8)</sup>:</p> <ul style="list-style-type: none"> <li>• Jährliche Rate VOC nach HU-Therapie (ja vs. nein) und kategorisierter Anzahl vorheriger VOC (2–4 vs. 5–10)</li> </ul>

<sup>1)</sup> Die Follow-up-Visite wurde in Protokoll 2 von Woche 56 nach Randomisierung auf Woche 58 nach Randomisierung geändert. Bis dahin wurde eine unbestimmte Zahl an Personen in die Studie eingeschlossen, die jedoch nach Angabe des Ersteintrags (24.07.2013) zum Zeitpunkt der Protokollveröffentlichung (07.04.2014) noch keine der beiden Follow-up-Phasendefinitionen erreicht haben sollten.

<sup>2)</sup> Studien- und Zulassungspopulation stimmen in diesem potentiell relevanten Fall (z. B. vorherige Schmerzkrisen, Therapieoption u. a.) nicht vollständig überein.

<sup>3)</sup> Eine Präzisierung der Begrifflichkeit „klinischer Akzeptabilität“ konnte nicht identifiziert werden.

<sup>4)</sup> Alle randomisierten Patientinnen und Patienten.

<sup>5)</sup> Alle randomisierten Patientinnen und Patienten, die mindestens eine Dosis der jeweiligen Studienmedikation erhalten haben.

<sup>6)</sup> Der Endpunkt „Jährliche Rate der Schmerzkrisentage“ wurde erstmals nach Studienende in Protokoll 4 definiert (siehe Kapitel 2.2 – Protokolländerungen).

<sup>7)</sup> Die Subgruppenanalyse wurde in Protokoll 3 vor Studienende präspezifiziert.

<sup>8)</sup> Weitere Post-hoc-Subgruppenanalysen wurden u. a. der EMA vorgelegt und sind nicht dargestellt [3].

Abkürzungen: ALT: Alaninaminotransferase; BPI-LF: Brief Pain Inventory – Long Form; EMA: European Medicines Agency; HU: Hydroxyurea/Hydroxycarbamid; ITT: Intention-to-Treat; LDH: Laktatdehydrogenase; pU: pharmazeutischer Unternehmer; SAP: Statistischer Analyseplan; SF-36: 36-item Short Form Health Survey; TEAE: Behandlungsbezogene unerwünschte Ereignisse; UE: Unerwünschtes Ereignis; ULN: Upper Limit of Normal; VOC: Vasookklusive Schmerzkrise.

Bei der Hochdosis mit 5,0 mg/kg Crizanlizumab (siehe Tabellen 2 und 3) handelt es sich um die Dosierung gemäß Fachinformation, worauf sich demnach die vorliegende Nutzenbewertung stützt. Da die Niedrigdosis mit 2,5 mg/kg Crizanlizumab nicht der zugelassenen Dosierung entspricht, werden diesbezügliche Informationen in den folgenden Kapiteln, Tabellen und Abbildungen nicht adressiert.

Tabelle 3: Charakterisierung der Intervention in der Studie SUSTAIN

Intervention	Kontrolle
<b>Crizanlizumab</b> Hochdosisgabe von 5,0 mg/kg KG (100ml-Infusion) <sup>1)</sup>	<b>Placebo</b> Sterile 0,9 % Natriumchlorid-Lösung (100ml-Infusion)
<p><b>Art und Dauer der Studienmedikationsgabe</b>  Die intravenöse Verabreichung von zwei Aufsättigungsdosen (5,0 mg/kg) erfolgte an Tag 1 und Tag 15 (<math>\pm</math> 3 Tage) mit einer anschließenden Erhaltungsdosisgabe alle 4 Wochen (<math>\pm</math> 3 Tage) bis zur letzten, planmäßig 14. Dosisgabe in Woche 50 oder bis zum Abbruch der Behandlung.</p> <p><b>Verabreichungsmodifikation oder -abbruch<sup>2)</sup></b>  Bei Auftreten einer unerwünschten Reaktion während der Verabreichung der Studienmedikation können die Infusionsrate verringert oder gestoppt sowie nach Einschätzung der Ärztin / des Arztes geeignete unterstützende Maßnahmen eingeleitet werden. Bei Verringerung der Infusionsrate soll die Gesamtdauer der Infusion 2 Stunde nicht überschreiten.</p> <p><b>Regeln zur Unterbrechung und Wiederaufnahme des Behandlungsplanes</b>  Unterbrechungen des Behandlungsplans waren aufgrund von Hospitalisierung, Nichteinverständnis, eines positiven Urin-Schwangerschaftstests oder aus anderen (in den Studienunterlagen nicht definierten) Gründen möglich. Bei positivem Urin-Schwangerschaftstest soll nach etwaigem Vorliegen eines negativen Serum-Schwangerschaftstests die Behandlung fortgesetzt werden. Bei anderen Gründen zur Unterbrechung des Behandlungsplans sollte bei Zustimmung des ärztlichen Prüfpersonals eine Wiederaufnahme der Behandlung folgendermaßen gehandhabt werden:</p> <ul style="list-style-type: none"> <li>• Bei einer Abweichung von mehr als 3 Tagen vom Dosisgabetermin sollte innerhalb von 2 Wochen vom ursprünglichen Dosisgabetermin die verpasste Dosisgabe erfolgen und anschließend die normale Dosisgabenplanung wieder zum nächsten vereinbarten Dosisgabetermin aufgenommen werden.</li> <li>• Bei einer Abweichung von mehr als 2 Wochen vom Dosisgabetermin soll die verpasste Dosisgabe übersprungen werden und die Dosisgabe wieder zum nächsten vereinbarten Dosisgabetermin aufgenommen werden.</li> </ul> <p>Zwischen zwei konsekutiven Dosisgaben sollten mindestens 2 Wochen liegen.</p> <p><b>Studienabbruch</b>  <u>Mögliche Gründe</u></p> <ul style="list-style-type: none"> <li>• Wunsch der Patientin / des Patienten bzw. des/der Bevollmächtigten;</li> <li>• UE;</li> <li>• Patientin/Patient offenbart einen fehlenden Willen oder die fehlende Fähigkeit die geplante Studienmedikationsgabe fortzuführen bzw. regelhaft zu beenden;</li> <li>• Lost to Follow-up;</li> <li>• Studienmedikationsabbruch nach Einschätzung des ärztlichen Prüfpersonals (z. B. aus Sicherheitsgründen oder Nichteinhaltung des Protokolls);</li> <li>• positiver Serum-Schwangerschaftstest oder</li> <li>• aufgrund spezifischer Aufforderung durch den Sponsor.</li> </ul> <p>Personen, welche im EKG ein wiederholt bestätigtes QTc-Intervall von <math>\geq</math> 500 ms oder eine Verlängerung von mindestens 60 ms im Vergleich zum Baselinewert aufwiesen, sollten die Studienmedikation sofort abbrechen. Den Studienunterlagen konnten keine differenzierten Angaben zu Behandlungs- und Studienabbruch entnommen werden.</p>	

Intervention	Kontrolle
<p><b>Regeln zur Medikation vor Studienbeginn sowie zur Begleitmedikation</b></p> <p>Medikationen entsprechend der Standardtherapie bei Patientinnen und Patienten mit Sichelzellerkrankung waren grundsätzlich erlaubt, unterlagen jedoch teilweise Einschränkungen.</p> <p>Bei einer gleichzeitigen Therapie mit HU und/oder Erythropoetin musste diese für mindestens 6 Monate vor der ersten Studienmedikationsgabe verschrieben worden sein sowie mit einer stabilen Dosis für mindestens 3 Monate vor der ersten Studienmedikationsgabe vorliegen. Etwaige Dosisänderungen und/oder ein Behandlungsabbruch von HU waren (außer aus Sicherheitsgründen) nicht vorgesehen und resultierten bei vorliegender Planung derselben für die Studienzeit im Ausschluss bei Rekrutierung.</p> <p>Die Teilnahme an einem chronischen Transfusionsprogramm, definiert als eine Reihe von geplanten (einfachen oder Austausch-) Transfusionen als Prophylaxe-Maßnahme, geplante Austauschtransfusion während des Studienzeitraums, eine chronische Therapie mit Antikoagulantien außer Aspirin und ein geplanter großer chirurgischer Eingriff während des Studienzeitraums führten ebenso zum Ausschluss bei Rekrutierung. Jegliche Medikation oder nicht pharmakologische Therapie von Patientinnen und Patienten sollte während der Studienteilnahme erfasst werden.</p>	

<sup>1)</sup> Dosierung und Verabreichung gemäß Angaben in der Fachinformation.

<sup>2)</sup> Regeln zu Verabreichungsmodifikation und -abbruch wurden ab Protokollversion 2 ins Protokoll aufgenommen.

Abkürzungen: HU: Hydroxyurea/Hydroxycarbamid; KG: Körpergewicht; UE: Unerwünschtes Ereignis.

## Protokolländerungen

Die Originalversion des Studienprotokolls vom 26. Februar 2013 wurde durch drei Amendments geändert bzw. ergänzt [11]. Die Protokollversionen 2 und 3 traten nach Studieneinschluss der ersten Patientin / des ersten Patienten (24. Juli 2013) und vor Abschluss der Rekrutierungsphase (Ende der Randomisierungsphase: Januar 2015) in Kraft [10]. Es konnte in den Studienunterlagen nicht identifiziert werden, wie viele Personen von den Protokolländerungen der Versionen 2 und 3 betroffen waren bzw. wie viele Personen zuvor bereits in die Studie eingeschlossen wurden. Protokoll 3 erlangte ca. 3–4 Monate vor Ende der Randomisierungsphase Gültigkeit. Die wesentlichen Änderungen der Protokollversionen sind nachfolgend in Tabelle 4 aufgeführt. Gemäß Protokoll 1 konnten etwa 8 Monate lang und gemäß Protokoll 3 für einen Zeitraum von 3–4 Monaten Personen im Altersbereich von 16 bis 18 Jahren rekrutiert werden. Weitere Angaben zur Anzahl der eingeschlossenen 16- bis 18-Jährigen liegen nicht vor. Protokollversion 4 (21. April 2016) erlangte erst nach Studienabschluss (der letzten Patientin / des letzten Patienten am 23. März 2016) Gültigkeit und bezieht sich somit ausschließlich auf den Analyseplan, welcher im später finalisierten SAP (2. Mai 2016) ausgeführt ist. Es konnten in den Studienunterlagen keine näheren Erklärungen zu den einzelnen Protokollversionen und keine Übersicht aller damit jeweils einhergehenden Protokolländerungen identifiziert werden. Hintergründe zu den vielfältigen Analyseplanänderungen im Rahmen von Protokoll 4, wie z. B. der Definition eines neuen Endpunkts und eines sekundären Schlüsselendpunkts, sind unklar. Diese Unklarheiten werden in Kapitel 2.4 näher erläutert.

Tabelle 4: Protokolländerungen

Protokollversion	Wesentliche Änderungen
Version 2 (07.04.2014)	<ul style="list-style-type: none"> <li>• Einschlusskriterium: Das Mindestalter wurde aufgrund von Rekrutierungsaufgaben der FDA bezüglich pädiatrischer Patientinnen und Patienten von 16 auf 18 Jahren angehoben.</li> <li>• Endpunktdefinition: Tod wurde als Definition eines VOC-Ereignisses entfernt. Während des Gültigkeitszeitraums von Protokoll 1 mussten Tode, die als Teil einer VOC auftraten, gleichzeitig auch als SUE erfasst werden.</li> </ul>
Version 3 (26.09.2014)	<ul style="list-style-type: none"> <li>• Einschlusskriterium: Nach Freigabe durch die FDA wurde das Mindestalter wieder auf 16 Jahre gesetzt.</li> <li>• Analyseplan: Ein stratifizierter Test soll Anwendung finden, um den primären Endpunkt zu analysieren und einen sekundären Schlüsselendpunkt zu identifizieren.</li> </ul>
Version 4 (21.04.2016 <sup>1)</sup> )	<ul style="list-style-type: none"> <li>• Analyseplan: <ul style="list-style-type: none"> <li>○ Die Analyse des primären Endpunkts wurde modifiziert, um die Stratifizierung bei Randomisierung zu berücksichtigen. Zudem wurde der Zeitraum der Auswertung von VOC, ursprünglich definiert vom Zeitpunkt der initialen Medikation bis zum letzten Kontakt, verändert. Der Zeitraum der Auswertung von VOC umfasste nun den Zeitpunkt der Randomisierung<sup>2)</sup> bis zur letzten Studienmedikationsgabe + 14 Tage bzw. bis zum Datum der Studienbeendigung (d. h. letzter Kontakt) für Personen, welche keine Studienmedikation erhalten haben.</li> <li>Die primäre Analyse der Wirksamkeitsendpunkte soll mit Ausnahme der patientenberichteten Endpunkte (BPI-LF und SF-36) laut Studienprotokoll analog zur Analyse des primären Endpunktes stattfinden.</li> <li>○ Anpassung der Sensitivitäts- und Subgruppenanalysen: Es wurden zwei neue Analysen definiert und alte Analysen entfernt (vgl. Tabelle 26 im Anhang).</li> <li>○ Endpunktkategorisierungen wurden umfangreich modifiziert, unter anderem: <ul style="list-style-type: none"> <li>▪ Der sekundäre Endpunkt „Jährliche Rate der Krankenhaustage“ wurde als sekundärer Schlüsselendpunkt definiert.</li> <li>▪ Der Endpunkt „Jährliche Rate der Schmerzkrisentage“ wurde erstmals definiert.</li> </ul> </li> </ul> </li> </ul>

<sup>1)</sup> Nach Angaben des pU gegenüber der EMA wurde die Datenbank am 13.05.2016 gesperrt [3].

<sup>2)</sup> Den Studienunterlagen war zu entnehmen, dass der Zeitpunkt der Randomisierung i. d. R. an Tag 1 der Behandlung stattgefunden hat. Unklar ist, ob es von dieser Regel Abweichungen gab, da den Unterlagen keine systematische Darstellung der Zeit zwischen Randomisierung und initialer Medikation zu entnehmen war.

Abkürzungen: BPI-LF: Brief Pain Inventory – Long Form; EMA: European Medicines Agency; FDA: U. S. Food and Drug Administration; pU: pharmazeutischer Unternehmer; SF-36: 36-item Short Form Health Survey; SUE: Schwerwiegendes unerwünschtes Ereignis; VOC: Vasookklusive Schmerzkrise.

## 2.3 Endpunkte

In diesem Kapitel wird die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier [10], der zugehörige Studienbericht [26], das Studienprotokoll [11,12,19,25], der SAP [18] und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. Tabelle 5 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 5: Zusammenfassung der Endpunktbewertung

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Mortalität <sup>1)</sup>	Mortalität	-	Ja
<i>Vasookklusive Schmerzkrisen</i> <ul style="list-style-type: none"> <li>• Jährliche Rate VOC<sup>2)</sup></li> <li>• Zeit bis zur ersten VOC</li> <li>• Zeit bis zur zweiten VOC</li> <li>• VOC-Subklassifikationen: <ul style="list-style-type: none"> <li>○ Jährliche Rate unkomplizierte VOC</li> <li>○ Jährliche Rate akute Thoraxsyndrome</li> <li>○ Lebersequestration</li> <li>○ Milzsequestration</li> <li>○ Priapismus</li> </ul> </li> <li>• Jährliche Rate der Schmerzkrisentage</li> <li>• Anteil der Personen ohne VOC während der Behandlungsphase</li> </ul>	Morbidität	Ja <sup>3)</sup>	Ja <sup>4)</sup>
Jährliche Rate der Krankenhaustage <ul style="list-style-type: none"> <li>• Jährliche Rate der Krankenhaustage aufgrund einer VOC</li> </ul>		Ja	Nein
Jährliche Rate der transfundierten Erythrozytenkonzentrate		Ja	Nein
BPI-LF		Ja	Ergänzend <sup>5)</sup>
SF-36	Lebensqualität	Ja	Ja
Unerwünschte Ereignisse	Sicherheit	Ja	Ja <sup>6)</sup>

<sup>1)</sup> Todesfälle wurden im Rahmen der Sicherheitserhebung erfasst.

<sup>2)</sup> Primärer Endpunkt der Studie.

<sup>3)</sup> Für die Subklassifikationen „Lebersequestration“, „Milzsequestration“ und „Priapismus“ wurden keine separaten statistischen Analysen in Form jährlicher Raten als eigenständige Endpunkte vorgenommen. Der Endpunkt „Jährliche Rate der Schmerzkrisentage“ wird durch den pU im Dossier ergänzend dargestellt.

<sup>4)</sup> Die Endpunkte „Jährliche Rate der Schmerzkrisentage“ und „Anteil der Personen ohne VOC während der Behandlungsphase“ werden in der Nutzenbewertung nicht berücksichtigt.

<sup>5)</sup> Ergebnisse des BPI-LF werden aufgrund geringer Rücklaufquoten nur zu Baseline, d. h. zu Tag 1 der Behandlungsphase als Teil der Studiencharakteristika dargestellt.

<sup>6)</sup> Patientenrelevanz von Laborparametern, Schwangerschaftstest u. a. unklar.

Abkürzungen: BPI-LF: Brief Pain Inventory – Long Form; pU: pharmazeutischer Unternehmer; SF-36: 36-item Short Form Health Survey; VOC: Vasookklusive Schmerzkrisen.

## 2.3.1 Mortalität

### Mortalität

Der Endpunkt „Mortalität“ wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Mortalität wurde in der Studie SUSTAIN nicht als separater Endpunkt erfasst. Todesfälle wurden im Rahmen der Sicherheitsbewertung und gemäß Protokoll 1 als Subklassifikation von VOC ausgewertet (siehe Kapitel 2.3.2 – Vasookklusive Schmerzkrisen und Kapitel 2.3.4). Todesfälle, die als VOC-Ereignis klassifiziert wurden, mussten auch als schwerwiegende unerwünschte Ereignisse erfasst werden.

##### *Erhebungszeitpunkte*

Todesfälle wurden u. a. im Rahmen der Sicherheitsbewertung anhand der unerwünschten Ereignisse mit Todesfolge (siehe Kapitel 2.3.4) erhoben.

##### *Auswertung*

Auf Basis der Sicherheitspopulation.

##### *Bewertung*

Die Operationalisierung ist weitestgehend nachvollziehbar.

#### Patientenrelevanz

Entsprechend § 2 Satz 3 Arzneimittel-Nutzenbewertungsverordnung (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“.

#### Validität

Mortalität war kein eigenständiger Endpunkt in der Studie SUSTAIN. Todesfälle waren jedoch Bestandteil der Erhebung von unerwünschten Ereignissen. Informationen zur Validität des Endpunkts „Unerwünschte Ereignisse“ können Kapitel 2.3.4 entnommen werden. Die externe Validität der Mortalitätserhebung wird durch die geringe Beobachtungsdauer eingeschränkt.

## 2.3.2 Morbidität

### Vasookklusive Schmerzkrisen

Die auf der Operationalisierung vasookklusiver Schmerzkrisen basierenden Endpunkte „Jährliche Rate VOC“, „Zeit bis zur ersten VOC“ und „Zeit bis zur zweiten VOC“ sowie die Subklassifikationen „Jährliche Rate unkomplizierte VOC“ und „Jährliche Rate akute Thoraxsyndrome“ sind unmittelbar patientenrelevant und werden in der Nutzenbewertung berücksichtigt. Die jeweilige Anzahl der VOC-Subklassifikationen „Lebersequestration“, „Milzsequestration“ und „Priapismus“ wurde ausschließlich deskriptiv aufgeführt, und sind vom pU nicht als eigenständige Endpunkte spezifiziert, werden aber dennoch als patientenrelevant eingestuft sowie in der Nutzenbewertung berücksichtigt. Alle Subklassifikationen sind zugleich im Endpunkt „Jährliche Rate VOC“ enthalten.

Der ebenfalls auf der VOC-Operationalisierung basierende Endpunkt „Jährliche Rate der Schmerzkrisentage“ und der post hoc dargestellte Endpunkt „Anteil der Personen ohne VOC während der Behandlungsphase“ werden als grundsätzlich patientenrelevant erachtet, jedoch aufgrund diesbezüglichen weitgehenden Schwächen bzw. Unklarheiten der Operationalisierung und damit einhergehenden Einschränkungen der Validität in der vorliegenden Nutzenbewertung nicht berücksichtigt.

### Operationalisierung

#### *Beschreibung*

VOC war definiert als

- a) eine akute Schmerzepisode,
- b) ohne andere Schmerzsache als ein vasookklusives Ereignis,
- c) bei der die Vorstellung in einer medizinischen Einrichtung und
- d) die orale bzw. parenterale Behandlung mit Opioiden oder die parenterale Behandlung mit nichtsteroidalen Antirheumatika (NSAID) notwendig war.

Alle vier Kriterien mussten zutreffen. Andere Symptome, die mit der Sichelzellerkrankung zusammenhängen, wie z. B. chronischer Schmerz aufgrund von Knöchel-/Beinulzationen, aseptische Knochennekrose oder Gicht, wurden nicht als VOC erfasst. Komplikationen (wie z. B. Lungen-, Nieren- oder Herzversagen), die vollständig unabhängig von einer VOC auftraten, wurden nicht als VOC erhoben. Sofern sich diese Komplikationen zu einer VOC entwickelten, wurde deren Klassifikation angepasst, d. h. diese zu „VOC“ geändert.

Jeglicher Besuch einer medizinischen Einrichtung aufgrund einer Schmerzepisode mit Verdacht auf VOC sollte erfasst werden und musste innerhalb von 5 Tagen nach Erstkenntnis des Studienzentrums durch das ärztliche Prüfpersonal im electronic Case Report Form (eCRF) einmal vermerkt werden. Dabei wurden folgende Kriterien erhoben:

- Diagnose, limitiert auf eine der prädefinierten Subklassifizierungen (siehe Tabelle 6);
- Datumsangaben zu Beginn und Ende der VOC;
- ergriffene Maßnahme (keine, benötigte Begleitmedikation, temporäre Unterbrechung der Studienmedikation, Abbruch der Studienmedikation oder andere);
- ob die Notwendigkeit einer Hospitalisierung gegeben war oder nicht;
- welche Begleitmedikation/en gegeben wurde/n;
- Resultat (Erholung ohne Folgen, Erholung mit Folgen, anhaltend, unbekannt oder Tod).

Diese Ereignisse sollten nicht als UE oder SUE berichtet werden.

Für Patientinnen und Patienten, die aufgrund einer Schmerzepisode mit Verdacht auf VOC in anderen medizinischen Einrichtungen als in ihrem jeweiligen Studienzentrum behandelt wurden, sollten jeweils Zusammenfassungen des Besuchs zusammengetragen werden. Dazu wurden den Patientinnen und Patienten Studienteilnahmekarten ausgegeben, welche die nötigen Informationen anfordern und bei jeder Vorstellung in einer medizinischen Einrichtung vorgelegt werden können. Vasookklusive Schmerzkrisen an (gemäß Protokoll) geplanten Visitentagen im Studienzentrum wurden nur als VOC-Ereignisse gezählt, wenn die Person angab, dass sie aufgrund des krisenverbundenen Schmerzes auch eine medizinische Einrichtung aufgesucht hätte, wenn keine Visite geplant gewesen wäre.

Alle vom jeweiligen ärztlichen Prüfpersonal als VOC eingestuft Ereignisse mussten zusätzlich durch ein vom Sponsor bestimmtes und verblindetes Crisis Review Committee (CRC), bestehend aus drei Ärztinnen/Ärzten, die in der Behandlung der Sichelzellerkrankung erfahren sind, adjudiziert

werden. Die für das CRC verfügbaren Unterlagen zur Adjudikation konnten u. a. klinische Aufzeichnungen (der Krankengeschichte und von körperlichen Untersuchungen), Berichte von Notfallambulanzen (zur Krankengeschichte und von körperlichen Untersuchung), Krankenhausentlassungsberichte, klinische Laborwerte, körperliche Untersuchungsergebnisse, Begleitmedikationen und Röntgenbilder der Brust (wenn ACS berichtet wurde) umfassen. Sollte das CRC zu dem Ergebnis gekommen sein, dass nicht genügend Evidenz zur Validierung eines Ereignisses vorlag, oder das CRC anderer Meinung war als die jeweilige ärztliche Prüfperson, so konnte letztere weitere relevante Informationen nachreichen, wobei die abschließende Diagnoseentscheidung jedoch beim CRC verblieb. Weitere Regeln bspw. zu Adjudikationsgründen wurden im Protokoll nicht präspezifiziert.

Unkomplizierte VOC sowie vasookklusiv-bedingte Notfallkomplikationen, d. h. akutes Thoraxsyndrom (ACS), Priapismus, Lebersequestration und Milzsequestration mit Notwendigkeit zur Vorstellung in einer medizinischen Einrichtung, wurden jeweils als eine Subklasse eines VOC-Ereignisses erfasst. Eine genaue Definition dieser Ereignisse kann Tabelle 6 entnommen werden.

*Tabelle 6: Subklassifizierungen von vasookklusiven Schmerzkrisenereignissen*

<b>Vasookklusives Schmerzkrisenereignis<sup>1)</sup></b>	<b>Definition</b>
Unkomplizierte VOC	Eine a) akute Schmerzepisode, welche die Kriterien b), c) und d) der Definition einer VOC erfüllt sowie nicht als ACS, Lebersequestration, eine Milzsequestration oder Priapismus klassifiziert ist.
ACS	Wurde basierend auf einem neuen pulmonalen Infiltrat gestellt, das mindestens ein vollständiges Lungensegment involvierte und im Thorax-Röntgenbild mit einer alveolären Konsolidierung (nicht aber Atelektase) konsistent war. Außerdem musste mindestens eines der folgenden Symptome vorliegen: <ul style="list-style-type: none"> <li>• Von Patientin/Patient berichteter Brustschmerz,</li> <li>• Körpertemperatur &gt; 38,5 °C,</li> <li>• Tachypnoe,</li> <li>• pfeifende Atmung oder</li> <li>• Husten.</li> </ul>
Lebersequestration	Diagnose basierend auf Schmerzen im oberen rechten Quadranten, einer vergrößerten Leber und einem akuten Abfall der Hämoglobin-Konzentration (z. B. um ~ 2 g/dl).
Milzsequestration	Diagnose basierend auf Schmerzen im linken oberen Quadranten, einer vergrößerten Milz und einem akuten Abfall der Hämoglobin-Konzentration (z. B. um ~ 2 g/dl).
Priapismus	Diagnose basierend auf einer anhaltenden penilen Erektion, welche der Vorstellung in einer medizinischen Einrichtung bedarf.

<sup>1)</sup> Zu Protokollversion 1 wurde „Tod aus jeglichem Grund“ – außer durch einen eindeutigen Fall eines Tötungsdelikts, Suizids oder Unfalls – ebenfalls als VOC-Ereignis gezählt und ab Protokoll 2 (nach erstem Studieneinschluss in Kraft getreten) aus der Definition von VOC entfernt.

Abkürzungen: ACS: Akutes Thoraxsyndrom; VOC Vasookklusive Schmerzkrise.

### *Erhebungszeitraum*

Die Erfassung von VOC sollte gemäß allen Protokollversionen und Case Report Forms (CRF) über die gesamte Behandlungsphase bis zur Follow-up-Visite erfolgen. Ein standardisiertes Vorgehen zu den Zeitpunkten der Benachrichtigung des jeweiligen Studienzentrums durch die Patientin / den Patienten über eine VOC in einer medizinischen Einrichtung außerhalb des Studienzentrums oder das standardisierte Auslesen der zur Dokumentation ausgegebenen Studienteilnahmekarten

konnte in den Studienunterlagen nicht identifiziert werden. Es wird davon ausgegangen, dass die jeweilige Kenntnisnahme spätestens während den Studienvisiten erfolgen sollten. Diese Studienvisiten umfassten:

- Behandlungsphase: Tag 1, Tag 15, dann im 4-Wochen-Rhythmus bis Woche 50 mit einer Abschlussuntersuchung in Woche 52
- Follow-up-Phase: Woche 58

Nach etwaigem Behandlungsabbruch sollte die Behandlungsabschlussvisite so zeitnah wie möglich durchgeführt werden. Zudem sollte die Follow-up-Visite nach Möglichkeit ungefähr 8 Wochen nach der letzten Dosisgabe erfolgen.

Es ist zu berücksichtigen, dass die retrospektive Erhebung von VOC, die während der Screeningphase (bis zu 30 Tage vor Behandlungsbeginn) aufgetreten sind, an Tag 1 der Behandlungsphase als Teil des Baselinecharakteristikums VOC während der vergangenen 12 Monate erfolgte und sich von der behandlungsphasenspezifischen Dokumentation der VOC insofern unterschied, dass das vergangene Vorliegen einer VOC entweder auf Basis medizinischer Unterlagen oder der Erinnerung der Betroffenen erfolgen konnte.

#### *Auswertung*

Die vorgelegten Auswertungen enthalten jährliche Raten, deskriptive Auswertungen und Time-to-Event-Analysen.

Die jährliche Rate wurde am Beispiel der primären Auswertung des Endpunkts „Jährliche Rate VOC“ (= JR VOC) folgendermaßen definiert:

$$JR\ VOC = \frac{\text{Gesamtzahl VOC einer Person während der Behandlungsphase} \times 365}{(\text{Behandlungsende} - \text{Datum der Randomisierung} + 1)}$$

Das Behandlungsende war definiert als der Zeitpunkt der letzten Dosis + 14 Tage (entspricht bei regulärer Studienteilnahme der Abschlussuntersuchung der Behandlungsphase in Woche 52). Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt). Änderungen des Auswertungszeitraums können Kapitel 2.2 (Abschnitt „Protokolländerungen“) entnommen werden. Die primäre Auswertung erfolgte auf Basis der ITT-Population.

Details zu Time-to-Event-Analysen sind Kapitel 2.4 zu entnehmen.

Folgende auf der Operationalisierung von VOC basierende Auswertungen wurden vorgelegt:

1. Jährliche Rate VOC  
Gemäß Studienunterlagen als primärer Endpunkt definiert; die primäre Auswertung erfolgte wie oben dargestellt.
2. Zeit bis zur ersten VOC  
Definiert als Anzahl der Monate von der Randomisierung bis zum Datum der ersten VOC. Der Beginn des Auswertungszeitraums wurde nach Studienende vom ersten Behandlungstag zum Zeitpunkt der Randomisierung geändert (siehe Kapitel 2.2 – Protokolländerungen). Gemäß Studienunterlagen sollten diese i. d. R. übereinstimmen. Eine Person ohne Schmerzkrise vor Studienabbruch oder Behandlungsende (d. h. Zeitpunkt der letzten Dosis + 14 Tage) wurde zum jeweiligen Zeitpunkt des Studienabbruchs bzw. Behandlungsendes zensiert (siehe „Anteil der Personen ohne VOC während der Behandlungsphase“). Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt). Die primäre Auswertung erfolgt auf Basis der ITT-Population.

### 3. Zeit bis zur zweiten VOC

Definiert als Anzahl der Monate von der Randomisierung bis zum Datum der zweiten VOC. Der Beginn des Auswertungszeitraums wurde nach Studienende vom ersten Behandlungstag zum Zeitpunkt der Randomisierung geändert (siehe Kapitel 2.2 – Protokolländerungen). Gemäß Studienunterlagen sollten diese i. d. R. übereinstimmen. Eine Person mit weniger als zwei Schmerzkrisen vor Studienabbruch oder Behandlungsende (d. h. Zeitpunkt der letzten Dosis + 14 Tage) wurde zum jeweiligen Zeitpunkt des Studienabbruchs bzw. Behandlungsendes zensiert. Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt). Die primäre Auswertung erfolgt auf Basis der ITT-Population.

### 4. VOC-Subklassifikationen

Jegliche VOC wurde als eine der definierten Subklassifikationen „Jährliche Rate unkomplizierte VOC“, „Jährliche Rate akute Thoraxsyndrome“, „Lebersequestration“, „Milzsequestration“ oder „Priapismus“ klassifiziert. Bis zur Gültigkeit von Protokoll 2 galt Tod auch als eine Subklassifikation (siehe Kapitel 2.2 – Protokolländerungen). Die primäre Auswertung der Subklassifikationen „Jährliche Rate unkomplizierte VOC“ und „Jährliche Rate akute Thoraxsyndrome“ erfolgte analog zu der des Endpunkts „Jährliche Rate VOC“. Zu den Subklassifikationen „Lebersequestration“, „Milzsequestration“ und „Priapismus“ wurden keine separaten statistischen Analysen definiert, jedoch wurden im Rahmen der Analyse zum Endpunkt „Jährliche Rate VOC“ post hoc deskriptive Auswertungen basierend auf der ITT-Population vorgelegt.

### 5. Jährliche Rate der Schmerzkrisentage

Der Endpunkt „Jährliche Rate der Schmerzkrisentage“ wurde nach Studienende definiert (siehe Kapitel 2.2 – Protokolländerungen). Nach Angaben des pU in Modul 4 sei die für diesen Endpunkt relevante Zahl der Krisentage den Angaben des ärztlichen Prüfpersonals zu Krisenbeginn und Krisenende im CRF entnommen worden. Die primäre Auswertung der Schmerzkrisentage erfolgte analog zu der des Endpunkts „Jährliche Rate VOC“, „Jährliche Rate VOC“.

### 6. Anteil der Personen ohne VOC während der Behandlungsphase

Der vorliegende Endpunkt wurde in Modul 4 post hoc berichtet und definiert als Personen, welche die Studie beendet haben und deren jährliche Rate VOC in Verbindung mit Aufsuchen einer medizinischen Einrichtung gleich 0 war.

#### *Bewertung*

Es ergeben sich Einschränkungen bezüglich der Nachvollziehbarkeit der Operationalisierung, die im Folgenden in Reihenfolge der Beschreibung der Operationalisierung (nicht nach Wertigkeit) beschrieben werden.

Zur Definition der VOC bestehen Unklarheiten bezüglich der Anwendung der Kriterien a), b) sowie vor allem der Bedingungen c), der Vorstellung in einer medizinischen Einrichtung, und d), der oralen bzw. parenteralen Behandlung mit Opioiden, auf die Subklassifizierungen der VOC. Unter der Bedingung, dass alle Kriterien a), b), c) und d) für jegliche VOC gegeben sein mussten, um dieses innerhalb der Studie als VOC zu erheben, ist nicht eindeutig ersichtlich, ob und wie Ereignisse erfasst wurden, die einer Subklassifizierung (z. B. ACS) weitestgehend entsprachen und möglicherweise nicht mindestens eines der generellen VOC-Kriterien (bspw. Behandlung mit Opioiden oder NSAID) erfüllten.

Während die Kriterien c) und d) sowie weitestgehend das Kriterium b) eindeutig operationalisiert erscheinen, bleiben Fragen hinsichtlich der Diagnose und insbesondere der Abgrenzung von akuten VOC gegenüber chronischen Schmerzen in Verbindung mit einer Sichelzellerkrankung oder

zu anderen akuten VOC (d. h. der Zählung von VOC) offen (siehe Validität). So liegen bspw. keine Angaben zu einer direkten Erfassung des Schmerzes und des jeweiligen Ausmaßes im Rahmen der VOC vor, noch werden andere potentiell mit der Sichelzellerkrankheit in Verbindung stehende schmerzauslösende Ereignisse, wie bspw. cerebro-vaskuläre Ereignisse, diskutiert [4]. Auch wird eine fehlende Standardisierung bei der Bestimmung von Krisenbeginn und Krisenende beschrieben, da diese weder in den Protokollen noch im CRF eindeutig definiert sind. Wie der pU im Dossier ausführt, war dieser Aspekt nicht Gegenstand des Reviews durch das CRC, sodass die Dauer einer Schmerzkrisse zwischen einzelnen Studienzentren bzw. Prüferinnen/Prüfern relevant variieren konnte [10]. Diese fehlende Standardisierung betrifft alle VOC-bezüglichen Endpunkte und insbesondere den Endpunkt „Jährliche Rate Schmerzkrisentage“ sowie die Endpunkte „Zeit bis zur ersten VOC“ bzw. „Zeit bis zur zweiten VOC“.

Die in Kapitel 2.2 (Abschnitt „Protokolländerungen“) beschriebenen Änderungen in Protokoll 4 bezüglich des Beginns des Auswertungszeitraums von VOC sind nicht näher erläutert, lassen sich jedoch grundsätzlich hinsichtlich einer Auswertung nach ITT-Prinzip nachvollziehen. Unklar erscheint, warum (vor als auch nach dieser Protokolländerung) VOC nicht unabhängig von der Studienmedikationsgabe bzw. dem etwaigen Behandlungsabbruch bis zum Zeitpunkt der präspezifizierten Wirksamkeitsanalyse zu Woche 52 erfasst werden sollten.

Der post hoc definierte Endpunkt „Anteil der Personen ohne VOC während der Behandlungsphase“ wird aus zwei Gründen nicht dargestellt: Einerseits liegen diesbezüglich zu große Unklarheiten hinsichtlich der Operationalisierung vor; andererseits geht der Anteil der Personen ohne VOC in die Time-to-Event-Analyse des a priori definierten Endpunkts „Zeit bis zur ersten VOC“ mit ein und wird daher bereits berücksichtigt.

#### 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“.

Trotz der genannten Einschränkungen und Unklarheiten bezüglich der Operationalisierung der Endpunkte werden „Jährliche Rate VOC“, „Jährliche Rate unkomplizierte VOC“, „Jährliche Rate akute Thoraxsyndrome“, die VOC-Subklassifikationen „Lebersequestration“, „Milzsequestration“ und „Priapismus“ sowie die „Zeit bis zur ersten VOC“ und „Zeit bis zur zweiten VOC“ als patientenrelevant erachtet. Der Endpunkt „Jährliche Rate der Schmerzkrisentage“ und der Endpunkt „Anteil der Personen ohne VOC während der Behandlungsphase“ werden grundsätzlich als patientenrelevant erachtet.

#### Validität

Hinsichtlich der Validität jeglicher VOC ist einschränkend zu berücksichtigen, dass die Voraussetzungen einer Vorstellung in einer medizinischen Einrichtung (Kriterium c)) sowie der oralen bzw. parenteralen Behandlung mit Opioiden oder einer parenteralen Behandlung mit NSAID (Kriterium d)) zum Erfassen einer VOC durch den jeweiligen Versorgungskontext beeinflusst worden sein könnte, da alle teilnehmenden Studienzentren (in den USA, Brasilien und Jamaika) außerhalb Deutschlands bzw. Europas liegen. Unabhängig vom Versorgungskontext bleibt vorwiegend unklar, inwiefern eine Eingrenzung auf VOC, die das Aufsuchen einer medizinischen Einrichtung und eine spezifische Behandlung bedingen, die quantitativen und qualitativen Ausprägungen vasookklusiver Schmerzkrisen der Patientinnen und Patienten hinreichend abbildet. So wurden

Behandlungen von vasookklusiven Ereignissen außerhalb medizinischer Einrichtungen, z. B. Zuhause mit oder ohne Selbstmedikation, durch die vorliegende Operationalisierung nicht erfasst. Das Aufsuchen einer medizinischen Einrichtung aufgrund einer VOC könnte somit – abgesehen vom Schweregrad des Ereignisses – jeweils in Abhängigkeit vom Krankenversicherungsstatus (bzw. der ökonomischen Situation), der Erreichbarkeit und Entfernung der nächsten medizinischen Einrichtung, dem gleichzeitigen Besitz von adäquaten Medikamenten sowie weiteren Faktoren variieren.

Alle von der jeweiligen ärztlichen Prüfperson als VOC eingestuften Ereignisse mussten zusätzlich durch ein verblindetes CRC adjudiziert werden. Im Rahmen des EPAR (European Public Assessment Report) wurden durch „Good Clinical Practice“ (GCP)-Inspektionen kritische Aspekte bezüglich des klinischen Adjudikationsprozesses und dessen Dokumentation aufgedeckt [3]. So besteht hauptsächlich eine Unsicherheit bezüglich einer 2-Wochenregel (VOC, die innerhalb von 14 Tagen auftraten, wurden nicht als unterschiedliche Ereignisse gezählt), da diese Zeitspannenregel nicht in den Studienprotokollen präspezifiziert war und nicht nachvollzogen werden kann, ob diese innerhalb der Studie konsistent Anwendung fand. Die GCP-Untersucher der EMA empfahlen daher die Daten des verblindeten CRC aufgrund der verbleibenden Unsicherheiten nicht zu akzeptieren. Da eine Stichprobe von zwei (der insgesamt sechzig) durch die Untersucher inspizierten Studienzentren nach deren Einschätzung relativ gute GCP-Compliance aufwies und die dort gesammelten Daten als glaubwürdig eingeschätzt wurden, schloss die EMA, dass stattdessen das Heranziehen von VOC-Einschätzungen gemäß ärztlichem Prüfpersonal zur Nutzen-Risiko-Bewertung akzeptabel sei [3]. Aufgrund der ausgeführten Unklarheiten bezüglich der Operationalisierung von VOC sowie der Limitation der GCP-Einschätzung der seitens der EMA eingesetzten Untersucher aufgrund einer Stichprobe von nur zwei Studienzentren, wird gefolgert, dass sowohl die Validität der VOC gemäß Einschätzung des ärztlichen Prüfpersonals als auch die Validität der CRC-adjudizierten VOC Einschränkungen unterliegen.

Im Kontext der Bewertung der Auswertungen jährlicher Raten, die eine Annualisierung von Zählwerten bedeuten, ist zu berücksichtigen, dass bei etwaigem frühzeitigem Behandlungsabbruch oder Lost to Follow-up die bereits vorliegenden Daten auf ein Jahr extrapoliert werden. Die Eignung dieses Vorgehens mit fehlenden Werten setzt die Annahme voraus, dass die relevanten annualisierten Ereignisse (bspw. VOC, ACS u. a.) nach Behandlungsabbruch eine gleiche oder hinreichend ähnliche Auftretensrate aufweisen wie vor Behandlungsabbruch. Da in der Studie SUSTAIN u. a. angestrebt war eine Wirksamkeit von Crizanlizumab hinsichtlich der Prävention wiederkehrender VOC gegenüber Placebo nachzuweisen, kann der Annahme einer gleichbleibenden VOC-Rate nach Studienabbruch nicht gefolgt werden. Daraus ergeben sich Implikationen für den Umgang mit fehlenden Daten im Zusammenhang mit den statistischen Auswertungen, die in Kapitel 2.4 näher erläutert werden. Es ist nicht ersichtlich, warum keine separaten statistischen Auswertungen hinsichtlich der VOC-Subklassifikationen „Lebersequestration“, „Milzsequestration“ oder „Priapismus“ präspezifiziert wurden.

Bezüglich des Endpunkts „Jährliche Rate der Schmerzkrisentage“ ist festzuhalten, dass der Einschätzung des pU hinsichtlich deutlicher Einschränkungen der Validität, welche dieser auf die fehlende Standardisierung bei der Bestimmung von Krisenbeginn und Krisenende (eine diesbezügliche Operationalisierung fehlt in den Protokollen gänzlich) und der Beeinflussung der Krisendauer durch Analgetika begründet, gefolgt wird. Angaben zu Erfassung und Graduierung von Schmerz im Rahmen der VOC-Erhebung fehlen und resultieren in einem hohen Variabilitätspotential. Des Weiteren sind die allgemein beschriebenen Einschränkungen basierend auf der

Definition von VOC und deren Validität zu berücksichtigen. Aufgrund der umfassenden Unklarheiten sowie der nicht gegebenen Validität des Endpunkts „Jährliche Rate der Schmerzkrisentage“, wird dieser in der Nutzenbewertung nicht berücksichtigt.

### **BPI-LF**

Der Endpunkt „BPI-LF“ wird in der Nutzenbewertung aufgrund der eingeschränkten Validität ergänzend dargestellt.

#### Operationalisierung

##### *Beschreibung*

Der Brief Pain Inventory (BPI) ist ein Fragebogen zur Selbsteinschätzung der wahrgenommenen Schmerzen. Der Fragebogen enthält Items zu Schmerzintensität (sensorische Dimension), Beeinträchtigung der täglichen Aktivitäten durch Schmerzen (reaktive Dimension) sowie Schmerzentlastung, Schmerzqualität und Einschätzung der Schmerzursache. Der Fragebogen liegt in einer Kurzversion (Short Form; neun Items) und einer Langversion (Long Form (LF); 17 Items und 6 demographische Items) vor und wurde in diverse Sprachen übersetzt. Einer der wichtigsten Unterschiede zwischen der Kurz- und der Langversion ist, dass sich die Kurzversion auf die wahrgenommenen Schmerzen der vergangenen 24 Stunden und die Langversion auf die wahrgenommenen Schmerzen der vergangenen Woche bezieht [1].

Aus dem BPI-LF können zwei Scores ermittelt werden: Der Score zur Schmerzintensität („pain severity score“) und zur Beeinträchtigung durch Schmerz („pain interference score“).

- Der Score zur Schmerzintensität setzt sich zusammen aus vier Items zur Schmerzintensität (Beschreibung der stärksten, geringsten, durchschnittlichen Schmerzen in den letzten sieben Tagen sowie Beschreibung der aktuellen Schmerzen), die jeweils auf einer elfstufigen Skala von 0 (kein Schmerz) bis 10 (stärkste vorstellbare Schmerzen) bewertet werden. Für den Score werden die Werte aufsummiert (Bereich: 0–40).
- Der Score zur Beeinträchtigung durch Schmerz setzt sich zusammen aus sieben Items zur Beeinträchtigung in den letzten sieben Tage hinsichtlich folgender Bereiche: Allgemeine Aktivität, Stimmung, Gehvermögen, normale Arbeit (sowohl außerhalb des Hauses als auch Hausarbeit), Beziehungen zu Menschen, Schlaf und Lebensfreude. Die Bewertung wird ebenso auf einer elfstufigen Skala vorgenommen von 0 (keine Beeinträchtigung) bis 10 (vollständige Beeinträchtigung). Für den Score werden die Werte aufsummiert (Bereich: 0–70).

Die übrigen Items des Fragebogens gehen nicht in einen Score ein. Da alle Score-relevanten Items in der Kurz- und Langversion enthalten sind, bezieht sich der Unterschied zwischen den Versionen nur auf den Zeitraum der wahrgenommenen Schmerzen.

In der Studie SUSTAIN wurde der BPI-LF verwendet. Gemäß Protokoll und Studienbericht erfolgte die Auswertung anhand des Manuals [1]. Im Studienbericht werden dementsprechend Ergebnisse für die beiden Skalen „Schmerzintensität“ und „Beeinträchtigung durch Schmerz“ berichtet.

Im Dossier hat der pU darüber hinaus eine post hoc definierte Auswertung zu dem Item „stärkster Schmerz“ („Kreisen Sie die Zahl ein, die Ihre stärksten Schmerzen in der letzten Woche beschreibt“) vorgelegt. Dieses Item ist bereits in der Skala zur Schmerzintensität enthalten (eines der vier Items, die in den Gesamtscore eingehen). Der pU begründet diese Auswertung mit Bezug auf zwei Nutzenbewertungen des IQWiG, in denen die Patientenrelevanz dieses Einzelitems hervorgehoben wurde [6,7].

### *Erhebungszeitpunkte*

- Behandlungsphase: Tag 1, Tag 15, zu Woche 14, 26, 38 mit einer Abschlussuntersuchung in Woche 52
- Follow-up-Phase: Woche 58

Nach etwaigem Behandlungsabbruch sollte die Behandlungsabschlussvisite so zeitnah wie möglich durchgeführt werden. Zudem sollte die Follow-up-Visite nach Möglichkeit ungefähr 8 Wochen nach der letzten Dosisgabe erfolgen.

### *Auswertung*

Die primäre Auswertung erfolgte auf Basis der ITT-Population bis Woche 58. Die Analyse basierte auf Mittelwertdifferenzen, die mittels eines linearen gemischten Modells mit wiederholten Messungen (MMRM) geschätzt wurden.

Weiterhin weist der pU im Dossier darauf hin, dass die Erhebung des Endpunkts zu a priori festgelegten Zeitpunkten und nicht im Zusammenhang mit dem Auftreten einer Schmerzkrise erfolgte [10]. Dementsprechend fiel nur ein Bruchteil der Erhebungen per Zufall mit dem Auftreten einer VOC zusammen.

### *Bewertung*

Die Operationalisierung des Endpunkts ist weitgehend nachvollziehbar.

Gemäß Manual [1] wird in klinischen Studien empfohlen den BPI-LF zur Baselinemessung anzuwenden und im Verlauf nur noch die Kurzversion (BPI-SF) zu nutzen. In der vorliegenden Studie wurde der BPI-LF sowohl zur Baselinemessung als auch zur Verlaufsmessung verwendet. Dies ist vor dem Hintergrund des unterschiedlichen Bezugszeitraums der wahrgenommenen Schmerzen der beiden Versionen nachvollziehbar.

Das Krankheitsbild der Sichelzellerkrankung ist charakterisiert durch Schmerzspitzen während der VOC. Im Verlauf der Erkrankung können sich auch chronische Schmerzsyndrome entwickeln. Gemäß der Anwendung des BPI-LF in der Studie SUSTAIN wurden a priori Zeitpunkte zur Erhebung festgelegt. Dadurch werden nicht systematisch die Schmerzen durch VOC erfasst, sondern Schmerzen anhand von Stichproben von sieben Erhebungszeitpunkten mit jeweils einer Woche Bezugszeitraum (inklusive Baselineerhebung), d. h. insgesamt 7 Wochen eines 59 Wochen umfassenden Zeitraums, zufällig erhoben. Die Werte des BPI-LF reflektieren also eher den Verlauf von chronischen Schmerzen während der Studie. Da Crizanlizumab zur Prävention wiederkehrender vasookklusiver Krisen zugelassen und nicht zu Behandlung chronischer Schmerzen vorgesehen ist, wird die Aussagekraft der Ergebnisse durch die Operationalisierung im vorliegenden Anwendungsgebiet eingeschränkt [9].

### 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“.

Die Erfassung von Schmerzen durch den BPI-LF wird als patientenrelevant bewertet.

### Validität

Der BPI wurde zunächst als Wisconsin Brief Pain Questionnaire (BPQ) bekannt [2] und ursprünglich zur Erhebung krebsbedingter Schmerzen entwickelt. Der BPI ist mittlerweile in einer Vielzahl klinischer Indikationen untersucht und validiert. Es existieren zahlreiche kultur- und sprachvalidierte Übersetzungen [1,22].

Die Untersuchungen zu den psychometrischen Eigenschaften des BPI beschränken sich vor allem auf die beiden Skalen „Schmerzintensität“ und „Beeinträchtigung durch Schmerz“, die sowohl in der Kurz- also auch in der Langversion enthalten sind. Die Test-Retest Reliabilität wird als gut [22] bis hoch [1] bewertet und die interne Konsistenz ist für beide Skalen hoch [1,22].

Hinsichtlich der Struktur ergibt sich eine Zwei-Faktoren-Struktur („Schmerzintensität“ und „Beeinträchtigung durch Schmerz“), wie von der Autorengruppe intendiert. Allerdings zeigte eine Studie auch eine Drei-Faktoren-Struktur, wobei die Skala „Beeinträchtigung durch Schmerz“ noch in die Faktoren Aktivität und Affekt getrennt werden kann [1,22]. Hinsichtlich der konvergenten Validität konnte gezeigt werden, dass die beiden Skalen moderat ( $r = 0,47-0,65$ ) mit dem körperlichen Summenscore (PCS) des SF-36 korrelierten sowie ebenso moderat ( $r = 0,57$ ) mit dem Roland and Morris Disability Questionnaire [22].

Auch zur Änderungssensitivität liegen Studien für verschiedene Krankheitsbilder vor, die zum Teil ein starkes Ansprechen auf eine Therapie zeigen konnten. Für das vorliegende Krankheitsbild der Sichelzellerkrankung liegen keine Informationen zu psychometrischen Eigenschaften, insbesondere zur Änderungssensitivität, vor.

Der pU stellt im Dossier neben den Auswertungen der a priori definierten Skalen „Schmerzintensität“ und „Beeinträchtigung durch Schmerz“ (vier Items) eine post hoc definierte Auswertung zu dem Item „stärkster Schmerz“ (Einzelitem) dar. Der Endpunkt „stärkster Schmerz“ (Einzelitem) ist jedoch schon in der Skala „Beeinträchtigung durch Schmerz“ berücksichtigt. Die Ergebnisse des Einzelitems „stärkster Schmerz“ werden daher in der Darstellung in Modul 4 doppelt berücksichtigt. Zudem war den Unterlagen des pU keine Begründung zu entnehmen, warum das Einzelitem „stärkster Schmerz“ im Anwendungsgebiet Sichelzellerkrankung die Schmerzen besser abbilden könnte als die Gesamtskala „Schmerzintensität“.

Insgesamt bleibt unklar, inwiefern mit der Erhebung des BPI-LF im vorliegenden Krankheitsbild der Sichelzellerkrankung die Häufigkeit und Intensität von VOC abgebildet werden sollte. Die Schmerzhäufigkeit wird bereits im Endpunkt „Vasookklusive Schmerzkrisen“ berücksichtigt. Da der Fragebogen nur zu a priori festgelegten 7 von 59 Studienwochen akute Schmerzen erfasst, wird auch die Intensität von VOC nicht ausreichend systematisch erfasst. Der BPI-LF bildet somit eher den chronisch wahrgenommenen Schmerz ab, das vorliegende Anwendungsgebiet bezieht sich jedoch auf akute VOC. Da die Rücklaufquoten bereits zum ersten Messzeitpunkt nach Studienbeginn (Tag 15) und auch im Verlauf unter 70 % im Interventionsarm bleiben (siehe auch Kapitel 3.3), ist die Aussagekraft der Ergebnisse im Studienverlauf nicht belastbar. Bezüglich des Endpunkts „BPI-LF“ werden daher nur die Ergebnisse der a priori definierten Skalen „Schmerzintensität“ und „Beeinträchtigung durch Schmerz“ zu Baseline im Rahmen der Studiencharakteristika ergänzend dargestellt.

### Jährliche Rate der transfundierten Erythrozytenkonzentrate

Der Endpunkt „Jährliche Rate der transfundierten Erythrozytenkonzentrate“ wird in der Nutzenbewertung aufgrund fehlender Patientenrelevanz nicht berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Wie der pU in Modul 4 ausführt, können Erythrozytenkonzentrate, in Form von Einfach- oder Austauschtransfusionen verabreicht, in Notfällen zur Behandlung von Akutkomplikationen der Sichelzellerkrankung, wie z. B. einer akuten Milz- oder Lebersequestration, einem akuten Thoraxsyndrom, einem Schlaganfall und einem Multiorganversagen erforderlich werden [4,10]. Transfusion von Erythrozytenkonzentraten können auch gelegentlich elektiv verabreicht werden oder Teil eines langfristig ausgelegten Therapiekonzepts zur Prophylaxe vorhersehbarer Krankheitskomplikationen sein [4]. Personen, die an einem chronisch elektiven Transfusionsprogramm teilnahmen, planten an diesem während der Studie teilzunehmen oder planten sich einer elektiven Austauschtransfusion während der Studie zu unterziehen, sollten gemäß Ausschlusskriterien nicht in die SUSTAIN-Studie eingeschlossen werden (siehe Tabelle 2) Ein chronisches Transfusionsprogramm war definiert als eine vorher geplante Serie an einfachen oder Austauschtransfusionen mit prophylaktischer Zielsetzung.

Hinsichtlich der statistischen Auswertung der transfundierten Erythrozytenkonzentrate wurde in den Studienunterlagen ein analoges Vorgehen zum primären Endpunkt beschrieben. Es konnten keine expliziten Angaben zu Erhebung, Erhebungszeitpunkten und ähnlichen Aspekten bezüglich des Endpunkts „Jährliche Rate der transfundierten Erythrozytenkonzentrate“ identifiziert werden. Jedoch fand im Rahmen der Erfassung von VOC eine Dokumentation der diesbezüglich in einer medizinischen Einrichtung gegebenen Begleitmedikation statt (siehe Abschnitt „Vasookklusive Schmerzkrisen“). Gleichzeitig liegen Angaben dazu vor, dass jegliche Medikation oder non-pharmakologische Therapie, welche durch eine Untersuchungsperson während der Studienteilnahme eingenommen bzw. mit welcher diese behandelt wurde, vom ärztlichen Prüfpersonal im eCRF (samt Dosis, Behandlungsschema, Vergabeart, Indikation und Anwendungsdatum) vermerkt werden musste. Der Vermerk von Begleitmedikationen und -therapien fand zu folgenden Zeitpunkten statt:

##### *Erhebungszeitpunkte (gemäß Erhebung der Begleitmedikation)*

- Screeningphase: bis zu 30 Tage vor Studieneinschluss
- Behandlungsphase: Tag 1, Tag 15, dann im 4-Wochen-Rhythmus bis Woche 50 mit einer Abschlussuntersuchung in Woche 52
- Follow-up-Phase: Woche 58

Nach etwaigem Behandlungsabbruch sollte die Behandlungsabschlussvisite so zeitnah wie möglich durchgeführt werden. Zudem sollte die Follow-up-Visite nach Möglichkeit ungefähr 8 Wochen nach der letzten Dosisgabe erfolgen.

##### *Auswertung*

Die primäre Auswertung des Endpunkts „Jährliche Rate der transfundierten Erythrozytenkonzentrate“ (= JR EK) erfolgte auf Basis der ITT-Population und war definiert als:

$$JR\ EK = \frac{\text{Gesamtzahl JR EK einer Person während der Behandlungsphase} \times 365}{(\text{Behandlungsende} - \text{Datum der Randomisierung} + 1)}$$

Das Behandlungsende war definiert als der Zeitpunkt der letzten Dosis + 14 Tage (entspricht bei regulärer Studienteilnahme der Abschlussuntersuchung der Behandlungsphase in Woche 52). Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt). Änderungen des Auswertungszeitraumes können Kapitel 2.2 (Abschnitt „Protokolländerungen“) entnommen werden. Die primäre Auswertung erfolgte auf Basis der ITT-Population.

### *Bewertung*

Aufgrund der Ausschlusskriterien wird im Rahmen der Wirksamkeitsevaluation darauf geschlossen, dass Transfusionen unregelmäßig bei Bedarf und nicht elektiv prophylaktisch stattfanden. Es konnten jedoch keine expliziten Informationen identifiziert werden, ob nach Studieneinschluss, d. h. während der Studienteilnahme, die Planung und Behandlung von Patientinnen und Patienten mit Transfusionen gemäß den genannten Ausschlusskriterien (z. B. im Rahmen eines chronischen Transfusionsprogramms) gestattet war oder bspw. in einem Studienabbruch resultierten. Durch die Ausschlusskriterien wurde aber in jedem Fall eine Wahrscheinlichkeitseingrenzung des Auftretens von elektiven Austausch- bzw. chronischen Transfusionen beider Arten vorgenommen und die Erfassung vermutlich weitestgehend auf Personen mit Notfalltransfusionen oder vereinzelt elektiven Einfachtransfusionen beschränkt.

Einschränkungen im Zusammenhang mit der Annualisierung von Zählzeiten, welche vor allem bei fehlenden Daten aufgrund von vorzeitigen Studienabbrüchen zum Tragen kommen und bereits zum primären Endpunkt (siehe Abschnitt „Vasookklusive Schmerzkrisen“) aufgeführt wurden, sind auch auf den vorliegenden Endpunkt anwendbar und diesbezügliche Implikationen sind in Kapitel 2.4 näher erläutert.

### Patientenrelevanz/Validität

„Jährliche Rate der transfundierten Erythrozytenkonzentrate“ bzw. die Reduktion von Transfusionen mit Erythrozytenkonzentraten wird nicht als patientenrelevant erachtet. Es wird vorrangig davon ausgegangen, dass sich etwaige Vorteile des zu bewertenden Arzneimittels, die sich aus einem geringeren Bedarf an anderen Therapieoptionen, wie bspw. transfundierten Erythrozytenkonzentraten, ergeben, in anderen, patientenrelevanten Endpunkten, wie z. B. hinsichtlich der Krankheitssymptome, Lebensqualität oder Nebenwirkungen, widerspiegeln sollten. Zudem ist die externe Validität limitiert, da bereits zu Studieneinschluss Einschränkungen bezüglich dieser Behandlungsoption Anwendung fanden (siehe Ausschlusskriterien in Tabelle 2).

### **Jährliche Rate der Krankenhaustage**

Der Endpunkt „Jährliche Rate der Krankenhaustage“ und die Post-hoc-Auswertung „Jährliche Rate der Krankenhaustage aufgrund einer VOC“ werden in der Nutzenbewertung nicht berücksichtigt.

### Operationalisierung

#### *Beschreibung*

Der Endpunkt „Jährliche Rate Krankenhaustage“ schließt gemäß Modul 4 alle Krankenhaustage ein, unabhängig von deren Ursache [10]. Post hoc wurde zudem in Modul 4 eine Analyse des Endpunkts „Jährliche Rate der Krankenhaustage aufgrund einer VOC“ dargestellt. Weitere Informationen zur Operationalisierung der Post-hoc-Analyse konnten im Dossier nicht identifiziert werden. Hospitalisierungen im eCRF werden durch das ärztliche Prüfpersonal im Rahmen anderer Endpunkte, bspw. von VOC und als Teil der Sicherheitserhebung, erfasst. Für die Hospitalisierungsdauer aufgrund von VOC, die für beide Auswertungen relevant ist, konnte nicht die gesamte

Systematik der Erhebungszeitpunkte identifiziert werden. VOC sollten jedoch über den gesamten Studienverlauf erfasst werden. Entsprechende Ausführungen zur Erhebung von VOC können dem Abschnitt „Vasookklusive Schmerzkrise“ entnommen werden. Es wird davon ausgegangen, dass sich die Post-hoc-Analyse ausschließlich auf diese Erhebungen (von VOC) stützt. Im Folgenden sind die Erhebungszeitpunkte von Hospitalisierungen auf Grundlage der Sicherheitserhebung und der Erhebung der medizinischen Vorgeschichte aufgeführt. Diese sind neben der Hospitalisierungsdauer aufgrund von VOC für die Auswertung der jährlichen Rate der Krankenhaustage relevant.

#### *Erhebungszeitpunkte (gemäß Erhebung der unerwünschten Ereignisse)*

- Screeningphase: bis zu 30 Tage vor Studieneinschluss
- Behandlungsphase: Tag 1, Tag 15, dann im 4-Wochen-Rhythmus bis Woche 50 mit einer Abschlussuntersuchung in Woche 52
- Follow-up-Phase: Woche 58

Nach etwaigem Behandlungsabbruch sollte die Behandlungsabschlussvisite so zeitnah wie möglich durchgeführt werden. Zudem sollte die Follow-up-Visite nach Möglichkeit ungefähr 8 Wochen nach der letzten Dosisgabe erfolgen.

#### *Auswertung*

Die primäre Auswertung des Endpunkts „Jährliche Rate Krankenhaustage“ (= JR KT) erfolgte auf Basis der ITT-Population und war definiert als:

$$JR\ KT = \frac{\text{Anzahl der Krankenhaustage einer Person} \times 365}{(\text{Behandlungsende} - \text{Datum der Randomisierung} + 1)}$$

Das Behandlungsende war definiert als der Zeitpunkt der letzten Dosis + 14 Tage (entspricht bei regulärer Studienteilnahme der Abschlussuntersuchung der Behandlungsphase in Woche 52). Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt). Änderungen des Auswertungszeitraumes können Kapitel 2.2 (Abschnitt „Protokolländerungen“) entnommen werden. Die primäre Auswertung erfolgte auf Basis der ITT-Population.

#### *Bewertung*

Die Operationalisierung des Endpunkts „Jährliche Rate der Krankenhaustage“ erfasst Hospitalisierungen jeglichen Grundes. Da dieses auch Ereignisse umfassen kann, die über indikations- und therapiespezifische Hospitalisierungen hinausgehen, erscheint die Operationalisierung nicht gänzlich nachvollziehbar. Hinsichtlich der Post-hoc-Analyse ist zu beachten, dass sich weitere Unklarheiten im Zusammenhang mit der uneindeutigen Abgrenzung der VOC voneinander (und deren Zählweise) sowie der fehlenden Standardisierung des Beginns und Endes ergeben (siehe Abschnitt „Vasookklusive Schmerzkrise“).

Einschränkungen im Zusammenhang mit der Annualisierung von Zählzeiten, welche vor allem bei fehlenden Daten aufgrund von vorzeitigem Studienabbrüchen zum Tragen kommen und bereits zum primären Endpunkt (siehe Abschnitt „Vasookklusive Schmerzkrise“) aufgeführt wurden, sind auch auf den vorliegenden Endpunkt anwendbar und diesbezügliche Implikationen sind in Kapitel 2.4 näher erläutert.

### Patientenrelevanz/Validität

Sowohl der Endpunkt „Jährliche Rate der Krankenhaustage“, d. h. die annualisierte Rate der Verweildauer im Krankenhaus im Allgemeinen, als auch die Post-hoc-Analyse zum Endpunkt „Jährliche Rate der Krankenhaustage aufgrund einer VOC“ werden nicht als patientenrelevant erachtet. Vorrangig wird davon ausgegangen, dass sich etwaige Vorteile des zu bewertenden Arzneimittels, die sich aus der Verkürzung der Krankheitsdauer ergeben, in anderen, patientenrelevanten Endpunkten, wie z. B. hinsichtlich der Krankheitssymptome, Lebensqualität oder Nebenwirkungen, widerspiegeln sollten. Außerdem können Krankenhaustage jeglicher Art durch regionale Unterschiede bzw. differierende Versorgungskontexte stark variieren. Dieser Aspekt wird durch eine Erhebung in 60 Studienzentren außerhalb Europas, vorwiegend in den USA, erhärtet. Es ergeben sich insbesondere hinsichtlich der Post-hoc-Analyse weitere Einschränkungen der Validität, welche auf die beschriebenen Mängel der Operationalisierung und die Beeinflussung der Krisendauer durch Analgetika zurückgehen. Die Vorstellung in einer (jeglichen) medizinischen Einrichtung aufgrund von VOC wird zudem bereits durch den primären Endpunkt („Jährliche Rate VOC“) annualisiert ausgewertet und unterliegt vergleichsweise geringeren Einschränkung bezüglich der systematischen Erfassung.

### **2.3.3 Lebensqualität**

#### **SF-36**

Der Endpunkt „SF-36“ wird in der Nutzenbewertung ergänzend im Anhang dargestellt.

#### Operationalisierung

##### *Beschreibung*

Zur Erhebung der Lebensqualität wurde Version 2 des 36-item Short Form Health Survey (SF-36) eingesetzt.

Der SF-36 beinhaltet acht Subskalen: „körperliche Funktionsfähigkeit“, „körperliche Rollenfunktion“, „körperlicher Schmerz“, „allgemeine Gesundheitswahrnehmung“, „Vitalität“, „soziale Funktionsfähigkeit“, „emotionale Rollenfunktion“ und „psychisches Wohlbefinden“ sowie zwei Summenskalen bzw. Summenscores. Der „Körperliche Summenscore“ (PCS) schließt die Subskalen „körperliche Funktionsfähigkeit“, „körperliche Rollenfunktion“, „körperliche Schmerzen“ und „allgemeine Gesundheitswahrnehmung“ ein; in den „Psychischen Summenscore“ (MCS) gehen die Subskalen für „Vitalität“, „soziale Funktionsfähigkeit“, „emotionale Rollenfunktion“ und „psychisches Wohlbefinden“ ein.

Die Werte der Summen- und Subskalen (T-Scores) sind auf einen Mittelwert von 50 und einer Standardabweichung von 10 der Normstichprobe (Referenzpopulation ist die US-Normalbevölkerung) transformiert. Ein höherer Wert spiegelt einen besseren Gesundheitszustand wider. In der Studie SUSTAIN bezogen sich die Items des SF-36 auf die Lebensqualität der vergangenen Woche.

##### *Erhebungszeitpunkte*

- Behandlungsphase: Tag 1, Tag 15, zu Woche 14, 26, 38 mit einer Abschlussuntersuchung in Woche 52
- Follow-up-Phase: Woche 58

Nach etwaigem Behandlungsabbruch sollte die Behandlungsabschlussvisite so zeitnah wie möglich durchgeführt werden. Zudem sollte die Follow-up-Visite nach Möglichkeit ungefähr 8 Wochen nach der letzten Dosisgabe erfolgen.

#### *Auswertung*

Die primäre Auswertung erfolgte auf Basis der ITT-Population bis Woche 58. Die Analyse basierte auf Mittelwertdifferenzen, die mittels eines linearen MMRM geschätzt wurden (Details siehe Kapitel 2.4 – Statistische Analysen). Zudem war eine Responderanalyse geplant, bei der eine Response-schwelle von einer Verbesserung von 10 oder mehr Punkten festgelegt wurde.

#### *Bewertung*

Die Operationalisierung ist insgesamt nachvollziehbar. Das Scoring erfolgte entsprechend Manual. Ähnlich wie bei dem Endpunkt „BPI-LF“ wurden die Erhebungszeitpunkte des Endpunkts „SF-36“ a priori festgelegt. Eine systematische Erfassung des SF-36 im Rahmen einer VOC, die die Einschätzung der Lebensqualität beeinflussen könnte, erfolgte nicht.

#### 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“.

Die Erfassung von gesundheitsbezogener Lebensqualität durch den SF-36 wird als patientenrelevant bewertet.

#### Validität

Der SF-36 ist ein häufig genutzter, valider, generischer Fragebogen zur gesundheitsbezogenen Lebensqualität.

Die Auswertung des SF-36 erfolgte durch Responderanalysen (eine Response wurde gewertet, wenn eine Verbesserung von  $\geq 10$  Punkten auf der Skala der T-Scores gegenüber der Baseline-messung vorlag) sowie der Analyse kontinuierlicher Werte mittels MMRM. Gemäß IQWiG-Methodenpapier 6.0 [5] sollte eine Responseschwelle, damit sie hinreichend sicher eine für den Patienten bzw. die Patientin spürbare Veränderung abbildet, mindestens 15 % der Skalenspannweite eines Instruments entsprechen. Bei Betrachtung der tatsächlich beobachteten Spannweiten der Summenskalen (PCS und MCS), entspricht diese Responseschwelle von 10 Punkten einer Responseschwelle von 15 % der empirisch ermittelten Skalenspannweiten aus der Normstichprobe. Das IQWiG hat für beide Summenskalen des SF-36 ein Responsekriterium von knapp 10 Punkten geschätzt (PCS: 9,4 und MCS: 9,6) [7]. Gemäß dieser Normstichprobe des SF-36 aus dem Jahr 2009, liegen die empirischen Skalenweiten der Subskalen i. d. R. unter den Skalenweiten der Summenskalen [8], sodass die vom pU verwendeten 10 Punkte für das Responsekriterium der Subskala „körperlicher Schmerz“ nicht den 15 % der tatsächlich genutzten Skalenspannweite entsprechen (sondern höheren Prozentzahlen). Darüber hinaus ist unklar, warum für die Responderanalyse ausschließlich die Subskala „körperlicher Schmerz“ berichtet wird, obwohl stets der komplette Fragebogen erhoben wurde. Die Responderanalysen werden daher aufgrund der fehlenden Vollständigkeit der Angaben und der fehlenden Begründung für die selektive Auswertung und Darstellung nur einer Subskala nicht dargestellt.

Die Auswertungen für die kontinuierlichen Daten gemäß MMRM sind gegenüber den Responderanalysen nachrangig zu bewerten und werden ergänzend im Anhang dargestellt.

## 2.3.4 Sicherheit

### Unerwünschte Ereignisse

Der Endpunkt „Unerwünschte Ereignisse“ wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

Unerwünschte Ereignisse (UE) waren definiert als jegliches unerwünschte medizinische Ereignis, d. h. jegliches unerwünschte und unbeabsichtigte Anzeichen (inklusive abnormaler Laborwerte), Symptom oder jegliche Erkrankung, die zeitlich mit der Studienmedikationsbehandlung zusammenhängt, unabhängig vom Vorliegen eines direkten Zusammenhangs mit der Studienmedikation.

Das ärztliche Prüfpersonal sollte jedes UE u. a. gemäß folgender Aspekte klassifizieren und evaluieren:

- Beschreibung des Ereignisses (bei Vorliegen einer Gruppe an Anzeichen und/oder Symptomen ist eine Diagnose (z. B. „Grippe“) zu berichten, anstatt jedes einzelne Anzeichen und Symptom)
- Datenangaben zu Beginn und Ende (wenn gegeben)
- Ernsthaftigkeit (siehe Definition von SUE)
- Schweregrad
  - mild: beeinträchtigt tägliche Aktivitäten nicht
  - moderat: beeinträchtigt tägliche Aktivitäten
  - schwer: verhindert tägliche Aktivitäten
- Ergriffene Maßnahme
  - keine
  - benötigte Begleitmedikation
  - temporäre Unterbrechung der Studienmedikation
  - Abbruch der Studienmedikation
  - andere (samt Erklärung)
- Resultat (Erholung ohne Folgen, Erholung mit Folgen, anhaltend, unbekannt, Tod)

Chirurgische Eingriffe stellten keine UE dar. Jedes UE war nur einmal und mit dem maximalen, während dem Ereignis aufgetretenen Schweregrad auf der entsprechenden Seite im eCRF zu vermerken. UE konnten durch spontane Berichte der Patientinnen und Patienten oder durch standardisierte, verbale Sondierungen identifiziert werden. Die Erhebung, Dokumentation und Berichterstattung erfolgte gemäß ICH GCP Guidelines. Alle UE, d. h. auch SUE, wurden während der gesamten Behandlungsphase und bis zur Follow-up-Visite erhoben, wobei je Behandlungstermin jeweils eine Erhebung vor, während und in Form einer Beobachtungszeit bis eine Stunde nach Studienmedikationsgabe stattfinden sollte. Außerdem mussten auffällige, klinisch relevante Ergebnisse der körperlichen, Vitalzeichen-, Biomarker-, Immunogenität-, EKG-, Labor- und Schwangerschaftsuntersuchungen als UE zu bestimmten Zeiten erfasst und berichtet werden. Gegebenenfalls wurden diese an Behandlungsterminen jeweils vor der Studienmedikationsgabe erhoben.

Behandlungsbezogene unerwünschte Ereignisse (TEAE) waren definiert als jegliches Anzeichen, Symptom, Syndrom oder jegliche neue Erkrankung, welche/s nach Behandlungsbeginn (und bis zur Follow-up-Visite) auftrat oder sich bei Vorliegen zu Baseline während der restlichen Studienbeobachtungszeit verschlimmert und das Wohlbefinden einer Patientin oder eines Patienten beeinträchtigen kann. Medizinisch relevante Ergebnisse von Laboruntersuchungen und anderen

Diagnoseuntersuchungen (die bspw. ungeplante Diagnoseuntersuchungen, Behandlungsanpassungen oder einen Studienabbruch zur Folge haben) waren von dieser Definition umfasst. Ein TEAE konnte in Abhängigkeit von der Studienmedikation oder unabhängig von der Studienteilnahme sein, musste jedoch zwischen dem ersten Behandlungstag und der Follow-up-Visite auftreten. Zudem wurde eine Schwangerschaft als TEAE erfasst und hatte bei diagnostischer Bestätigung durch einen Serum-Schwangerschaftstest, den Behandlungsabbruch, die Aufhebung der Verblindung sowie die Nachbeobachtung bis zur Geburt zur Folge.

Ein schwerwiegendes unerwünschtes Ereignis (SUE) war ein jegliches UE, das während jeglichem Zeitpunkt in der Beobachtungszeit eines der folgenden Kriterien erfüllte:

- führte zum Tod
- war direkt lebensbedrohlich
- erforderte oder verlängerte eine Hospitalisierung
- führte zu einer anhaltenden oder schwerwiegenden Behinderung/Unfähigkeit oder einer substantiellen Beeinträchtigung der Fähigkeit normale Lebensfunktionen aufrechtzuerhalten
- führte zu einer kongenitalen Anomalie / angeborenen Fehlbildung
- war medizinisch bedeutsam
- erforderte eine medizinische Intervention, um permanente Beeinträchtigungen oder Schäden zu vermeiden

Krankenhausaufnahmen und/oder chirurgische Eingriffe, die vor oder während der Studie geplant wurden, wurden nicht als UE / SUE eingestuft, solange die zugrunde liegende Erkrankung vor Studieneinschluss bereits vorlag und sich während der Studie nicht auf unerwartete Weise verschlechterte.

UE und SUE welche der Definition der Subklassifizierungen von VOC (siehe Tabelle 5) gemäß primärem Endpunkt entsprechen, wurden im Rahmen desselben erhoben und sollten nicht als UE oder SUE berichtet werden.

Im Rahmen des Zulassungsprozesses wurden seitens des pU post hoc UE von besonderem Interesse definiert und ausgewertet [3]. Gemäß Modul 4 umfassten diese „Infektionen“, „Infusionsbedingte Reaktionen“, „Schwere infusionsbedingte Reaktionen“ und „Wirkung auf die Blutgerinnung – Blutung (einschließlich veränderter Gerinnungsparameter)“ [10].

#### *Erhebungszeitpunkte*

- Behandlungsphase: Tag 1, Tag 15, dann im 4-Wochen-Rhythmus bis Woche 50 mit einer Abschlussuntersuchung in Woche 52
- Follow-up-Phase: Woche 58

Alle UE, welche den Studienabbruch einer Person zur Folge hatten, sollten bis zur abschließenden Klärung oder Stabilisation weiterverfolgt werden. Bei Behandlungsabbruch aus anderen Gründen sollte die Behandlungsabschlussvisite so zeitnah wie möglich durchgeführt werden. Zudem sollte die Follow-up-Visite nach Möglichkeit ungefähr 8 Wochen nach der letzten Dosisgabe erfolgen.

### *Auswertung*

Die Auswertung erfolgte auf Basis der Sicherheits-Population. UE werden zur Auswertung folgendermaßen klassifiziert:

- Auftreten vor Behandlung:  
UE, die nach Unterschrift der Einwilligungserklärung, aber vor erster Studienmedikationsgabe auftraten bzw. sich verschlimmerten
- Auftreten während der Behandlung:  
UE, die nach der ersten Studienmedikationsgabe bis zum Abschluss der Follow-up-Visite auftraten bzw. sich verschlimmerten
- Auftreten nach der Behandlung:  
UE, die nach Abschluss der Follow-up-Visite auftraten bzw. sich verschlimmerten

Für UE, welche fehlende Werte in Datumsangaben zu ihrem jeweiligen Beginn aufwiesen, wurde für eine fehlende Monatsangabe der Monat Januar und für eine fehlende Tagesangabe eine 1 imputiert und bei fehlenden Datumsangaben zum Ende wurden die jeweiligen Datumsangaben zum spätesten möglichen Zeitpunkt (Dezember oder letzter Tag des Monats) imputiert. War es nicht möglich ein UE anhand der vorliegenden Informationen als behandlungsbezogen oder nicht zu klassifizieren, so wurde dieses als behandlungsbezogen (d. h. als TEAE) klassifiziert.

Zur Auswertung von UE wurden ausschließlich zusammenfassende Tabellen zu TEAE präspezifiziert und im Rahmen dessen nach Systemorganklassen und Preferred Terms mit MedDRA (Medical Dictionary for Regulatory Activities), gemäß Studienbericht Version 16, kodiert. Diese Tabellen sollten folgendes umfassen:

- Alle TEAE, alle TEAE nach Schweregrad oder alle TEAE nach Zusammenhang
- Mit der Studienmedikation zusammenhängende TEAE
- TEAE, die zu einem Behandlungsabbruch führten
- Schwerwiegende TEAE
- Häufig berichtete TEAE

Alle TEAE wurden zudem jeweils nach Subgruppe der HU-Therapie und kategorisierter vaso-okklusiver Schmerzkrisenanzahl dargestellt.

Einem vom Sponsor eingesetzten Data Monitoring Committee (DMC), bestehend aus drei Klinikern (zwei Hämatologen/Sichelzellexperten und einem Spezialisten für Infektionserkrankungen) und einem Biostatistiker, oblag es während der Studie regelmäßig ausschließlich deskriptive Analysen von Sicherheitsdaten und Sicherheitsdaten-stützende Daten der Studie anhand pseudonymisierter Behandlungsgruppen (d. h. unter Benutzung der Buchstaben A, B und C) zu überwachen, jedoch keinerlei gruppenvergleichenden statistischen Tests durchzuführen.

### *Bewertung*

Die Operationalisierung ist weitestgehend nachvollziehbar. Einschränkungen ergeben sich aufgrund der Differenzierungs- und Spezifizierungsgrade der Einschätzungen von Schweregraden. Diese erscheinen bspw. gegenüber einer CTCAE-Gradierung limitiert. Auch erscheint unklar, inwiefern bspw. Laborwerte basierend auf Auswirkungen auf Alltagsaktivitäten klassifiziert werden. Es liegen keine Informationen vor, inwiefern UE, die nach Abschluss der Follow-up-Visite auftraten, erfasst wurden, da entsprechende Erhebungszeitpunkte nicht vorgesehen waren.

### Patientenrelevanz/Validität

Das Auftreten unerwünschter Ereignisse wird i. d. R. als patientenrelevant und die Erhebung in der vorliegenden Studie SUSTAIN als größtenteils valide angesehen. Einschränkungen bestehen hinsichtlich der Aussagekraft einzelner UE (wie bspw. Laborparameter oder Schwangerschaftstestergebnisse) deren Patientenrelevanz, bspw. als potentielle Surrogatparameter, nicht dargestellt ist. Außerdem erscheint die Reliabilität der Schweregradeinteilung aufgrund der Operationalisierung (siehe Bewertung) limitiert.

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“.

Die Patientenrelevanz von Laborparametern ist unklar.

### 2.3.5 Übersicht der Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 7. Erhebungszeitpunkte und Auswertungszeitraum können voneinander abweichen.

*Tabelle 7: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie SUSTAIN*

Studien- visite  Endpunkt	Screening	Behandlungsphase										Follow- up	
	Tag -30 bis -1	Behandlung										Abschluss- unter- suchung	Unter- suchung
		Tag 1	Tag 15	Woche									
				6,10	14	18, 22	26	30, 34	38	42, 46, 50	Woche 52	Woche 58	
Mortalität <sup>1)</sup>		x	x	x	x	x	x	x	x	x	x	x	x
VOC <sup>2)</sup>	x <sup>3)</sup>	kontinuierliche Erfassung <sup>3)4)</sup>											
BPI-LF SF-36		x	x		x		x		x			x	x
Unerwünschte Ereignisse <sup>5)</sup>		x	x	x	x	x	x	x	x	x	x	x	x

<sup>1)</sup> Mortalität war kein eigenständiger Endpunkt und wurde im Rahmen der „Unerwünschten Ereignisse“ erhoben.

<sup>2)</sup> Die dargestellten Informationen bilden den innerhalb der Studie vorliegenden Erfassungszeitraum aller in der Studie operationalisierten Arten von VOC ab.

<sup>3)</sup> Zur Screeningvisite wurde eine retrospektive Erhebung der VOC in den vergangenen 12 Monaten basierend auf medizinischen Berichten und/oder Erinnerungen der Patientinnen und Patienten durchgeführt. Diese Erhebung wurde zum ersten Behandlungstag aktualisiert und unterscheidet sich von der standardisierten Erfassung im weiteren Studienverlauf.

<sup>4)</sup> Während der gesamten Behandlungsphase bis zur Follow-up-Visite sollten VOC erhoben werden. VOC, die in medizinischen Einrichtungen außerhalb der Studienzentren behandelt wurden, konnten anhand einer individuellen Studienteilnahmekarte erfasst werden. Die Auslesezeitpunkte der Studienteilnahmekarten bzw. die Kenntnisnahmezeitpunkte über diese VOC seitens der Studienzentren sind unklar.

<sup>5)</sup> Zu allen dargestellten Zeitpunkten erfolgte als Teil der Sicherheitserhebung u. a. auch die Erhebung von klinischen Laboruntersuchungen, zusätzlichen Laboruntersuchungen, Untersuchungen zu EKG, Vitalzeichen und Körpergewicht, die Durchführung von körperlichen Untersuchungen (variierten Umfangs) sowie für Patientinnen die Durchführung von Schwangerschaftstests.

Abkürzungen: BIP-LF: Brief Pain Inventory – Long Form; SF-36: 36-item Short Form Health Survey; VOC: Vasookklusive Schmerzkrise.

## 2.4 Statistische Methoden

In der Studie SUSTAIN wurde die Überlegenheit von Crizanlizumab gegenüber Placebo in einem randomisierten, doppelblinden, zweifach stratifizierten Studiendesign getestet. Die Fallzahlplanung mit insgesamt 174 Patientinnen und Patienten (oder 58 Personen je Studienarm) beruhte auf Überlegungen zum Endpunkt „Jährliche Rate VOC“ und erfolgte unter Berücksichtigung der Stratifizierung nach Begleitbehandlung mit HU und der Anzahl an VOC in den vergangenen 12 Monaten. Die Patientinnen und Patienten sollten an insgesamt 70 Studienzentren über eine Studienlaufzeit von insgesamt 25 Monaten (12 Monate Rekrutierungs- sowie 13 Monate Behandlungs- und Follow-up-Phase) rekrutiert und untersucht werden. In den Studienunterlagen konnte nicht identifiziert werden, warum über eine längere Gesamtstudiendauer (vom ersten Studieneinschluss bis zum Studienabschluss der letzten Person) von etwa 32 Monaten, eine größere Anzahl an Patientinnen und Patienten (N = 198) an weniger Studienzentren (N = 60) rekrutiert und untersucht wurden.

Die vorliegenden Analysen erfolgten auf Basis des finalen SAP vom 2. Mai 2016 und weiteren Post-hoc-Analysen, die im Rahmen des Zulassungsprozesses durchgeführt wurden oder im Rahmen des Dossiers dargestellt sind. Das Datum des Protokolls 4 (21.04.2016), welches Änderungen zur statistischen Auswertung (siehe Tabelle 4) beinhaltet, sowie des finalen SAP liegen zeitlich nach dem Studienende (Studienabschluss letzte Person 23.03.2016). Nach Angaben der EMA wurde die Datenbank am 13. Mai 2016 gesperrt [3]. Der pU gibt in Modul 4 an, dass Patientinnen und Patienten, ärztliches Prüfpersonal, Mitglieder des CRC sowie der Sponsor bis dahin verblindet gewesen seien [10]. Gemäß Angaben gegenüber der EMA wiesen davon möglicherweise etwa 7 % der Patientinnen und Patienten Protokollverletzungen im Zusammenhang mit potentieller Entblindung auf [3]. Es liegen keine Informationen über die Verblindung bzw. explizit den Entblindungszeitpunkt der Statistiker vor.

### Analysepopulationen

Intention-to-Treat (ITT; N = 132): Alle Personen, die auf die Studienarme randomisiert wurden. Patientendaten wurden nach der Behandlungsgruppe ausgewertet, in die sie jeweils randomisiert wurden.

Modifizierte Intention-to-Treat (mITT; N = 128): Alle Personen, die auf die Studienarme randomisiert wurden und mindestens eine Dosis Studienmedikation erhalten haben. Patientendaten wurden nach der Behandlungsgruppe ausgewertet, in die sie randomisiert wurden.

Sicherheitspopulation (N = 128): Alle auf die Studienarme randomisierten Personen, die mindestens eine Dosis der Studienmedikation erhalten haben. Patientendaten wurden nach tatsächlich erhaltener Behandlung ausgewertet.

Per Protocol (PP; N = 81): Alle ITT-Personen, die mindestens 12 der geplanten 14 Studienmedikationsgaben erhalten haben, die Abschlussuntersuchung mindestens 14 Tage nach der letzten Studienmedikationsgabe wahrgenommen haben und die Studie ohne Protokollverletzungen hinsichtlich der Anzahl der VOC im vergangenen Jahr vor Studieneinschluss (d. h. weder < 2 noch > 10 VOC), welche die Bewertung der Wirksamkeit beeinträchtigen, abgeschlossen haben. Die PP-Population wurde vor Datenbanksperre dokumentiert. Patientendaten wurden nach der Behandlungsgruppe ausgewertet, in die sie randomisiert wurden.

### **Protokolländerungen: Auswertungen, statistische Analysen und Verblindung**

Aus den Studienunterlagen ergeben sich Unklarheiten bezüglich der Durchführung etwaiger Interimsanalysen, inwieweit diese ggf. verblindet waren und ob diese Einfluss auf Änderungen in Protokollversion 4 sowie die Erstellung der finalen Version des SAP hatten. Hintergründe zu diesen Unklarheiten sind im Folgenden erläutert.

Laut der zusammenfassenden Darstellung im Studienbericht sahen Änderungen in Protokollversion 3, welche vor Studienende Gültigkeit erlangte, unter anderem vor, dass ein stratifizierter Test Anwendung finden sollte, um (den primären Endpunkt zu analysieren und) einen sekundären Schlüsselendpunkt zu identifizieren (siehe Tabelle 4). Diese referenzierte Änderung und etwaige diesbezügliche Erklärungen konnten in Protokollversion 3 nicht identifiziert werden. In Protokollversion 4 (nach Studienende) wird jedoch – wie in Protokoll 3 angekündigt – ein sekundärer Schlüsselendpunkt definiert, welches auf die Durchführung eines stratifizierten Tests hindeutet. Wie in Kapitel 2.3.4 dargestellt sollten während der Studie jedoch ausschließlich deskriptive Analysen von Sicherheitsdaten und Sicherheitsdaten-stützende Daten der Studie anhand pseudonymisierter Behandlungsgruppen durch das DMC und keinerlei gruppenvergleichenden statistischen Tests durchgeführt werden. Sicherheitsdaten-stützende Daten waren nicht definiert. Während der Studie wurden gemäß den Protokollen explizit keinerlei Interimsanalysen hinsichtlich der Wirksamkeit geplant, noch konnten diesbezügliche Ergebnisse in den Studienunterlagen identifiziert werden. Demzufolge erscheint nicht nachvollziehbar, warum in der Protokollzusammenfassung zu Version 3 ein statistischer Test zur Identifikation eines sekundären Schlüsselendpunktes angekündigt wurde sowie ob, inwiefern, unter welchen Bedingungen – insbesondere bezüglich der Verblindung – und wann statistische Tests vor der abschließenden Datenanalyse stattgefunden haben. Ob etwaige Ergebnisse weitere Änderungen in Protokoll 4 und/oder SAP begründeten, die beide nach Studienende finalisiert wurden, ist demnach unklar. Dieses würde potentiellen Einfluss auf die Bewertung der Änderungen bezüglich des Analysevorgehens haben. Die im Rahmen der Nutzenbewertung diesbezüglich relevanten Änderungen betreffen insbesondere die primäre Analyse (z. B. die Berücksichtigung der Randomisierungsstrata) aller Wirksamkeitsendpunkte (mit Ausnahme von BPI-LF und SF-36) sowie Veränderungen der Sensitivitäts- und Subgruppenanalysen (siehe Tabelle 4 und Tabelle 26 im Anhang). Gegenüber der EMA berichtete der pU zwar die Datenbanksperrung habe zum 13. Mai 2016 (nach Finalisierung von Protokoll 4 und SAP) stattgefunden und gemäß Dossier habe die Aufhebung der Verblindung für Patientinnen und Patienten, ärztliches Prüfpersonal, Mitglieder des CRC sowie den Sponsor erst zu diesem Zeitpunkt stattgefunden, jedoch finden sich keine Erklärungen für die berichteten Inkonsistenzen. Ob die Wirksamkeitsdaten vor Finalisierung von Protokoll 4 und SAP und, wenn ja, von wem (bspw. vom DMC, welches nur eingeschränkt verblindet war) bereits analysiert wurden, bleibt unklar. Angaben zur Aufhebung der Verblindung der Statistiker liegen zudem nicht vor.

### **Statistische Analysen**

Der primäre Endpunkt „Jährliche Rate VOC“ wurde zur primären Analyse der Daten zum Ende der 52-wöchigen Behandlungsphase mittels eines Wilcoxon-Rangsummen-Tests getestet. Die Analyse wurde unter Berücksichtigung der Randomisierungsstrata, der Anzahl der VOC in den letzten 12 Monaten (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein), vorgenommen. Die Berechnung der Mediane und die Schätzung des Medians der Differenzen der Studienarme einschließlich zugehörigem 95%-Konfidenzintervall (KI) erfolgte mithilfe der Hodges-

Lehmann (HL)-Methode. Um die Ergebnisse eines Wilcoxon-Rangsummen-Tests zu interpretieren, müssen u. a. statistische Annahmen über die Verteilungen der beiden Vergleichsgruppen berücksichtigt werden. Sollten die Verteilungsformen der beiden Vergleichsgruppen einander ähnlich (bspw. sind beide unimodal und rechtsschief) und deren Lage voneinander verschoben sein (auch „location shift“ genannt), so eignet sich der Wilcoxon-Rangsummen-Test i. d. R. dazu, Interpretationen über die etwaigen Unterschiede der beiden Mediane anzustellen. Ist eine dieser beiden Annahmen verletzt, so könnte bspw. ein signifikantes Ergebnis trotz gleicher bzw. ähnlicher Mediane vorliegen und stattdessen auf unterschiedliche Verteilungen hindeuten. Die Überprüfung der beiden statistischen Annahmen kann u. a. anhand von Histogrammen vorgenommen werden, wobei zu Studie SUSTAIN nur eines für den Endpunkt „Jährliche Rate VOC“ berichtet wurde.

Neben der beschriebenen primären Analyse des Endpunkts „Jährliche Rate VOC“ wurde eine Vielzahl von Sensitivitätsanalysen geplant und teilweise durchgeführt, während zudem viele weitere Sensitivitätsanalysen post hoc vorgenommen wurden. Diese Sensitivitätsanalysenplanung und -durchführung umfasste u. a. die in Tabelle 26 im Anhang dargestellten Analysen. Von den prä-spezifizierten Sensitivitätsanalysen (PS) wird die Auswertung des primären Endpunkts „Jährliche Rate VOC“ mithilfe eines Negativ-Binomial-Regressionsmodells (Log-Link-Funktion) (PS-1) in der vorliegenden Nutzenbewertung herangezogen. Dieses erfolgte mit dem Logarithmus der Studiendauer als Offset-Variable und mit der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10), dem Einsatz einer begleitenden HU Therapie (ja vs. nein) sowie der Behandlung als unabhängige Variablen.

Die EMA konnte eine ergebnissteuerte Auswahl mancher Post-hoc-Sensitivitätsanalysen (PhS), welche im Rahmen des Zulassungsprozesses vorgelegt wurden, nicht ausschließen [3]. Daher erfolgt in der vorliegenden Nutzenbewertung eine Darstellung jener Sensitivitätsanalysen, deren Rationale weitestgehend nachvollziehbar begründet erscheinen. Folgende Negativ-Binomial-Regressionsmodelle mit Imputationen (PhS-M5, PhS-M6a und PhS-M6b) werden demnach herangezogen:

- PhS-M5 gemäß CRC  
Imputation der Werte von Personen mit Therapieabbruch innerhalb der ersten 6 Monate durch die Zahl der Krisenereignisse vor der Randomisierung
- PhS-M6a gemäß CRC  
Imputation der Zahl der VOC bei vorzeitigem Behandlungsabbruch für Personen im Crizanlizumab-Arm über eine Jump-to-Reference-Methode basierend auf Daten von Personen aus dem Placebo-Arm. Für Personen im Placebo-Arm erfolgte die Ersetzung unter der Missing-at-Random-Annahme basierend auf Daten von Personen desselben Studienarmes vor Behandlungsabbruch
- PhS-M6b gemäß ärztlichem Prüfpersonal  
Wie PhS-M6a, aber basierend auf der Bewertung durch das ärztliche Prüfpersonal und unter Ausschluss einer Person im Crizanlizumab-Arm, bei der es sich um einen statistischen Ausreißer mit 37 VOC in 6 Monaten handelt und die zudem die Einschlusskriterien nicht erfüllte, da sie bereits im Jahr vor der Randomisierung 30 Schmerzkrisen erlitten hatte.

Die mittels Wilcoxon-Rangsummen-Test und Hodges-Lehmann-Schätzer durchgeführte PhS-1 (siehe Tabelle 26 im Anhang) des Endpunkts „Jährliche Rate VOC“ gemäß der Bewertung durch das ärztliche Prüfpersonal sowie die primäre Analyse des Endpunkts „Jährliche Rate VOC“ gemäß CRC werden ebenfalls dargestellt. Diese unterliegen jedoch den oben beschriebenen Einschränkungen der statistischen Annahme ähnlicher voneinander verschobener Verteilungsformen.

Die primären Analysen der Endpunkte „Jährliche Rate unkomplizierte VOC“ und „Jährliche Rate akute Thoraxsyndrome“ wurden gemäß Präspezifizierung analog zum primären Endpunkt ausgewertet, d. h. mittels eines Wilcoxon-Rangsummen-Tests unter Berücksichtigung der Randomisierungsstrata bzw. mittels Hodges Lehman-Methode zur Schätzung des Medians der Differenzen der Studienarme (s. o.). Post-hoc-Sensitivitätsanalysen, wie sie zum primären Endpunkt Anwendung gefunden haben, konnten zu „Jährliche Rate unkomplizierte VOC“ lediglich gemäß der Imputationsmethode M6b im EPAR und im Bericht der EUnetHTA identifiziert werden [3,22] (Daten nicht im Rahmen des Dossiers vorgelegt). Für die sekundären Wirksamkeitsendpunkte „Jährliche Rate unkomplizierte VOC“ und „Jährliche Rate akute Thoraxsyndrome“ werden jeweils die primäre Analyse sowie für die Subklassifikation „Jährliche Rate unkomplizierte VOC“ die PhS-M6b dargestellt. Für den Endpunkt „Jährliche Rate akute Thoraxsyndrome“ konnte keine Sensitivitätsanalyse gemäß Imputationsmodell M5 oder M6a/b identifiziert werden. Einschränkungen zur primären Analyse der diskutierten sekundären Endpunkte müssen analog zu jenen der Analyse des primären Endpunkts berücksichtigt werden.

Die sekundären Endpunkte „Zeit bis zur ersten VOC“ und „Zeit bis zur zweiten VOC“ wurden anhand der Kaplan-Meier-Methode geschätzt. Der Vergleich der Studienarme erfolgte anhand eines Logrank-Tests, stratifiziert nach den Randomisierungsstrata. Die Schätzung des Hazard Ratios (HR) sowie des 95%-KI wurde anhand des Cox-Proportional-Hazards-Modells mit den Randomisierungsstrata und der Behandlung als Kovariaten vorgenommen. Die Auswertung zu diesen beiden Endpunkten basieren auf der Annahme der nicht-informativen Zensierung. Es wird zudem die im EPAR berichtete Post-hoc-Sensitivitätsanalyse nach Imputationsmodell M6b zur Nutzenbewertung herangezogen [3] (Daten nicht im Rahmen des Dossiers vorgelegt). Im vorliegenden Fall erfolgte diese, anders als bei den Endpunkten „Jährliche Rate VOC“ und „Jährliche Rate unkomplizierter VOC“, basierend auf der Schätzung des Kaplan-Meier-Medians und der Berechnung des HR mithilfe des Cox-Proportional-Hazards-Modells mit den Randomisierungsstrata und der Behandlung als Kovariaten.

Der patientenberichtete Endpunkt „SF-36“ wurde mit einem linearen MMRM hinsichtlich der Änderung gegenüber Baseline und mit Baselinewert, Randomisierungsstrata, Visite, Behandlung sowie der Interaktion von Behandlung und Visite als feste Effekte und der Patientin / dem Patienten als zufälligen Effekt analysiert. Als Ergebnis des MMRM wurden adjustierte Mittelwertdifferenzen (least squares mean differences) inklusive des 95%-KI für den Gruppenvergleich zu jeder Post-Baseline-Erhebung aus dem MMRM geschätzt. Eine Imputation für fehlende Daten wurde in den MMRM-Analysen nicht vorgenommen. Weiterhin wurde eine Responderanalyse mit einem a priori definierten Responsekriterium (eine Verbesserung von  $\geq 10$  Punkten gegenüber der Baselinemessung) ausschließlich für die Subskala „körperlicher Schmerz“ vorgelegt. Diese Analyse wird aufgrund der in Abschnitt 2.3.3 genannten Einschränkungen nicht dargestellt.

Die Darstellung der unerwünschten Ereignisse (TEAE und UE von besonderem Interesse) erfolgte deskriptiv. Zum Vergleich der beiden Studienarme wurde post hoc der Cochran-Mantel-Haenszel-Test unter Berücksichtigung der Randomisierungsstrata herangezogen.

Es wurden präspezifizierte Subgruppenanalysen für den primären Endpunkt differenziert nach HU-Therapie (ja vs. nein), kategorisierter Anzahl vorheriger VOC (2–4 vs. 5–10) und nach Genotyp (HbSS vs. alle Genotypen ausschließlich HbSS) durchgeführt. Außerdem erfolgten post hoc Berechnungen von Interaktionstermen (jeweils HU-Therapie, Anzahl vorheriger VOC, Sichelzellerkrankungstyp, Geschlecht, Ethnie und Alter) basierend auf einer Erweiterung des Negativ-Binomial-Regressionsmodells.

## 2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 8: Verzerrungspotential der Studie SUSTAIN

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patientin/Patient	Verblindung Behandlerin/Behandler	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studien-ebene
SUSTAIN	Ja <sup>1)</sup>	Ja <sup>2)</sup>	Ja <sup>3)4)</sup>	Ja <sup>4)5)</sup>	Ja <sup>6)7)</sup>	Ja <sup>8)9)10)</sup>	Hoch

<sup>1)</sup> Die Randomisierung erfolgte stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein). Die Randomisierungscodes und -schemata wurde durch den IXRS-Anbieter über ein automatisches System erstellt. Nach Angaben der EMA erfolgte diese anhand einer Blockrandomisierung mit einer Blockgröße von 9, wobei Studienzentren nicht in die Stratifizierung eingingen [3].

<sup>2)</sup> Es wurde ein zentrales Randomisierungsschema verwendet, d. h. das IXRS-System wies der Probandin / dem Probanden immer den nächsten verfügbaren Randomisierungsslot zu. Daraufhin wurde der unverblindete Pharmazeuten im jeweiligen Studienzentrum ggf. über den Einschluss einer Person informiert und trat mit dem externen Anbieter des IXRS-Systems in Kontakt, um die Studienmedikation vorzubereiten. Das ärztliche Prüfpersonal war verantwortlich für die Aufrechterhaltung der Verblindung am jeweiligen Studienzentrum. Diese durfte im Fall eines SUE bei dringender Notwendigkeit für den weiteren Behandlungsverlauf nach Einschätzung des ärztlichen Prüfpersonals im Kontakt mit dem Sponsor aufgehoben werden.

<sup>3)</sup> Patientinnen und Patienten, ärztliches Prüfpersonal, Mitglieder des CRC und Sponsor blieben i. d. R. vom Zeitpunkt der Randomisierung bis zum Datenbankschluss (am 13.05.2016 gemäß EPAR) verblindet [3]. Die Studienmedikation (Crizanlizumab oder Placebo) wurde in äußerlich nicht unterscheidbaren Infusionsbehältern verabreicht. Die Art und Weise, wie die Infusion verabreicht wurden, war ebenfalls identisch.

<sup>4)</sup> Es lagen Protokollverletzungen im Zusammenhang mit potentieller Entblindung vor, welche etwa 7 % aller Patientinnen und Patienten betreffen konnten (Daten nicht mit Dossier vorgelegt).

<sup>5)</sup> Ein unverblindeter Pharmazeut im jeweiligen Studienzentrum war für die Vorbereitung der Studienmedikation verantwortlich. Die jeweilige ärztliche Prüfperson war verantwortlich für die Aufrechterhaltung der Verblindung am Studienzentrum und konnte deren Aufhebung in bestimmten Fällen (bspw. bestimmte SUE, Notfälle, Schwangerschaft) veranlassen.

<sup>6)</sup> Es bestehen Unklarheiten hinsichtlich der etwaigen Durchführung und des Ausmaßes der Verblindung von Analysen nach Studienende aber vor Finalisierung von Protokoll 4 und SAP sowie vor der Datenbanksperre. Solche Analysen hätten ggf. Änderungen des Analyseplans beeinflussen können (siehe Kapitel 2.4 – Protokolländerungen: Auswertungen, statistische Analysen und Verblindung).

<sup>7)</sup> Präspezifizierte Analysen mit Imputationen wurden nicht durchgeführt und es wurde abweichend vorgegangen bzw. wurde eine dieser Analysen der EMA auf Aufforderung post hoc nachgereicht (M5). Die EMA konnte eine ergebnisgesteuerte Auswahl mancher präspezifizierten und Post-hoc-Sensitivitätsanalysen nicht ausschließen [3].

<sup>8)</sup> Es bestehen Unklarheiten hinsichtlich der Begründungen von Protokolländerungen, insbesondere der Auswertungs- und Analyseplanänderungen nach Studienende (siehe Fußnote 6).

<sup>9)</sup> Es konnte nicht identifiziert werden, warum gegenüber der Planung über eine längere Gesamtstudiendauer (~ 32 Monate ggü. 25 Monate) eine größere Anzahl an Personen (N = 198 ggü. N = 174) an weniger Studienzentren (N = 60 ggü. N = 70) rekrutiert, untersucht und ausgewertet wurden.

<sup>10)</sup> Zu den Baselinecharakteristika der eingeschlossenen Patientinnen und Patienten liegen nur limitierte Informationen vor. Insbesondere spezifischen Einschätzungen zur Krankheitsschwere und Angaben zur Dauer seit Diagnosestellung fehlen, sodass etwaige diesbezüglich Imbalancen der Behandlungsarme nicht überprüft werden konnten.

Abkürzungen: CRC: Crisis Review Committee; EMA: European Medicines Agency; EPAR: European Public Assessment Report; IXRS: Interactive Voice/Web Response System; SAP: Statistischer Analyseplan; SUE: Schwerwiegendes unerwünschtes Ereignis; VOC: Vasookklusive Schmerzkrise.

Das Verzerrungspotential auf Studienebene wird als hoch eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 9 dargestellt, wobei insbesondere die hohen Studienabbruchraten in beiden Studienarmen von mehr als einem Drittel zu berücksichtigen sind.

**Tabelle 9: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie SUSTAIN**

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Mortalität <sup>1)</sup>	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)5)</sup>	Unklar
<i>Vasookklusive Schmerzkrisen</i>					
Jährliche Rate VOC	Ja	Nein <sup>2)6)</sup>	Ja <sup>7)8)</sup>	Ja <sup>3)4)7)</sup>	Hoch
Zeit bis zur ersten VOC	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)9)</sup>	Hoch
Zeit bis zur zweiten VOC	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)9)</sup>	Hoch
VOC-Subklassifikationen:					
Jährliche Rate unkomplizierte VOC	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)7)</sup>	Hoch
Jährliche Rate akute Thoraxsyndrome	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)7)</sup>	Hoch
Lebersequestration	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)</sup>	Hoch
Milzsequestration	Ja	Nein <sup>2)6)10)</sup>	Nein	Ja <sup>3)4)</sup>	Hoch
Priapismus	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)</sup>	Hoch
SF-36	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)11)</sup>	Hoch
Unerwünschte Ereignisse	Ja	Nein <sup>2)6)</sup>	Nein	Ja <sup>3)4)12)</sup>	Hoch

<sup>1)</sup> Mortalität war kein eigenständiger Endpunkt und wurde im Rahmen der unerwünschten Ereignisse erhoben.

<sup>2)</sup> Etwas mehr als ein Drittel der Patientinnen und Patienten (jeweils 24 Personen, 35,8 vs. 36,9 %) brach die Studie vorzeitig ab. Es ist unklar, in welchem Umfang diese Abbrüche im Zusammenhang mit der Wirksamkeit oder Sicherheit standen. Es lagen keine Informationen zu den Charakteristiken (wie bspw. prognostische Faktoren, VOC, UE) der Studienabbrecher vor. Bei den Abbruchbegründungen kann es sich teilweise um konkurrierende Ereignisse zu VOC-spezifischen Endpunkten und Sicherheitserfassungen handeln, insbesondere, weil Personen, welche die Behandlung abbrachen, kürzer beobachtet wurden als regelhafte Studienteilnehmende.

<sup>3)</sup> Zu den Baselinecharakteristika der eingeschlossenen Patientinnen und Patienten liegen nur limitierte Informationen vor. Insbesondere spezifische Einschätzungen zur Krankheitsschwere und Angaben zur Dauer seit Diagnosestellung fehlen, sodass etwaige diesbezüglich Imbalancen der Behandlungsarme nicht überprüft werden konnten.

<sup>4)</sup> Während Angaben über die Behandlungsdauer und Studienmedikationsgaben vorliegen, konnten keine spezifischen Angaben zur Beobachtungsdauer identifiziert werden. Dieses ist insbesondere bezüglich Auswertungen der Sicherheitspopulation relevant, d. h. betreffend der Mortalität, der UE und manchen deskriptiven Auswertungen von VOC, da die Sicherheitsanalysen am Ende der regulären Beobachtungszeit einen Zeitraum von 8 Wochen ohne Behandlungen umfassen.

<sup>5)</sup> Im Rahmen der vorliegenden Indikation erscheint bei einer geplanten Beobachtungsdauer von 58 Wochen die zu erwartende Inzidenz von Todesfällen begrenzt zu sein, weshalb Auswirkungen des Verzerrungspotentials limitiert erscheinen.

<sup>6)</sup> Fünf Personen der ITT-Population wiesen Verstöße gegen das Einschlusskriterium von 2 bis 10 VOC in den vergangenen 12 Monaten vor Studieneinschluss auf. Verstöße gegen andere Einschlusskriterien wurden nicht dargestellt. Eine Person im Crizanlizumab-Arm wies bereits im Jahr vor der Randomisierung 30 Schmerzkrisen auf und stellte einen statistischen Ausreißer mit 37 behandlungsphasenbezogenen VOC in 6 Monaten dar. Diese Person wurde aus einer Sensitivitätsanalyse (PhS-M6b) ausgeschlossen während die anderen vier Personen in der Analyse verblieben.

<sup>7)</sup> Im Rahmen der primären Analyse deuten sowohl das Konfidenzintervall des Hodges-Lehmann-Schätzers als auch im Fall des Endpunkts „Jährliche Rate VOC“ eine graphische Darstellung darauf hin, dass die statistischen Annahmen über das Vorliegen ähnlicher Verteilungsformen beider Studienarme, die voneinander verschoben sein müssen, nicht erfüllt sind [3]. Unter diesen Voraussetzungen erscheint eine eindeutige Interpretation der Ergebnisse hinsichtlich der Signifikanz des Unterschieds der Mediane (Wilcoxon-Rangsummen-Test) und eine verlässliche Schätzung des Medians der Differenzen (Hodges-Lehmann-Schätzer) nicht möglich.

<sup>8)</sup> Die EMA konnte eine ergebnisgesteuerte Auswahl mancher Post-hoc-Sensitivitätsanalysen nicht ausschließen, weshalb in der vorliegenden Nutzenbewertung nur jene Post-hoc-Sensitivitätsanalysen berücksichtigt wurden, deren Rationale weitestgehend nachvollziehbar erscheinen.

<sup>9)</sup> Im Rahmen der Ergebnisdarstellung liegen nur eingeschränkte Informationen vor. Detaillierte Angaben zu Zensierungsgründen fehlen.

<sup>10)</sup> Es ist zu berücksichtigen, dass gemäß medizinischer Vorgeschichte 14,9 % der Personen im Crizanlizumab-Arm und 10,8 % der Personen im Placebo-Arm bereits vor Studienbeginn einer Splenektomie unterzogen wurden.

<sup>11)</sup> Aufgrund der prädefinierten Erhebungszeitpunkten erfolgte keine systematische Erfassung des SF-36 im Rahmen von VOC. Post-hoc-Analysen offenbarten, dass nur etwa 6,7 % der abgefragten (7-Tage-)Zeiträume eine VOC beinhalteten. Angaben, inwiefern sich diese auf die beiden Arme verteilen, liegen nicht vor.

<sup>12)</sup> Die Patientenrelevanz von Laborparametern, Schwangerschaftstest u. a. ist unklar.

Abkürzungen: EMA: European Medicines Agency; ITT: Intention-to-Treat; M: Imputationsmethode; PhS: Post-hoc-Sensitivitätsanalyse; SF-36: 36-item Short Form Health Survey; UE: Unerwünschtes Ereignis; VOC: Vasookklusive Schmerzkrise.

### 3 Ergebnisse der eingeschlossenen Studie

Mit der Studie SUSTAIN liegt eine Studie vor, welche die Wirksamkeit und Sicherheit von Crizanlizumab (Adakveo®) zur Prävention von wiederkehrenden vasookklusiven Schmerzkrise bei Sichelzellerkrankung untersucht.

#### 3.1 Studiencharakteristika und Studienmedikation

Die für die vorliegende Nutzenbewertung relevante Studienpopulation der Studie SUSTAIN besteht aus 67 Patientinnen und Patienten, die auf den Crizanlizumab-Arm mit einer Dosis von 5,0 mg/kg randomisiert wurden und 65 Patientinnen und Patienten, die auf dem Placebo-Arm randomisiert wurden. Von diesen brachen 35,8 % im Crizanlizumab-Arm (n = 24) und 36,9 % im Placebo-Arm (n = 24) die Studie vorzeitig ab (siehe Tabelle 10). Gemäß Studienangaben wiesen Personen im Crizanlizumab-Arm (40,3 %) etwas häufiger mindestens eine relevante Protokollverletzung auf als Personen im Placebo-Arm (36,9 %). Deskriptive Auswertungen zu etwaigen weiteren Protokollverletzungen, wie bspw. Dosisverzögerung, Aufhebung der Verblindung oder der Verletzung von anderen Ein- und Ausschlusskriterien (als der Anzahl der VOC in den vergangenen 12 Monaten) konnten nicht identifiziert werden. Während die Behandlungsdauer und Studienmedikationsgaben beider Arme deskriptiv vergleichbar waren, liegen keine näheren Angaben zur jeweiligen Beobachtungsdauer vor.

Tabelle 10: Allgemeine Angaben der Studie SUSTAIN

Studie SUSTAIN	Crizanlizumab 5,0 mg/kg	Placebo
ITT-Population <sup>1)</sup> , n	67	65
mITT-Population <sup>2)</sup> , n (%) <sup>3)</sup>	66 (98,5)	62 (95,4)
Sicherheitspopulation <sup>4)</sup> , n (%) <sup>3)</sup>	66 (98,5)	62 (95,4)
PP-Population <sup>5)</sup> , n (%) <sup>3)</sup>	40 (59,7)	41 (63,1)
Vorzeitiger Studienabbruch <sup>6)</sup> , n (%) <sup>3)</sup>	24 (35,8)	24 (36,9)
Aufgrund von UE	1 (1,5)	3 (4,6)
Entscheidung des Arztes	2 (3,0)	2 (3,1)
Lost to Follow-up	4 (6,0)	6 (9,2)
Tod	2 (3,0)	2 (3,1)
Fehlende Compliance	1 (1,5)	1 (1,5)
Protokollverletzung	0	0
Entzug der Einverständniserklärung	7 (10,4)	6 (9,2)
Fehlende Wirksamkeit	0	0
Durch den Sponsor abgebrochen	0	0
Sonstige Gründe	7(10,4)	4 (6,2)
Mediane Behandlungsdauer in Tagen (min; max) <sup>7)8)</sup>	364 (15; 385)	365 (15; 396)
Mediane Beobachtungsdauer	k. A.	k. A.
Mediane Anzahl erhaltener Dosen (min; max) <sup>7)</sup>	13 (1; 14)	13,5 (1; 14)
Personen mit ≥ 1 relevanten Protokollverletzung, n (%) <sup>3)</sup>	27 (40,3)	24 (36,9)
Nicht mindestens 12 von maximal 14 Dosen erhalten, n (%) <sup>9)</sup>	25 (37,3)	24 (36,9)
Keine Visite mindestens 14 Tage nach letzter Dosisgabe abgeschlossen, n (%) <sup>9)</sup>	10 (14,9)	12 (18,5)
Weniger als 2 oder mehr als 10 VOC vor Einschluss in die Studie, n (%) <sup>9)</sup>	3 (4,5)	2 (3,1)

- 1) Alle randomisierten Personen.
- 2) Allen randomisierten Personen, die mindestens eine Dosis Studienmedikation erhalten haben. Patientendaten wurden nach der Behandlungsgruppe ausgewertet, in die sie randomisiert wurden.
- 3) Prozentuale Angaben beziehen sich auf den Anteil an allen randomisierten Personen (ITT) in der jeweiligen Behandlungsgruppe.
- 4) Allen randomisierten Personen, die mindestens eine Dosis der Studienmedikation erhalten haben. Patientendaten wurden nach tatsächlich erhaltener Behandlung ausgewertet.
- 5) Alle randomisierten Personen, die keine der in der vorliegenden Tabelle angegebenen relevanten Protokollverletzungen aufwiesen. Patientendaten wurden nach der Behandlungsgruppe ausgewertet, in die sie randomisiert wurden.
- 6) Vorliegend ist der primäre Grund angegeben; es konnten mehrere Begründungen zutreffen.
- 7) Bezogen auf die Sicherheitspopulation der jeweiligen Behandlungsgruppe.
- 8) Die Behandlungsdauer war definiert als der letzte Tag der Studienmedikationsgabe - dem Datum der ersten Studiengabe + 15 Tagen. Behandlungsunterbrechungen und Compliance wurden nicht berücksichtigt.
- 9) Es waren Nennungen verschiedener Protokollverletzungen pro Person möglich.

Abkürzungen: ITT: Intention-to-Treat; k. A.: keine Angabe; mITT: modifizierte ITT-Population; PP: Per Protocol; UE: Unerwünschtes Ereignis; VOC: Vasookklusive Schmerzkrise.

Zu Baseline waren die Patientinnen und Patienten im Median zwischen 26 Jahren im Placebo-Arm und 29 Jahren im Crizanlizumab-Arm alt. Bei einer annähernd balancierten Geschlechterverteilung innerhalb des Crizanlizumab-Arms mit 52,2 % Frauen und 47,8 % Männern, waren mit 58,5 % vergleichsweise mehr Frauen im Placebo-Arm vertreten. In beiden Armen waren die Patientinnen und Patienten zu ca. 90 % schwarz oder afroamerikanischer Abstammung (siehe Tabelle 11). Während mit etwa 70 % überwiegend und nahezu gleichverteilt Personen mit HbSS-Genotyp eingeschlossen wurden, zeigten sich bei Personen mit HbSβ0-Thalassämie (4,5 % im Crizanlizumab-Arm und 10,8 % im Placebo-Arm) sowie HbSβ+-Thalassämie (10,4 % im Crizanlizumab-Arm und 1,5 % im Placebo-Arm) deskriptiv Unterschiede zwischen den Behandlungsgruppen. Gemäß Interactive Voice/Web Response System (IXRS) scheint eine ausgewogene Stratifikation (gemäß begleitender HU-Therapie und VOC während der vergangenen 12 Monate) vorzuliegen. Deskriptive Angaben gemäß eCRF stellen jeweils eine geringfügig höhere Personenanzahl mit weniger VOC in den vergangenen 12 Monaten bzw. mit keiner begleitenden HU-Therapie im Crizanlizumab-Arm gegenüber dem Placebo-Arm dar (siehe Tabelle 11). Eine intervallskalierte Deskription der VOC während der vergangenen 12 Monate (gemäß eCRF) wird in Tabelle 11 zur ITT-Population unter Ausschluss einer Person im Crizanlizumab-Arm, welche einen statistischen Ausreißer (37 VOC in 6 Monaten) darstellt, aufgezeigt [3] (Daten nicht im Rahmen des Dossiers vorgelegt). Diese Population bildet die Grundlage für die Sensitivitätsanalyse M6b. Jeweils 2 weitere Personen beider Behandlungsarme, welche Protokollverletzungen bezüglich der VOC während den vergangenen 12 Monaten aufwiesen (siehe Tabelle 10) wurden von dieser Auswertung nicht ausgeschlossen. Es konnten keine Informationen dazu identifiziert werden, inwiefern diese Personen im Rahmen der Stratifizierung nach VOC während der vergangenen 12 Monate (2–4 vs. 5–10) berücksichtigt wurden, da sowohl die Angaben dazu nach IXRS als auch gemäß eCRF eine sachgerechte Stratifizierung aller Personen (N = 132) innerhalb von 2 bis 10 VOC postulieren. In Verbindung mit der intervallskalierten Deskription der VOC während der vergangenen 12 Monate (gemäß eCRF) ist davon auszugehen, dass jeweils eine Person in beiden Armen bei Einschluss nur eine dieser historischen VOC aufwies und eine Person im Crizanlizumab-Arm 13 VOC während der vergangenen 12 Monate gegenüber einer Person im Placebo-Arm 24 dieser VOC berichtete. Angaben zu etwaigen statistischen Ausreißern dieser Personen während der Studie liegen nicht vor.

Die durchschnittliche Dauer seit Diagnosestellung der Patientinnen und Patienten zu Baseline ist nicht bekannt. Abgesehen von Angaben zu VOC in den vergangenen 12 Monaten sowie des Genotyps liegen keine spezifischen Einschätzungen von Krankheitsschwere oder Risikofaktoren vor.

Tabelle 11: Charakterisierung der Studienpopulation der Studie SUSTAIN – ITT-Population

<b>Studie SUSTAIN</b>	<b>Crizanlizumab 5,0 mg/kg N = 67</b>	<b>Placebo N = 65</b>
<i>Alter (Jahre)</i> MW (SD) Median (min; max)	30,9 (10,89) 29,0 (16; 63)	29,3 (10,36) 26,0 (16; 56)
<i>Geschlecht, n (%)<sup>1)</sup></i> männlich weiblich	32 (47,8) 35 (52,2)	27 (41,5) 38 (58,5)
<i>Abstammung (genetisch), n (%)<sup>2)</sup></i> indigen amerikanisch asiatisch schwarz oder afroamerikanisch indigen hawaiianisch oder von anderen pazifischen Inseln weiß andere	0 0 60 (89,6) 0 4 (6,0) 3 (4,5)	0 0 60 (92,3) 0 3 (4,6) 2 (3,1)
<i>BMI in kg/m<sup>2</sup>, n (%)<sup>1)</sup></i> MW (SD) Median (min; max)	64 (95,5) 24,32 (5,45) 23 (14,5; 36)	62 (95,4) 23,75 (4,8) 23,2 (17; 37,9)
<i>Sichelzellerkrankungstyp, n (%)<sup>1)</sup></i> HbSS HbSC HbS $\beta$ 0-Thalassämie HbS $\beta$ +Thalassämie Anderer	47 (70,1) 9 (13,4) 3 (4,5) 7 (10,4) 1 (1,5)	47 (72,3) 8 (12,3) 7 (10,8) 1 (1,5) 2 (3,1)
<b>Stratifikationsfaktoren (separat)</b> <i>Begleitende HU-Therapie (gemäß IXRS), n (%)<sup>1)</sup></i> Ja Nein <i>VOC während den vergangenen 12 Monaten (gemäß IXRS), n (%)<sup>1)</sup></i> 2–4 5–10	42 (62,7) 25 (37,3) 42 (62,7) 25 (37,3)	40 (61,5) 25 (38,5) 41 (63,1) 24 (36,9)
<b>Stratifikationsfaktoren</b> <i>Begleitende HU-Therapie und Anzahl VOC während der vergangenen 12 Monate (gemäß IXRS), n (%)<sup>1)</sup></i> Ja und 2–4 Ja und 5–10 Nein und 2–4 Nein und 5–10	25 (37,3) 17 (25,4) 17 (25,4) 8 (11,9)	24 (36,9) 16 (24,6) 17 (26,2) 8 (12,3)
<i>Begleitende HU-Therapie (gemäß eCRF), n (%)<sup>1)</sup></i> Ja Nein <i>VOC während den vergangenen 12 Monaten (gemäß eCRF), n (%)<sup>1)</sup></i> 2–4 5–10	43 (64,2) 24 (35,8) 45 (67,2) 22 (32,8)	40 (61,5) 25 (38,5) 39 (60,0) 26 (40,0)
<i>Begleitende HU-Therapie und Anzahl VOC während der vergangenen 12 Monate (gemäß eCRF), n (%)<sup>1)</sup></i> Ja und 2–4 Ja und 5–10 Nein und 2–4 Nein und 5–10	28 (41,8) 15 (22,4) 17 (25,4) 7 (10,4)	23 (35,4) 17 (26,2) 16 (24,6) 9 (13,8)

<b>Studie SUSTAIN</b>	<b>Crizanlizumab 5,0 mg/kg N = 67</b>	<b>Placebo N = 65</b>
VOC während der vergangenen 12 Monate (gemäß eCRF), n (%) <sup>1)3)</sup> MW (SD) Median (min; max)	66 <sup>4)</sup> (98,5) 3,95 (2,46) 3 (1; 13)	65 (100) 4,43 (3,3) 4 (1; 24)
BPI-LF, n (%) <sup>1)</sup> Summenscore „Schmerzintensität“ <sup>5)</sup> , MW (SD) Summenscore „Beeinträchtigung durch Schmerz“ <sup>6)</sup> , MW (SD)	48 (71,6) 4,36 (2,12) 4,64 (2,57)	55 (84,6) 4,13 (2,01) 5 (2,95)
Dauer seit Diagnosestellung bis Randomisierung in Jahren MW (SD) Median (min; max)	k. A. k. A.	k. A. k. A.

<sup>1)</sup> Prozentuale Angaben beziehen sich auf den Anteil an allen randomisierten Personen (ITT) in der jeweiligen Behandlungsgruppe.

<sup>2)</sup> Prozentuale Angaben beziehen sich auf den Anteil an den vorliegenden Angaben zur Abstammung. Mehrfachangaben zur genetischen Abstammung waren möglich.

<sup>3)</sup> Diese Daten wurden nicht mit dem Dossier vorgelegt und entstammen dem EPAR [3].

<sup>4)</sup> Hierbei handelt es sich um die ITT-Population unter Ausschluss einer Person von der Auswertung. Die Person erfüllte die Einschlusskriterien nicht, da diese bereits im Jahr vor der Randomisierung 30 VOC erlitten hatte. Während der Studienzeit wies die Person innerhalb von 6 Monaten 37 VOC auf, was einen statistischen Ausreißer darstellte [3].

<sup>5)</sup> Der Score zur Schmerzintensität setzt sich zusammen aus vier Items, die jeweils auf einer elfstufigen Skala von 0 (kein Schmerz) bis 10 (stärkste vorstellbare Schmerzen) bewertet werden. Der Summenscore umfasst Werte von 0–40.

<sup>6)</sup> Der Score zur Beeinträchtigung durch Schmerz setzt sich zusammen aus sieben Items, die jeweils auf einer elfstufigen Skala von 0 (keine Beeinträchtigung) bis 10 (vollständige Beeinträchtigung) bewertet werden. Der Summenscore umfasst Werte von 0–70.

Abkürzungen: BPI-LF: Brief Pain Inventory – Long Form; eCRF: electronic Case Report Form; EPAR: European Public Assessment Report; HU: Hydroxyurea/Hydroxycarbamid; ITT: Intention-to-Treat; IXRS: Interactive Voice/Web Response System; k. A.: keine Angabe; MW: Mittelwert; SD: Standardabweichung; VOC: Vasookklusive Schmerzkrise.

Aufgrund fehlender spezifischer Einschätzungen der Krankheitsschwere, sind in Tabelle 12 die Ereignisse der medizinischen Vorgeschichte und Komorbiditäten der Patientinnen und Patienten zu Baseline dargestellt. Neben Systemorganklassen nach MedDRA (Version 16) mit Ereignissen von  $\geq 10\%$  wurden Preferred Terms nach MedDRA aufgeführt, die aufgrund ihres direkten Zusammenhangs zu den erhobenen VOC-spezifischen Endpunkten ausgewählt wurden. Bei der Interpretation ist zu beachten, dass keine Informationen dazu vorliegen, wie umfassend die medizinische Vorgeschichte der Patientinnen und Patienten erfasst und überprüft wurde. Deskriptiv zeigen sich Auffälligkeiten mit einem Unterschied von  $\geq 10$  Prozentpunkten zwischen den beiden Behandlungsarmen hinsichtlich der Systemorganklassen „Herzerkrankungen“ ( $n_C = 10$ ,  $n_P = 22$ ), „Infektionen und parasitäre Erkrankungen“ ( $n_C = 25$ ,  $n_P = 31$ ), „Stoffwechsel- und Ernährungsstörungen“ ( $n_C = 17$ ,  $n_P = 23$ ), „Erkrankungen der Atemwege, des Brustraums und Mediastinums“ ( $n_C = 25$ ,  $n_P = 35$ ). Zu diesen Systemorganklassen wurden deskriptiv häufiger Patientinnen und Patienten mit mindestens einem Ereignis oder einer Kondition der medizinischen Vorgeschichte im Placebo-Arm berichtet.

Alle Patientinnen und Patienten wiesen mindestens ein Ereignis im Rahmen der gesamten medizinischen Vorgeschichte auf. Auch Schmerzkrise wurden bezüglich der medizinischen Vorgeschichte von allen Patientinnen und Patienten berichtet. Erklärungen für die Zählweise der Preferred Terms „Sichelzellanämie“, „Sichelzellanämie mit Krise“ und „Beta-Thalassämie“ in Relation zu den in Tabelle 12 beschriebenen deskriptiven Häufigkeiten von Sichelzellerkrankungstypen liegen nicht vor. Zu den Darstellungen der ausgewählten Preferred Terms „Sichelzellanämie“, „Sichelzellanämie mit Krise“, „Beta-Thalassämie“, „Priapismus“ und „Splenektomie“ scheinen im Rahmen der medizinischen Vorgeschichte und Komorbiditäten deskriptiv weitestgehend vergleichbare Fallzahlen vorzuliegen. Zum Preferred Term „Akutes Thoraxsyndrom“

berichten deskriptiv häufiger Personen im Placebo-Arm (n = 14) von mindestens einem Ereignis in der medizinischen Vorgeschichte als Personen im Crizanlizumab-Arm (n = 8).

*Tabelle 12: Medizinische Vorgeschichte und Komorbiditäten mit  $\geq 10\%$  Ereignisse nach MedDRA-Systemorganklassen und ausgewählten Preferred Terms<sup>1)</sup> zu Baseline der Studie SUSTAIN – ITT-Population*

<b>MedDRA-Systemorganklasse<sup>2)</sup></b> Preferred Term	<b>Crizanlizumab 5,0 mg/kg</b> <b>N = 67</b> <i>n (%)<sup>3/4)</sup></i>	<b>Placebo</b> <b>N = 65</b> <i>n (%)<sup>3/4)</sup></i>
Patientinnen und Patienten mit $\geq 1$ Ereignis	167(100)	65 (100)
<b>Erkrankungen des Blutes und des Lymphsystems</b>	<b>12 (17,9)</b>	<b>12 (18,5)</b>
<b>Herzerkrankungen</b>	<b>10 (14,9)</b>	<b>22 (33,8)</b>
<b>Kongenitale, familiäre und genetische Erkrankungen</b>	<b>10 (14,9)</b>	<b>12 (18,5)</b>
Sichelzellanämie	1 (1,5)	3 (4,6)
Sichelzellanämie mit Krise	5 (7,5)	8 (12,3)
Beta-Thalassämie	1 (1,5)	0
<b>Augenerkrankungen</b>	<b>12 (17,9)</b>	<b>12 (18,5)</b>
<b>Erkrankungen des Gastrointestinaltrakts</b>	<b>22 (32,8)</b>	<b>18 (27,7)</b>
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	<b>9 (13,4)</b>	<b>11 (16,9)</b>
<b>Leber- und Gallenerkrankungen</b>	<b>13 (19,4)</b>	<b>16 (24,6)</b>
<b>Erkrankungen des Immunsystems</b>	<b>12 (17,9)</b>	<b>9 (13,8)</b>
<b>Infektionen und parasitäre Erkrankungen</b>	<b>25 (37,3)</b>	<b>31 (47,7)</b>
<b>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</b>	<b>7 (10,4)</b>	<b>4 (6,2)</b>
<b>Untersuchungen</b>	<b>14 (20,9)</b>	<b>17 (26,2)</b>
<b>Stoffwechsel- und Ernährungsstörungen</b>	<b>17 (25,4)</b>	<b>23 (35,4)</b>
<b>Skelettmuskulatur- und Bindegewebs- und Knochenerkrankungen</b>	<b>29 (43,3)</b>	<b>30 (46,2)</b>
<b>Erkrankungen des Nervensystems</b>	<b>20 (29,9)</b>	<b>21 (32,3)</b>
<b>Psychiatrische Erkrankungen</b>	<b>16 (23,9)</b>	<b>17 (26,2)</b>
<b>Erkrankungen der Nieren und Harnwege</b>	<b>10 (14,9)</b>	<b>14 (21,5)</b>
<b>Erkrankungen der Geschlechtsorgane und der Brustdrüse</b>	<b>10 (14,9)</b>	<b>8 (12,3)</b>
Priapismus	3 (4,5)	2 (3,1)
<b>Erkrankungen der Atemwege, des Brustraums und Mediastinums</b>	<b>25 (37,3)</b>	<b>35 (53,8)</b>
Akutes Thoraxsyndrom	8 (11,9)	14 (21,5)
<b>Erkrankungen der Haut und des Unterhautzellgewebes</b>	<b>10 (14,9)</b>	<b>6 (9,2)</b>
<b>Chirurgische und medizinische Eingriffe</b>	<b>47 (70,1)</b>	<b>42 (64,6)</b>
Splenektomie	10 (14,9)	7 (10,8)
<b>Gefäßerkrankungen</b>	<b>17 (25,4)</b>	<b>12 (18,5)</b>

<sup>1)</sup> Die dargestellten Preferred Terms wurden aufgrund ihres direkten Zusammenhangs zu den erhobenen VOC-spezifischen Endpunkten ausgewählt.

<sup>2)</sup> Systemorganklassen und Preferred Terms gemäß MedDRA (Version 16).

<sup>3)</sup> Personen mit mehrfachen Ereignissen in der gleichen Kategorie werden nur einmal in dieser gezählt.

<sup>4)</sup> Prozentuale Angaben beziehen sich auf den Anteil an allen randomisierten Personen (ITT) in der jeweiligen Behandlungsgruppe.

Abkürzungen: ITT: Intention-to-Treat; MedDRA: Medical Dictionary for Regulatory Activities; VOC: Vasookklusive Schmerzkrise.

Es liegen im Rahmen des Dossiers voneinander abweichende Informationen zur Begleitmedikation vor. Gemäß Studienbericht wird davon ausgegangen, dass 66 Personen im Crizanlizumab-Arm und 62 im Placebo-Arm für eine unbestimmte Dauer mindestens eine medikamentöse Begleittherapie während der Studiendauer erhalten haben.

Zu Begleitmedikationen, welche in mindestens einem Studienarm von  $\geq 20$  % der Patientinnen und Patienten angewendet wurden, ist zu berücksichtigen, dass Heparin für eine unbestimmte Zeit von 8 Personen (11,9 %) im Crizanlizumab-Arm und von 16 Personen (24,6 %) im Placebo-Arm trotz des Ausschlusskriteriums einer chronischen Therapie mit Antikoagulantien (außer Aspirin) angewendet wurde. Bei der Interpretation der Ergebnisse zur Begleitmedikation im Studienbericht ist zu berücksichtigen, dass diese durch eine äußerst limitierte Aufbereitung erschwert scheint, da zum einen keine Darstellung zusammengefasst nach allen jeweiligen Arzneimittelgruppen (bspw. Opioide) erfolgte und zum anderen aus den Rohdaten hervorgeht, dass spezifische Arzneimittelangaben nicht zusammengefasst wurden [26]. So wurden bspw. „Morfina“, „Morfine“, „Morphine Sulfate“, „Morphine Sulphate“ und das in Tabelle 13 dargestellte „Morphin“ jeweils einzeln erfasst, Diskrepanzen zwischen der vorliegenden Tabelle und Darstellungen zur Begleitmedikation in Modul 4 konnten dadurch nicht aufgeklärt werden (vgl. [10]).

Tabelle 13: Begleitmedikation in der Studie SUSTAIN – ITT-Population

Medikamentöse Begleitmedikation <sup>1)</sup> bei $\geq 20$ Personen gemäß WHODD <sup>2)</sup>	Crizanlizumab 5,0 mg/kg N = 67	Placebo N = 65
Personen mit $\geq 1$ Begleitmedikation, n (%)	66 (98,5)	62 (95,4)
Acetaminophen	17 (25,4)	16 (24,6)
Benadryl	18 (26,9)	20 (30,8)
Dilaudid	27 (40,3)	29 (44,6)
Diphenhydramin	11 (16,4)	17 (26,2)
Folic acid	50 (74,6)	45 (69,2)
Heparin	8 (11,9)	16 (24,6)
Hydromorphone	13 (19,4)	20 (30,8)
Hydroxyurea (HU) <sup>3)</sup>	33 (49,3)	36 (55,4)
Ibuprofen	25 (37,3)	24 (36,9)
Ketorolac	12 (17,9)	14 (21,5)
Miralax	6 (9)	15 (23,1)
Morphin	30 (44,8)	31 (47,7)
Odansentron	10 (14,9)	16 (24,6)
Perocet	12 (17,9)	17 (26,2)
Phenergan	10 (14,9)	15 (23,1)
Potassium chlorid	5 (7,5)	13 (20)
Sodium chlorid	12 (17,9)	19 (29,2)
Toradol	15 (22,4)	21 (32,3)
Zofran	18 (26,9)	22 (33,8)

<sup>1)</sup> Alle Medikamente, die seit Beginn der Studienbehandlung bis zur letzten Follow-up-Visite angewandt wurden, einschließlich denjenigen, die vor Beginn der Studienbehandlung begonnen und darüber hinaus fortgesetzt wurden.

<sup>2)</sup> Kodierung gemäß Version 01DEC2013E.

<sup>3)</sup> Eine Begleitmedikation mit „Hydrea“ wurde zusätzlich für 8 Personen (11,9 %) im Crizanlizumab- und 4 (6,2 %) im Placebo-Arm berichtet und eine Begleitmedikation mit „Hidoxiurea“ wurde zusätzlich für keine Person im Crizanlizumab- und 1 Person (1,5 %) im Placebo-Arm vermerkt.

Abkürzungen: ITT: Intention-to-Treat; WHODD: World Health Organization Drug Dictionary.

## 3.2 Mortalität

Todesfälle wurden in der Studie SUSTAIN im Rahmen der Sicherheitserhebung erfasst. Es wurden pro Arm jeweils 2 Todesfälle deskriptiv berichtet, welches 3 % der Sicherheitspopulation (N = 66) des Crizanlizumab-Arm bzw. 3,2 % der Sicherheitspopulation (N = 62) des Placebo-Arms entspricht. Im Crizanlizumab-Arm stand ein Todesfall im Zusammenhang mit einer Sichelzellanämie mit Krise und der andere Todesfall trat mit einer Endokartitis und einer Sepsis auf. Die beiden Todesfälle im Placebo-Arm traten bei der einen Person in Verbindung mit einer rechtsventrikulären Herzinsuffizienz und bei der anderen Person mit einer Sichelzellanämie mit Krise, Beinvenenthrombose, Sepsis, Koma und ischämischem Schlaganfall auf.

## 3.3 Morbidität

### Vasookklusive Schmerzkrisen

Im Rahmen der Studie SUSTAIN wurde die Veränderung des Auftretens vasookklusiver Schmerzkrisen untersucht. Dabei wurden jegliche vasookklusive Schmerzkrisenereignisse u. a. anhand der Endpunkte „Jährliche Rate VOC“, „Zeit bis zur ersten VOC“ und „Zeit bis zu zweiten VOC“ untersucht. Von den fünf Subklassifikationen sind zudem „Jährliche Rate unkomplizierte VOC“ und „Jährliche Rate akute Thoraxsyndrome“ als eigenständige Endpunkte anhand jährlicher Raten ausgewertet worden, während Ergebnisse zu den Subklassifikationen „Lebersequestration“, „Milzsequestration“ und „Priapismus“ ausschließlich deskriptiv dargestellt wurden. Für eine deskriptive Übersicht zu allen Subklassifikationen gemäß ärztlichem Prüfpersonal und CRC bezüglich der Sicherheitspopulation siehe Tabelle 14; diese war nicht Grundlage der Wirksamkeitsanalysen. Bei der Interpretation ist weiterhin zu beachten, dass die deskriptive Darstellung nicht die Länge potentiell variabler Beobachtungsdauern berücksichtigt. Unabhängig von der Einschätzungsvariante (gemäß ärztlichem Prüfpersonal oder CRC) wurden keinerlei Leber- und Milzsequestrationen sowie gemäß ärztlichem Prüfpersonal nur jeweils ein Fall von Priapismus in beiden Armen beobachtet. Es ist zu berücksichtigen, dass 14,9 % der Personen im Crizanlizumab-Arm und 10,8 % der Personen im Placebo-Arm im Rahmen der Erhebung ihrer jeweiligen medizinischen Vorgeschichte von einer Splenektomie berichteten, sodass die Auftretenswahrscheinlichkeit von Milzsequestration durch die Populationscharakteristika verringert war.

Beim Vergleich der Einschätzung zu VOC gemäß ärztlichem Prüfpersonal gegenüber der adjudizierten VOC-Ereignisanzahl gemäß CRC zeigen sich erhebliche Unterschiede zwischen den beiden Einschätzungsvarianten, die in beiden Armen, jedoch v. a. im Crizanlizumab-Arm, zu beobachten sind. Der Vergleich basiert auf Angaben zur Sicherheitspopulation, da zur ITT-Population Informationen zu beiden Einschätzungsvarianten nicht vorliegen. Außerdem konnten im Dossier keine Behandlungsarm-spezifischen Angaben zur Konkordanz identifiziert werden, sodass diesbezüglich Unklarheiten bestehen. Angaben im EPAR weisen auf eine 64,6%ige Konkordanz im aktiven Behandlungsarm gegenüber einer Konkordanz von 81,6 % im Kontrollarm hin [3], welches die vorgelegten deskriptiven Angaben zu Diskrepanzen zwischen den VOC-Ereignissen gemäß ärztlichem Prüfpersonal ( $n_C = 225$ ;  $n_P = 244$ ) gegenüber Einschätzungen des CRC ( $n_C = 148$ ;  $n_P = 202$ ) stützt (siehe Tabelle 14). Die Konkordanz im Crizanlizumab-Arm konnte gemäß EPAR durch den analysespezifischen Ausschluss einer Person, die 37 VOC in 6 Monaten aufwies und deren VOC aufgrund zeitlicher Nähe durch das CRC gruppiert wurden, auf 77,6 % gesteigert werden [3]. Der Ausschluss einer Person aus den Analysen gemäß PhS-M6b wurde mit

einer hohen Anzahl an VOC im Jahr vor Studienbeginn (was eine Verletzung der Einschlusskriterien darstellt) sowie während 6 Monate der Studienzeit (was eher auf chronische Schmerzen als auf akute VOC hindeuten könnte) begründet. Auswertungen zu den Adjudizierungsgründen liegen zur Nutzenbewertung nicht vor, sodass weitere Unklarheiten bestehen.

Tabelle 14: Nach Behandlungsbeginn aufgetretene VOC – Sicherheitspopulation

Studie SUSTAIN	Crizanlizumab 5,0 mg/kg N = 66		Placebo N = 62	
	Personen n (%) <sup>1)2)</sup>	Ereignisse n (%) <sup>3)4)</sup>	Personen n (%) <sup>1)2)</sup>	Ereignisse n (%) <sup>3)4)</sup>
<b>Nach Behandlungsbeginn aufgetretene VOC gemäß ärztlichem Prüfpersonal</b>				
VOC	50 (75,8)	225 (100)	54 (87,1)	244 (100)
Unkomplizierte VOC	46 (69,7)	206 (91,6)	52 (83,9)	230 (94,3)
Akutes Thoraxsyndrom	15 (22,7)	18 (8)	10 (16,1)	13 (5,3)
Lebersequestration	0	0	0	0
Milzsequestration	0	0	0	0
Priapismus	1 (1,5)	1 (0,4)	1 (1,6)	1 (0,4)
<b>Nach Behandlungsbeginn aufgetretene VOC gemäß CRC</b>				
VOC	48 (72,7)	148 (100)	54 (87,1)	202 (100)
Unkomplizierte VOC	45 (68,2)	129 (87,2)	50 (80,6)	184 (91,1)
Akutes Thoraxsyndrom	14 (21,2)	18 (12,2)	13 (21)	15 (7,4)
Lebersequestration	0	0	0	0
Milzsequestration	0	0	0	0
Priapismus	0	0	1 (1,6)	1 (0,5)
Tod <sup>5)</sup>	1 (1,5)	1 (0,6)	2 (3,2)	2 (1)

<sup>1)</sup> Personen mit mehrfachen Ereignissen in der gleichen Kategorie werden nur einmal in dieser gezählt. Personen mit Ereignissen in mehr als einer Kategorie werden einmal je Kategorie gezählt.

<sup>2)</sup> Prozentuale Angaben beziehen sich auf den Anteil der betroffenen Personen an der Sicherheitspopulation.

<sup>3)</sup> Mehre Ereignisse einer Person in einer Kategorie werden mehrfach in dieser gezählt.

<sup>4)</sup> Prozentuale Angaben beziehen sich auf den Anteil der Ereignisse an allen VOC.

<sup>5)</sup> Die VOC-Subklassifikation „Tod“ wurde aus dem zweiten Protokoll als VOC-definierendes Ereignis entfernt, jedoch entschied das CRC in 3 Fällen, dass jeweils die Kriterien einer VOC vorgelegen hätten und daher die Subklassifikation „Tod“ vergeben werden sollte.

Abkürzungen: CRC: Crisis Review Committee; VOC: Vasookklusive Schmerzkrise.

Die primäre Analyse des Endpunkts „Jährliche Rate VOC“ basierte auf VOC gemäß CRC und basierend auf der ITT-Population mit einem regelhaften Erfassungszeitraum von Randomisierung bis zum Zeitpunkt der letzten Dosis + 14 Tage (siehe Kapitel 2.3.2 – Vasookklusive Schmerzkrise für potentielle Abweichungen). In Tabelle 15 sind die Ergebnisse der primären Analyse und die darin eingeflossenen VOC (gemäß ITT) deskriptiv dargestellt. Bei jeglicher Ergebnisinterpretation im Rahmen der Analyse von VOC und entsprechenden Subklassifikationen ist der hohe Anteil an Studienabbrüchen von über einem Drittel sowie das Fehlen der jeweiligen Beobachtungsdauern der beiden Studienarme zu berücksichtigen. Fehlende Beobachtungszeiten wurden unter der Annahme einer gleichbleibenden Ereignisrate durch eine Annualisierung der Zählzeiten auf ein Jahr extrapoliert. Zur Interpretation der Ergebnisse der primären Analyse (anhand Wilcoxon-Rangsummen-Test und Hodges-Lehmann-Schätzer) muss weiterhin berücksichtigt werden, dass die statistischen Annahmen über das Vorliegen über ähnliche Verteilungsformen des Endpunkts „Jährliche Rate VOC“ beider Studienarme, die voneinander verschoben sein müssen, nicht erfüllt scheinen [3]. Unter diesen Voraussetzungen erscheint eine eindeutige Interpretation der Ergebnisse hinsichtlich der Signifikanz des Unterschieds der Mediane (Wilcoxon-Rangsummen-Test) und eine verlässliche Schätzung des Medians der Differenzen (Hodges-Lehmann-Schätzer) nicht möglich.

**Tabelle 15: Primäre Analyse des Endpunkts „Jährliche Rate VOC“ und die berücksichtigte Anzahl und Subklassifikation der VOC gemäß CRC – ITT-Population**

	<b>Crizanlizumab 5,0 mg/kg N = 67</b>	<b>Placebo N = 65</b>	<b>HL-Schätzer des Medians der Differenzen [95%-KI], p-Wert</b>
<b>In der primären Analyse berücksichtigte VOC gemäß CRC</b>			
VOC, n <sup>1)</sup>	121	179	
Unkomplizierte VOC, n (%) <sup>1)2)</sup>	109 (90,1)	166 (92,7)	
Akutes Thoraxsyndrom, n (%) <sup>1)2)</sup>	12 (9,9)	12 (6,7)	
Lebersequestration, n (%) <sup>1)2)</sup>	0 (0)	0 (0)	
Milzsequestration, n (%) <sup>1)2)</sup>	0 (0)	0 (0)	
Priapismus, n (%) <sup>1)2)</sup>	0 (0)	1 (0,6)	
<b>Primäre Analyse des Endpunkts „Jährliche Rate VOC“<sup>3)</sup></b>			
Jährliche Rate VOC <sup>4)</sup>			
Median (min; max)	1,63 (0; 24,3)	2,98 (0; 24,3)	
MW (SD)	2,89 (4,2)	4,43 (4,86)	
HL-Schätzer Median	2 <sup>5)</sup>	3,49 <sup>5)</sup>	-1,01 [-2; 0] <sup>6)</sup> ; 0,01 <sup>7)</sup>

<sup>1)</sup> Hierbei handelt es sich um die Anzahl der Ereignisse (und nicht der Personen).

<sup>2)</sup> Prozentuale Angaben beziehen sich auf den Anteil der subklassifizierten Ereignisse an allen VOC.

<sup>3)</sup> In dieser Analyse wurden keine Imputationen vorgenommen. Fehlende Daten wurden durch Extrapolation der vorliegenden Daten nicht berücksichtigt. Vorzeitige Studienabbrüche betrafen jeweils 24 Personen in beiden Armen. 27 Personen im Crizanlizumab-Arm und 24 Personen im Placebo-Arm wiesen  $\geq 1$  der berichteten Protokollverletzungen auf.

<sup>4)</sup> Definiert als die Anzahl der VOC einer Person während der Behandlungsphase  $\times 365 / (\text{Datum des Behandlungsendes} - \text{Datum der Randomisierung} + 1)$ , wobei das Behandlungsende definiert war als der Zeitpunkt der letzten Dosis + 14 Tage. Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt).

<sup>5)</sup> Eine non-parametrische Schätzung des Medians erfolgte mithilfe der Hodges-Lehmann-Methode.

<sup>6)</sup> Eine non-parametrische Schätzung des Medians der Differenzen erfolgte mithilfe der Hodges-Lehmann-Methode.

<sup>7)</sup> Berechnung des p-Werts mithilfe eines Wilcoxon-Rangsummen-Tests stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).

Abkürzungen: CRC: Critical Review Committee; HL: Hodges-Lehmann; ITT: Intention-to-Treat; MW: Mittelwert; SD: Standardabweichung; VOC: Vasookklusive Schmerzkrise.

Die zur Nutzenbewertung herangezogenen Sensitivitätsanalysen (siehe Tabelle 16) basieren auf negativen Binomial-Regressionen und zeigen Effektschätzer in Richtung einer Wirksamkeit von Crizanlizumab gegenüber Placebo, dessen jeweilige 95%-KI jedoch in den Imputationsmodellen PhS-M6a und PhS-M6b die 1 umschließen. Bei der Interpretation ist zu berücksichtigen, dass widersprüchliche Angaben im Rahmen der Ergebnispräsentation von PS-1 und PhS-2 hinsichtlich der Durchführung einer Annualisierung der VOC vorliegen, welche in der Diskussion näher erläutert werden [10,26].

Die Interaktionsterme für die a priori definierten Subgruppenanalysen waren nicht signifikant. Es erfolgt daher keine Darstellung der Subgruppenanalysen.

**Tabelle 16: Sensitivitätsanalysen zu VOC basierend auf Negativ-Binomial-Regression – ITT-Population**

	<b>Crizanlizumab 5,0 mg/kg N = 67</b>	<b>Placebo N = 65</b>	<b>Rate Ratio [95%-KI]; p-Wert</b>
<b>Sensitivitätsanalysen ohne Berücksichtigung fehlender Werte<sup>1)</sup></b>			
PS-1: VOC gemäß CRC <sup>2)</sup> Geschätzte JR VOC [95%-KI] <sup>3)4)</sup>	2,43 [1,9; 3,11]	3,75 [2,99; 4,71]	0,65 [0,47; 0,9]; 0,008
PhS-2: VOC gemäß ärztlichem Prüfpersonal <sup>4)</sup> Geschätzte JR VOC [95%-KI] <sup>3)4)</sup>	N = 66 <sup>9)</sup> 3,14 [2,42; 4,06]	4,93 [3,84; 6,34]	0,64 [0,45; 0,9]; 0,011
<b>Sensitivitätsanalysen mit Imputationsmodellen</b>			
PhS-M5: VOC gemäß CRC <sup>4)5)6)</sup> Median (min; max) MW (SD)	k. A. k. A.	k. A. k. A.	0,72 [0,54; 0,95]; 0,022
PhS-M6a: VOC gemäß CRC <sup>4)5)7)</sup> Median (min; max) MW (SD)	k. A. k. A.	k. A. k. A.	0,74 [0,54; 1,03]; k. A.
PhS-M6b: VOC gemäß ärztlichem Prüfpersonal <sup>4)5)8)9)</sup> Median (min; max) MW (SD) <sup>10)</sup>	N = 66 <sup>9)</sup> k. A. 3,62 (4,12)	k. A. k. A. 4,95 (5,29)	0,74 [0,52; 1,06]; k. A.

<sup>1)</sup> In diesen Analysen wurden keine Imputationen vorgenommen. Fehlende Beobachtungszeiten wurden nicht berücksichtigt. Vorzeitige Studienabbrüche betrafen jeweils 24 Personen in beiden Armen. 27 Personen im Crizanlizumab-Arm und 24 im Placebo-Arm wiesen  $\geq 1$  der berichteten Protokollverletzungen auf.

<sup>2)</sup> Hierbei handelt es sich um eine Präspezifizierte Sensitivitätsanalyse (PS).

<sup>3)</sup> Der Bezugszeitraum beginnt zum Datum der Randomisierung und endet zum Behandlungsende, wobei das Behandlungsende definiert war als der Zeitpunkt der letzten Dosis + 14 Tage. Inwiefern eine Annualisierung der VOC stattgefunden hat, ist aufgrund widersprüchlicher Ergebnisdarstellungen unklar. Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt).

<sup>4)</sup> Raten und Ratenverhältnisse einschließlich 95%-KI sowie p-Wert geschätzt mithilfe eines Negativ-Binomial-Regressionsmodells (Log-Link-Funktion) mit dem Logarithmus der Studiendauer als Offset-Variable und der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10), dem Einsatz einer begleitenden HU-Therapie (ja vs. nein) sowie der Behandlung als unabhängige Variablen.

<sup>5)</sup> Hierbei handelt es sich um eine Post-hoc-Sensitivitätsanalyse (PhS).

<sup>6)</sup> Eine Imputation fehlender Werte von Personen mit Therapieabbruch innerhalb der ersten 6 Monate erfolgte durch die Anzahl der VOC vor der Randomisierung.

<sup>7)</sup> Eine Imputation der Anzahl der VOC erfolgte bei vorzeitigem Behandlungsabbruch für Personen im Crizanlizumab-Arm über eine Jump-to-Reference-Methode basierend auf VOC von Personen aus dem Placebo-Arm. Für Personen im Placebo-Arm erfolgte die Ersetzung unter der Missing-at-Random-Annahme basierend auf Daten von Personen desselben Studienarmes vor Behandlungsabbruch.

<sup>8)</sup> Das Vorgehen entspricht methodisch PhS-M6a (siehe Fußnote 7), erfolgt aber basierend auf Daten zu VOC gemäß ärztlichem Prüfpersonal und unter Ausschluss einer Person im Crizanlizumab-Arm.

<sup>9)</sup> Bei der von der Analyse ausgeschlossenen Person handelt es sich um einen statistischen Ausreißer mit 37 VOC in 6 Monaten, die zudem die Einschlusskriterien nicht erfüllte, da sie bereits im Jahr vor der Randomisierung 30 Schmerzkrisen erlitten hatte.

<sup>10)</sup> MW und SD inklusive aller Imputationen.

Abkürzungen: CRC: Critical Review Committee; HU: Hydroxyurea/Hydroxycarbamid; ITT: Intention-to-Treat; JR: Jährliche Rate; k. A.: keine Angabe; KI: Konfidenzintervall; M: Imputationsmethode; MW: Mittelwert; PhS: Post-hoc-Sensitivitätsanalyse; PS: Präspezifizierte Sensitivitätsanalyse; SD: Standardabweichung; VOC: Vasookklusive Schmerzkrise.

„Zeit bis zur ersten VOC“ und „Zeit bis zur zweiten VOC“

Die mediane Zeit bis zur ersten VOC zeigte in der primären Analyse einen Vorteil für den Crizanlizumab-Arm gegenüber dem Placebo-Arm. Bei Interpretation der Ergebnisse ist zu beachten, dass Time-to-Event-Analysen auf der statistischen Annahme einer nicht-informativen Zensierung basieren, die im vorliegenden Fall durch die hohe Rate von Studienabbrüchen sowie deren Hintergrund (bspw. „Entzug der Einverständniserklärung“, „Sonstige Gründe“; siehe Tabelle 10) in beiden Behandlungsarmen nicht gesichert erscheint. Angaben zu Beobachtungszeiten sowie spezifischen Hintergründen aller Zensierungen wurden nicht aufgeführt. Zensierungen konnten teilweise der Post-hoc-Auswertung „Anteil der Personen ohne VOC während der Behandlungsphase“ entnommen werden.

Der Unterschied zugunsten von Crizanlizumab wird durch Ergebnisse der Sensitivitätsanalyse mit Imputationsmodell PhS-M6b ([3]; Daten nicht im Rahmen des Dossiers vorgelegt), in Form einer Differenz zur ersten VOC von 2,63 Monaten, gestützt. Bezüglich des Endpunkts „Zeit bis zur zweiten VOC“, welcher unter Berücksichtigung der fehlenden Angaben und einer potentiellen informativen Zensierung zu interpretieren ist, werden 10,32 Monate im Crizanlizumab-Aarm gegenüber 5,09 Monaten im Placebo-Arm berichtet. Eine Post-hoc-Sensitivitätsanalyse analog zum Endpunkt „Zeit bis zur ersten VOC“ wird nicht berichtet.

Tabelle 17: Analysen zu den Endpunkten „Zeit bis zur ersten VOC“ und „Zeit bis zur zweiten VOC“ – ITT-Population

Studie SUSTAIN	Crizanlizumab 5,0 mg/kg N = 67	Placebo N = 65	HR [95%-KI]; p-Wert
<b>Zeit bis zur ersten VOC<sup>1)</sup></b>			
Primäre Analyse – VOC gemäß CRC			
Personen mit Ereignis, n (%)	43 (64,2)	54 (83,1)	0,5 [0,33; 0,74] <sup>3)</sup> ; 0,001 <sup>4)</sup>
Mediane Zeit bis zum Ereignis <sup>2)</sup> [25.; 75. Perzentil] (Monate)	4,07 [1,31; n. b.]	1,38 [0,39; 4,9]	
Mediane Beobachtungszeit	k. A.	k. A.	
Anzahl der Zensierungen, n (%) <sup>5)</sup>	24 (35,8) <sup>5)</sup>	11 (16,9) <sup>5)</sup>	
Studienabbruch vor erster VOC, n (%) <sup>6)</sup>	9 (13,4) <sup>6)</sup>	6 (9,2) <sup>6)</sup>	
Studienabbruchgründe, n (%)	k. A.	k. A.	
Erreichen des Behandlungsendes ohne VOC, n (%) <sup>7)</sup>	15 (22,4) <sup>7)</sup>	5 (7,7) <sup>7)</sup>	
Sensitivitätsanalyse mit Imputationsmodell – VOC gemäß ärztlichem Prüfpersonal			
PhS-M6b: Mediane Zeit bis zum Ereignis <sup>2)8)9)10)</sup> (Monate)	N = 66 <sup>7)</sup> 3,78	1,15	0,54 <sup>3)</sup> [0,36; 0,81]; k. A.
<b>Zeit bis zur zweiten VOC<sup>1)</sup></b>			
Primäre Analyse – VOC gemäß CRC			
Personen mit Ereignis, n (%)	30 (44,8)	38 (58,5)	0,53 [0,33; 0,87] <sup>3)</sup> ; 0,022 <sup>4)</sup>
Mediane Zeit bis zum Ereignis <sup>2)</sup> [25.; 75. Perzentil] (Monate)	10,32 [4,47; n. b.]	5,09 [2,96; 11,01]	
Mediane Beobachtungszeit	k. A.	k. A.	
Anzahl der Zensierungen, n (%) <sup>5)</sup>	37 (55,2) <sup>5)</sup>	27 (41,5) <sup>5)</sup>	
Studienabbruch vor zweiter VOC, n (%) <sup>6)</sup>	k. A.	k. A.	
Studienabbruchgründe, n (%)	k. A.	k. A.	
Erreichen des Behandlungsendes ohne zweite VOC, n (%) <sup>7)</sup>	k. A.	k. A.	

- <sup>1)</sup> Die Zeit bis zur ersten bzw. zweiten VOC war definiert als die Zeit von der Randomisierung bis zum Auftreten der ersten bzw. zweiten VOC (VOC-Definition gemäß primärem Endpunkt). Falls kein entsprechendes Ereignis vor etwaigem Studienabbruch oder Behandlungsende (definiert als Zeitpunkt der letzten Dosis + 14 Tage) auftrat, erfolgte eine Zensurierung zum Zeitpunkt des Abbruchs der Behandlung bzw. zum Behandlungsende. Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt).
- <sup>2)</sup> Die Schätzung erfolgte mithilfe der Kaplan-Meier-Methode.
- <sup>3)</sup> Schätzung des HR mithilfe eines Cox-Regressionsmodells mit der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10), dem Einsatz einer begleitenden HU-Therapie (ja vs. nein) und der Behandlung als Kovariaten.
- <sup>4)</sup> Berechnung des p-Werts mittels Logrank-Test stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).
- <sup>5)</sup> Eigenberechnung: Spezifische Angaben lagen im Rahmen der Ergebnisdarstellungen der präspezifizierten Überlebenszeitanalyse nicht vor. Auf die Anzahl und den Anteil der Zensurierungen wurde aufgrund der Gesamtzahl der Personen im Behandlungsarm – der Anzahl der Personen mit Ereignis im jeweiligen Behandlungsarm geschlossen. Etwaige Zensurierungsgründe gehen aus der Präspezifizierung hervor (siehe Fußnote 1). Im Crizanlizumab-Arm wurden eine Person und im Placebo-Arm drei Personen nie behandelt.
- <sup>6)</sup> Eigenberechnung: Spezifische Angaben lagen im Rahmen der Ergebnisdarstellungen der präspezifizierten Überlebenszeitanalyse nicht vor. Auf die Anzahl und den Anteil der Zensurierungen wurde aufgrund der Gesamtzahl der Personen mit Zensurierungen – der Anzahl der Personen, welche das Behandlungsende ohne VOC erreichten, unter Berücksichtigung der präspezifizierten Zensurierungsgründe geschlossen.
- <sup>7)</sup> Die Anzahl der Personen mit „Erreichen des Behandlungsendes ohne VOC“ und deren Anteil ist den Angaben im der Post-hoc-Auswertung „Anteil der Personen ohne VOC während der Behandlungsphase“ entnommen.
- <sup>8)</sup> Die vorliegende Analyse wurde nicht mit dem Dossier vorgelegt. Angaben stammen aus dem EPAR [3].
- <sup>9)</sup> Hierbei handelt es sich um eine Post-hoc-Sensitivitätsanalyse (PhS). Dabei wurde eine Imputation der Anzahl der VOC bei vorzeitigem Behandlungsabbruch für Personen im Crizanlizumab-Arm über eine Jump-to-Reference-Methode basierend auf VOC von Personen aus dem Placebo-Arm vorgenommen. Für Personen im Placebo-Arm erfolgte die Ersetzung unter der Missing-at-Random-Annahme basierend auf Daten von Personen desselben Studienarmes vor Behandlungsabbruch. Eine Person wurde von der Analyse ausgeschlossen.
- <sup>10)</sup> Bei der von der Analyse ausgeschlossenen Person handelt es sich um einen statistischen Ausreißer mit 37 VOC in 6 Monaten, die zudem die Einschlusskriterien nicht erfüllte, da sie bereits im Jahr vor der Randomisierung 30 Schmerzkrisen erlitten hatte.

Abkürzungen: CRC: Crisis Review Committee; EPAR: European Public Assessment Report; HR: Hazard Ratio; HU: Hydroxyurea/Hydroxycarbamid; ITT: Intention-to-Treat; k. A.: keine Angabe; KI: Konfidenzintervall; M: Imputationsmethode; n. b.: nicht berechnet oder nicht berechenbar; PhS: Post-hoc-Sensitivitätsanalyse; VOC: Vasookklusive Schmerzkrisen.

In Abbildung 2 ist die Kaplan Meier-Kurve für den Endpunkt „Zeit bis zur ersten VOC“ basierend auf der primären Analyse dargestellt. Spezifische Angaben zur Differenzierung der Anzahl der Personen, welche nicht mehr unter Risiko waren (bspw. aufgrund von Zensurierungen), liegen nicht vor (siehe Tabelle 17).

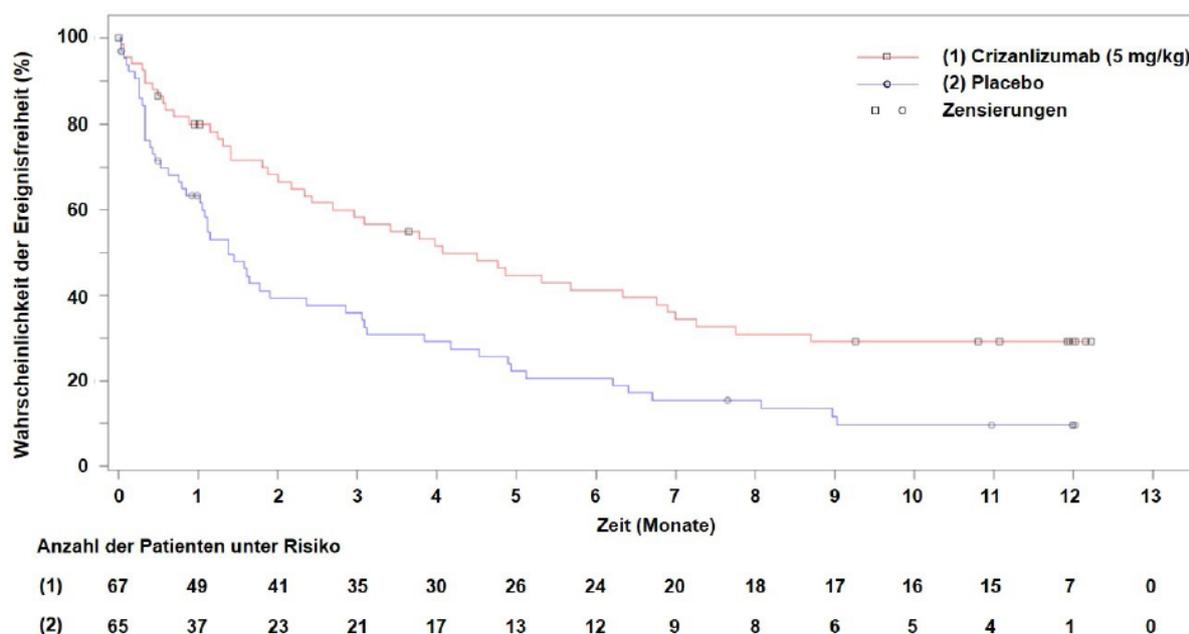


Abbildung 2: Kaplan Meier-Kurve für die Zeit bis zur ersten VOC (basierend auf primärer Analyse) [10]

Subklassifikationen: „Jährliche Rate unkomplizierte VOC“ und „Jährliche Rate akute Thoraxsyndrome“

Unkomplizierte VOC repräsentierten den Großteil der erfassten VOC und variierten über beide Studienarme hinweg je nach Einschätzungsvariante zwischen 89,1 % gemäß CRC und 92,95 % gemäß ärztlichem Prüfpersonal (Angaben zur Sicherheitspopulation, da zu ITT nicht vorliegend; siehe Tabelle 14). Der Anteil der unkomplizierten VOC an allen VOC war etwas höher im Placebo-Arm gegenüber dem Crizanlizumab-Arm (siehe Tabelle 14 und Tabelle 15). Einschränkungen zur Interpretierbarkeit der primären Analyse zu „Jährliche Rate unkomplizierte VOC“ ergeben sich weitestgehend analog zur primären Analyse des Endpunkts „Jährliche Rate VOC“. Die Richtung des Effektschätzers zur Wirksamkeit von Crizanlizumab gegenüber Placebo bleibt in der Sensitivitätsanalyse auch hinsichtlich der Subklassifikation „Jährliche Rate unkomplizierte VOC“ bestehen, wobei das 95%-KI analog zur Sensitivitätsanalyse Ph-SM6b aller VOC die 1 umfasst ([3], Daten nicht im Rahmen des Dossiers vorgelegt).

*Tabelle 18: Primäre Analyse und Sensitivitätsanalyse der Subklassifikation „Jährliche Rate unkomplizierte VOC“ – ITT-Population*

Studie SUSTAIN	Crizanlizumab 5,0 mg/kg N = 67	Placebo N = 65	HL-Schätzer des Medians der Differenzen [95%-KI]; p-Wert
Primäre Analyse ohne Berücksichtigung fehlender Werte <sup>1)</sup>			
„Jährliche Rate unkomplizierte VOC“ gemäß CRC <sup>2)</sup>			
Median (min; max)	1,08 (0; 3,96) <sup>3)</sup>	2,91 (1; 5) <sup>3)</sup>	
MW (SD)	2,66 (4,17)	3,88 (4,23)	
HL-Schätzer Median	1,97 <sup>4)</sup>	3 <sup>4)</sup>	-1 [-1,98; 0] <sup>4)</sup> ; 0,015 <sup>5)</sup>
			<b>Rate Ratio [95%-KI]; p-Wert</b>
Sensitivitätsanalyse mit Imputationsmodell			
PhS-M6b: „Jährliche Rate unkomplizierte VOC“ gemäß ärztlichem Prüfpersonal <sup>6)7)8)9)</sup> MW (SD) <sup>9)</sup>	3,39 (3,99)	4,79 (5,49)	0,72 [0,49; 1,05]; k. A.

<sup>1)</sup> In dieser Analyse wurden keine Imputationen vorgenommen. Fehlende Daten wurden durch Extrapolation der vorliegenden Daten auf ein Jahr nicht berücksichtigt. Vorzeitige Studienabbrüche betrafen jeweils 24 Personen in beiden Armen. 27 Personen im Crizanlizumab-Arm und 24 Personen im Placebo-Arm wiesen  $\geq 1$  der berichteten Protokollverletzungen auf.

<sup>2)</sup> „Jährliche Rate unkomplizierte VOC“ war definiert als die Anzahl unkomplizierter VOC einer Person während der Behandlungsphase  $\times 365 / (\text{Datum des Behandlungsendes} - \text{Datum der Randomisierung} + 1)$ , wobei das Behandlungsende definiert war als der Zeitpunkt der letzten Dosis + 14 Tage. Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt).

<sup>3)</sup> Angaben zu (min; max) lagen nicht vor und entstammen dem EunethHTA-Bericht [23].

<sup>4)</sup> Eine non-parametrische Schätzung erfolgte mithilfe der Hodges-Lehmann-Methode.

<sup>5)</sup> Berechnung des p-Werts mithilfe eines Wilcoxon-Rangsummen-Tests stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).

<sup>6)</sup> Diese Analyse wurde nicht mit dem Dossier vorgelegt. Die vorliegenden Angaben stammen aus dem EPAR [3].

<sup>7)</sup> Hierbei handelt es sich um eine Post-hoc-Sensitivitätsanalyse. Dabei wurde eine Imputation der Anzahl unkomplizierter VOC bei vorzeitigem Behandlungsabbruch für Personen im Crizanlizumab-Arm über eine Jump-to-Reference-Methode basierend auf unkomplizierten VOC von Personen aus dem Placebo-Arm vorgenommen. Für Personen im Placebo-Arm erfolgte die Ersetzung unter der Missing-at-Random-Annahme basierend auf Daten von Personen desselben Studienarmes vor Behandlungsabbruch. Eine Person wurde von der Analyse ausgeschlossen.

<sup>8)</sup> Bei der von der Analyse ausgeschlossenen Person handelt es sich um einen statistischen Ausreißer mit 37 VOC in 6 Monaten, die zudem die Einschlusskriterien nicht erfüllte, da sie bereits im Jahr vor der Randomisierung 30 Schmerzkrisen erlitten hatte.

<sup>9)</sup> MW und SD inklusive aller Imputationen.

Abkürzungen: CRC: Critical Review Committee; EPAR: European Public Assessment Report; HL: Hodges-Lehmann; HU: Hydroxyurea/Hydroxycarbamid; ITT: Intention-to-Treat; k. A.: keine Angabe; KI: Konfidenzintervall; M: Imputationsmethode; MW: Mittelwert; PhS: Post-hoc-Sensitivitätsanalyse; SD: Standardabweichung; VOC: Vasookklusive Schmerzkrise.

Deskriptiv waren ACS-Ereignisse anteilig an allen VOC etwas häufiger im Crizanlizumab- als im Placebo-Arm zu beobachten (siehe Tabelle 15), obwohl gemäß medizinischer Vorgeschichte in der Vergangenheit ACS-Ereignisse fast doppelt so häufig im Placebo-Arm (21,5 %) als im Crizanlizumab-Arm (11,9 %) auftraten (siehe Tabelle 12). Es ist zu beachten, dass die ACS-Häufigkeit in beiden Studienarmen niedrig war, welches sich in einem Median der Subklassifikation „Jährliche Rate akute Thoraxsyndrome“ von 0 in beiden Behandlungsgruppen widerspiegelt. Einschränkungen zur Interpretierbarkeit der primären Analyse zu „Jährliche Rate akute Thoraxsyndrome“ (siehe Tabelle 19) ergeben sich weitestgehend analog zur primären Analyse des Endpunkts „Jährliche Rate VOC“.

*Tabelle 19: Primäre Analyse der Subklassifikation „Jährliche Rate akute Thoraxsyndrome“ – ITT-Population*

	<b>Crizanlizumab 5,0 mg/kg N = 67</b>	<b>Placebo N = 65</b>	<b>HL-Schätzer des Medians der Differenzen [95%-KI]; p-Wert</b>
Primäre Analyse ohne Berücksichtigung fehlender Werte <sup>1)</sup>			
„Jährliche Rate akute Thoraxsyndrome“ gemäß CRC			
Median (min; max)	0 (k. A.; k. A.)	0 (k. A.; k. A.)	
MW (SD)	0,24 (0,65)	0,6 (3,05)	
HL-Schätzer Median	0 <sup>3)</sup>	0 <sup>3)</sup>	0 [0; 0] <sup>3)</sup> ; 0,78 <sup>4)</sup>

<sup>1)</sup> In dieser Analyse wurden keine Imputationen vorgenommen. Fehlende Daten wurden durch Extrapolation der vorliegenden Daten auf ein Jahr nicht berücksichtigt. Vorzeitige Studienabbrüche betrafen jeweils 24 Personen in beiden Armen. 27 Personen im Crizanlizumab-Arm und 24 im Placebo-Arm wiesen  $\geq 1$  der berichteten Protokollverletzungen auf.

<sup>2)</sup> Der Endpunkt „Jährliche Rate akute Thoraxsyndrome“ war definiert als die Anzahl akuter Thoraxsyndrome einer Person während der Behandlungsphase  $\times 365 / (\text{Datum des Behandlungsendes} - \text{Datum der Randomisierung} + 1)$ , wobei das Behandlungsende definiert war als der Zeitpunkt der letzten Dosis + 14 Tage. Für Personen, die nie behandelt wurden, war das Behandlungsende definiert als Datum der Studienbeendigung (d. h. letzter Kontakt).

<sup>3)</sup> Eine non-parametrische Schätzung erfolgte mithilfe der Hodges-Lehmann-Methode.

<sup>4)</sup> Berechnung des p-Werts mithilfe eines Wilcoxon-Rangsummen-Tests stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).

Abkürzungen: CRC: Critical Review Committee; HL: Hodges-Lehmann; HU: Hydroxyurea/Hydroxycarbamid; k. A.: keine Angabe; KI: Konfidenzintervall; ITT: Intention-to-Treat; MW: Mittelwert; SD: Standardabweichung; VOC: Vasookklusive Schmerzkrise.

## BPI-LF

Daten zu wahrgenommenen Schmerzen wurden mithilfe des BPI-LF erhoben.

*Tabelle 20: Rücklaufquoten des BPI-LF in der Studie SUSTAIN – ITT-Population*

Zeitpunkt	Crizanlizumab 5,0 mg/kg N = 67		Placebo N = 65	
	Anzahl der ausgefüllten Fragebogen (n)	Rücklauf- quote (%)	Anzahl der ausgefüllten Fragebogen (n)	Rücklauf- quote (%)
Baseline	48	71,6	55	84,6
Tag 15	46	68,7	52	80,0
Woche 14	42	62,7	37	56,9
Woche 26	33	49,3	33	50,8
Woche 38	33	49,3	33	50,8
Woche 52	22	32,8	24	36,9
Follow-up-Visite <sup>1)</sup>	34	50,7	35	53,8

<sup>1)</sup> Die Angaben umfassen jeweils die Teilnehmenden, welche entweder an der regulären Follow-up-Visite in Woche 58 oder der Follow-up-Visite mindestens 6 Wochen nach Behandlungsabbruch teilgenommen haben.

Abkürzungen: BPI-LF: Brief Pain Inventory – Long Form; ITT: Intention-to-Treat.

Aufgrund der geringen Rücklaufquoten (< 70 %) werden die Ergebnisse zum BPI-LF nicht dargestellt.

### 3.4 Lebensqualität

Die Ergebnisse zur Lebensqualität werden ergänzend im Anhang dargestellt.

### 3.5 Sicherheit

In der Studie SUSTAIN war die mediane Expositionsdauer mit der jeweiligen Studienintervention vergleichbar zwischen den Behandlungsarmen. Angaben zur medianen Beobachtungsdauer liegen nicht vor (siehe Tabelle 10). Im Vergleich zur ITT-Population umfasst die Sicherheitspopulation 1 Person weniger im Crizanlizumab-Arm und 3 Personen weniger im Placebo-Arm. Die deskriptiven Ergebnisdarstellungen deuten bezüglich der Gesamtrate der UE, schweren UE, SUE und UE, die zum Abbruch führten, auf vergleichbare Raten zwischen den Behandlungsarmen hin. Bei der Interpretation der Ergebnisse zu schweren UE ist zu berücksichtigen, dass diese als Verhinderung von Alltagsaktivitäten definiert waren, deren Reliabilität eingeschränkt erscheint.

Tabelle 21: Gesamtraten unerwünschter Ereignisse – Sicherheitspopulation

Personen mit mindestens einem ...	Crizanlizumab 5,0 mg/kg N = 66 n (%) <sup>1)</sup>	Placebo N = 62 n (%) <sup>1)</sup>	RR [95%-KI] <sup>2)</sup> ; p-Wert <sup>2)3)</sup>
UE	57 (86,4)	55 (88,7)	... <sup>4)</sup>
Schwere UE <sup>5)</sup>	12 (18,2)	12 (19,4)	0,94 [0,46; 1,93]; 0,858
SUE	17 (25,8)	17 (27,4)	0,94 [0,53; 1,67]; 0,820
UE, das zum Abbruch führte	2 (3,0) <sup>6)</sup>	3 (4,8)	0,63 [0,11; 3,62]; 0,595

<sup>1)</sup> Personen mit mehrfachen Ereignissen in der gleichen Kategorie werden nur einmal in dieser gezählt.

<sup>2)</sup> Effektschätzer und p-Werte wurden post hoc berechnet.

<sup>3)</sup> Berechnung des p-Werts mithilfe eines Cochran-Mantel-Haenszel-Tests stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).

<sup>4)</sup> Aufgrund unklarer Patientenrelevanz von in die Kategorie UE eingeflossenen Ereignissen, werden weder Effektschätzer noch p-Wert angegeben.

<sup>5)</sup> Die Bewertung des Schweregrads erfolgte nach den Einstufungen leicht (keine Beeinträchtigung von Alltagsaktivitäten), moderat (Beeinträchtigung von Alltagsaktivitäten) und schwer (Verhinderung von Alltagsaktivitäten).

<sup>6)</sup> Eine der beiden Personen beendete die Behandlung mit Crizanlizumab aufgrund des UE „Sichelzellanämie mit Krise“ und verstarb in der Folge. Bei den Angaben zu den Studiencharakteristika in Tabelle 10 wurde für diese Person jedoch „Tod“ als Abbruchgrund gelistet, weshalb dort „Unerwünschtes Ereignis“ lediglich bei einer Person als Abbruchgrund angegeben wurde.

Abkürzungen: HU: Hydroxyurea/Hydroxycarbamid; KI: Konfidenzintervall; RR: Relatives Risiko; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis; VOC: Vasookklusive Schmerzkrise.

## Unerwünschte Ereignisse und schwere unerwünschte Ereignisse nach Systemorgan- klassen und Preferred Terms

UE mit einer Inzidenz  $\geq 10\%$  nach Version 16 der MedDRA-Systemorganklassen und Preferred Terms sind in Tabelle 22 dargestellt. Ein statistisch signifikanter Unterschied in den post hoc erfolgten Berechnungen wurde ausschließlich zur Systemorganklasse „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ zuungunsten von Crizanlizumab mit einem relativen Risiko (RR) von 4,70 und breitem 95%-KI [1,07; 20,59] berichtet. Die Anzahl der in diese Systemorganklasse einfließenden Preferred Terms ( $n = 9$ ) zeigt eine hohe Variabilität auf, wobei lediglich die diesbezüglichen Preferred Terms „Handfraktur“ ( $n_C = 2$ ;  $n_P = 0$ ) und „Infusionsbedingte Reaktion“ ( $n_C = 2$ ;  $n_P = 0$ ) innerhalb eines Studienarmes mehrfach genannt wurden.

SUE mit einer Inzidenz von  $\geq 5\%$  wurden in beiden Behandlungsarmen nur bezüglich der Systemorganklasse „Infektionen und parasitäre Erkrankungen“ berichtet, wobei jedoch kein statistisch signifikanter Unterschied im Rahmen der Post-hoc-Auswertungen vorgefunden wurde.

Tabelle 22: UE mit Inzidenz  $\geq 10\%$  und schwere UE mit Inzidenz  $\geq 5\%$  nach MedDRA-Systemorganklassen und Preferred Terms in der Studie SUSTAIN – Sicherheitspopulation

MedDRA-Systemorganklasse <sup>1)</sup> Preferred Term	Crizanlizumab 5,0 mg/kg N = 66 n (%) <sup>2)</sup>	Placebo N = 62 n (%) <sup>2)</sup>	RR [95%-KI] <sup>3)</sup> ; p-Wert <sup>3)4)</sup>
<b>Unerwünschte Ereignisse mit Inzidenz <math>\geq 10\%</math></b>			
<b>Erkrankungen des Blutes und des Lymphsystems</b>	<b>8 (12,1)</b>	<b>10 (16,1)</b>	<b>0,75 [0,32; 1,78]; 0,526</b>
<b>Erkrankungen des Gastrointestinaltrakts</b>	<b>26 (39,4)</b>	<b>15 (24,2)</b>	<b>1,63 [0,96; 2,77]; 0,069</b>
Diarrhö	7 (10,6)	2 (3,2)	3,29 [0,71; 15,23]; 0,102
Übelkeit	12 (18,2)	7 (11,3)	1,61 [0,68; 3,83]; 0,281
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	<b>24 (36,4)</b>	<b>18 (29,0)</b>	<b>1,25 [0,76; 2,07]; 0,389</b>
Pyrexie	7 (10,6)	4 (6,5)	1,64 [0,51; 5,34]; 0,397
<b>Infektionen und parasitäre Erkrankungen</b>	<b>34 (51,5)</b>	<b>33 (53,2)</b>	<b>0,97 [0,70; 1,35]; 0,844</b>
Infektion der oberen Atemwege	7 (10,6)	6 (9,7)	1,10 [0,39; 3,08]; 0,847
Harnwegsinfektion	9 (13,6)	7 (11,3)	1,21 [0,48; 3,05]; 0,684
<b>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</b>	<b>10 (15,2)</b>	<b>2 (3,2)</b>	<b>4,70 [1,07; 20,59]; 0,022</b>
<b>Untersuchungen</b>	<b>14 (21,2)</b>	<b>18 (29,0)</b>	<b>0,73 [0,40; 1,34]; 0,315</b>
<b>Stoffwechsel- und Ernährungsstörungen</b>	<b>4 (6,1)</b>	<b>9 (14,5)</b>	<b>0,42 [0,14; 1,29]; 0,118</b>
<b>Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen</b>	<b>27 (40,9)</b>	<b>18 (29,0)</b>	<b>1,41 [0,87; 2,29]; 0,162</b>
Arthralgie	12 (18,2)	5 (8,1)	2,25 [0,84; 6,03]; 0,093
Rückenschmerzen	10 (15,2)	7 (11,3)	1,34 [0,54; 3,31]; 0,508
Schmerzen des Muskel- und Skelettsystems	8 (12,1)	6 (9,7)	1,25 [0,46; 3,41]; 0,672
Schmerzen in einer Extremität	11 (16,7)	10 (16,1)	1,03 [0,47; 2,26]; 0,926

MedDRA-Systemorganklasse <sup>1)</sup> Preferred Term	Crizanlizumab 5,0 mg/kg N = 66 n (%) <sup>2)</sup>	Placebo N = 62 n (%) <sup>2)</sup>	RR [95%-KI] <sup>3)</sup> ; p-Wert <sup>3)4)</sup>
<b>Erkrankungen des Nervensystems</b>	<b>21 (31,8)</b>	<b>15 (24,2)</b>	<b>1,32 [0,75; 2,31]; 0,339</b>
Kopfschmerz	11 (16,7)	10 (16,1)	1,03 [0,47; 2,26]; 0,924
<b>Erkrankungen der Atemwege, des Brustraums und Mediastinums</b>	<b>13 (19,7)</b>	<b>16 (25,8)</b>	<b>0,76 [0,40; 1,45]; 0,419</b>
Husten	4 (6,1)	7 (11,3)	0,54 [0,17; 1,74]; 0,301
<b>Erkrankungen der Haut und des Unterhautzellgewebes</b>	<b>12 (18,2)</b>	<b>9 (14,5)</b>	<b>1,25 [0,57; 2,76]; 0,587</b>
<b>Schwere unerwünschte Ereignisse mit Inzidenz <math>\geq</math> 5 %<sup>5)</sup></b>			
<b>Infektionen und parasitäre Erkrankungen</b>	<b>5 (7,6)</b>	<b>3 (4,8)</b>	<b>1,57 [0,39; 6,28]; 0,539</b>

<sup>1)</sup> Systemorganklassen und Preferred Terms gemäß MedDRA (Version 16.1).

<sup>2)</sup> Personen mit mehrfachen Ereignissen in der gleichen Kategorie werden nur einmal in dieser gezählt.

<sup>3)</sup> Effektschätzer und p-Werte wurden post hoc berechnet.

<sup>4)</sup> Berechnung des p-Werts mithilfe eines Cochran-Mantel-Haenszel-Tests stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).

<sup>5)</sup> Die Bewertung des Schweregrads erfolgte nach den Einstufungen leicht (keine Beeinträchtigung von Alltagsaktivitäten), moderat (Beeinträchtigung von Alltagsaktivitäten) und schwer (Verhinderung von Alltagsaktivitäten).

Abkürzungen: HU: Hydroxyurea/Hydroxycarbamid; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; RR: Relatives Risiko; UE: Unerwünschtes Ereignis; VOC: Vasookklusive Schmerzkrise.

### Schwerwiegende Unerwünschte Ereignisse nach Systemorganklassen

Es zeigten sich keine statistisch signifikanten Unterschiede in SUE zwischen den Behandlungsarmen Crizanlizumab 5,0 mg/kg und Placebo (Tabelle 23).

Tabelle 23: SUE mit Inzidenz  $\geq$  5 % nach MedDRA-Systemorganklassen und Preferred Terms in der Studie SUSTAIN – Sicherheitspopulation

MedDRA-Systemorganklasse <sup>1)</sup>	Crizanlizumab 5,0 mg/kg N = 66	Placebo N = 62	RR [95%-KI] <sup>2)</sup> ; p-Wert <sup>2)3)</sup>
Infektionen u. parasitäre Erkrankungen, n (%) <sup>4)</sup>	8 (12,1)	10 (16,1)	0,75 [0,32; 1,78]; 0,512

<sup>1)</sup> Gemäß MedDRA-Version 16.1.

<sup>2)</sup> Effektschätzer und p-Werte wurden post hoc berechnet.

<sup>3)</sup> Berechnung des p-Werts mithilfe eines Cochran-Mantel-Haenszel-Tests stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).

<sup>4)</sup> Personen mit mehrfachen Ereignissen in der gleichen Kategorie werden nur einmal in dieser gezählt.

Abkürzungen: HU: Hydroxyurea/Hydroxycarbamid; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; RR: Relatives Risiko; SUE: Schwerwiegendes unerwünschtes Ereignis; VOC: Vasookklusive Schmerzkrise.

### UE von besonderem Interesse

Post hoc wurden „Infektionen“, „Infusionsbedingte Reaktionen“ und „Schwere infusionsbedingte Reaktionen“ als UE von besonderem Interesse definiert. Es zeigten sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsarmen Crizanlizumab 5,0 mg/kg und Placebo (siehe Tabelle 24). Infektionen wurden bei etwas mehr als der Hälfte aller Patientinnen und Patienten berichtet. Deskriptive Auffälligkeiten zeigten sich bezüglich infusionsbedingter Reaktionen, welche mit 23 Fällen gegenüber 13 Fällen häufiger im Crizanlizumab-Arm auftraten. Es wurden 2 Fälle schwerer infusionsbedingter Reaktionen im Crizanlizumab-Arm gegenüber keinen Fällen im Placebo-Arm berichtet.

Tabelle 24: UE von besonderem Interesse

UE von besonderem Interesse <sup>1)</sup>	Crizanlizumab 5,0 mg/kg N = 66 n (%) <sup>2)</sup>	Placebo N = 62 n (%) <sup>2)</sup>	RR [95%-KI] <sup>3)</sup> ; p-Wert <sup>3)4)</sup>
Infektionen	35 (53)	33 (53,2)	1,00 [0,72; 1,38]; 0,979
Infusionsbedingte Reaktionen	23 (34,8)	13 (21)	1,66 [0,93; 2,98]; 0,084
Schwere infusionsbedingte Reaktionen	2 (3)	0	n. b.; 0,168
Wirkung auf Blutgerinnung – Blutung <sup>5)</sup>	11 (16,7)	8 (12,9)	1,29 [0,56; 3,00]; 0,557

<sup>1)</sup> Systemorganklassen und Preferred Terms gemäß MedDRA (Version 21.1). Es liegen keine weiteren Angaben vor.

<sup>2)</sup> Personen mit mehrfachen Ereignissen in der gleichen Kategorie werden nur einmal in dieser gezählt.

<sup>3)</sup> Effektschätzer und p-Werte wurden post hoc berechnet.

<sup>4)</sup> Berechnung des p-Werts mithilfe eines Cochran-Mantel-Haenszel-Tests stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).

<sup>5)</sup> Einschließlich veränderter Gerinnungsparameter

Abkürzungen: HU: Hydroxyurea/Hydroxycarbamid; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; n. b.: nicht berechnet oder nicht berechenbar; RR: Relatives Risiko; UE: Unerwünschtes Ereignis; VOC: Vaso-okklusive Schmerzkrise.

## 4 Diskussion der Methodik und Ergebnisse

### 4.1 Zulassungsstatus und Zulassungspopulation von Crizanlizumab

Gemäß Zulassung kann Crizanlizumab (Adakveo®) zusammen mit Hydroxyurea/Hydroxycarbamid (HU) oder als Monotherapie bei Patientinnen und Patienten, bei denen die Anwendung von HU nicht geeignet oder unzureichend ist, gegeben werden [9]. Die vorliegende Bewertung bezieht sich auf die Indikation der Prävention wiederkehrender vasookklusiver Krisen (VOC) bei Patientinnen und Patienten ab 16 Jahren mit Sichelzellerkrankung.

Für die Bewertung liegen Daten der pivotalen Studie SUSTAIN vor, in welcher Patientinnen und Patienten zwischen 16 und 65 Jahren mit Sichelzellerkrankung eingeschlossen wurden, die in den 12 Monaten vor Studieneinschluss zwischen 2 und 10 VOC aufweisen sollten. Patientinnen und Patienten mit bspw. mehr als 10 VOC in den vergangenen 12 Monaten sollten demnach nicht eingeschlossen werden, welches eine Abweichung von der Zulassungspopulation darstellt. Es wurde Crizanlizumab in den Dosierungen 5,0 und 2,5 mg/kg eingesetzt. Für die Nutzenbewertung wird nur der Behandlungsarm mit Crizanlizumab, in dem dieses in der Fachinformation-konformen Dosierung von 5,0 mg/kg untersucht wurde, im Vergleich zu Placebo herangezogen. Die Anzahl von Patientinnen und Patienten im Alter zwischen 16 und 18 Jahren war aufgrund von Protokolländerungen in der Studie SUSTAIN limitiert, jedoch liegen keine spezifischen Angaben zu Altersgruppen in der Studienpopulation vor. In der Studienpopulation waren keine Personen über 63 Jahren enthalten. Weitere Unterschiede der Studienpopulation zur Zulassungspopulation bestehen hinsichtlich der erlaubten Therapieoptionen, da die Planung einer Teilnahme an einem chronischen Transfusionsprogramm und die Durchführung einer Austauschtransfusion bei Rekrutierung sowie die chronische Behandlung mit Antikoagulantien (außer Aspirin) während der Studienteilnahme ausgeschlossen wurden. Aufgrund des Ausschlusses von Personen mit geplanten großen chirurgischen Eingriffen während des Studienzeitraums, scheint die Durchführung einer hämatopoetischen Stammzelltransplantation als einzige kurative Therapieoption in der Studie SUSTAIN keine Therapiemöglichkeit gewesen zu sein. Es wurden 118 Personen von insgesamt 329 Personen aufgrund der Ein- und Ausschlusskriterien nicht in die Studie eingeschlossen. Differenzierte Angaben dazu liegen nicht vor. Daher können etwaige Auswirkungen auf die externe Validität der Studienpopulation bezüglich der Gesamtpopulation von Patientinnen und Patienten mit Sichelzellerkrankung und wiederkehrenden VOC nicht abgeschätzt werden.

Eine Behandlung mit Medikamenten, die der Standardtherapie von Patientinnen und Patienten mit Sichelzellerkrankung entsprechen, war während des Studienzeitraums grundsätzlich erlaubt. Gemäß Zulassung kann Crizanlizumab zusammen mit HU gegeben werden oder als Monotherapie bei Personen, bei denen die Anwendung von HU nicht geeignet oder unzureichend ist. Informationen zur Studienpopulation hinsichtlich der Eignung einer HU-Therapie bei Personen ohne diese Begleitmedikation sowie zur Hinlänglichkeit einer HU-Therapie bei Personen mit HU-Behandlung (die anteilmäßig auf 65 % beschränkt waren) konnten nicht identifiziert werden. Die Prävention von wiederkehrenden VOC als Studienendpunkt wurde in der Studie SUSTAIN auf jene VOC beschränkt, welche u. a. die Vorstellung in einer medizinischen Einrichtung und die orale bzw. parenterale Behandlung mit Opioiden oder die parenterale Behandlung mit NSAID erforderten. Die Mehrheit der Studienzentren befindet sich in den USA und die weiteren Studienzentren in Südamerika. Etwa 91 % der Studienpopulation war schwarz oder afroamerikanischer Abstammung. Die Übertragbarkeit der Studienpopulation auf den deutschen Versorgungskontext ist unklar.

Die Genehmigung für das Inverkehrbringen von Crizanlizumab wurde unter „besonderen Bedingungen“ erteilt [3]. Crizanlizumab wurde ursprünglich von Selexys Pharmaceuticals Corporation (bzw. Reprixys Pharmaceuticals) unter dem Namen „SelG1“ entwickelt und wird nach Übernahme durch Novartis in veränderter Form unter dem Namen „SEG101“ vertrieben. In der pivotalen Studie SUSTAIN kam Crizanlizumab in Form des ursprünglichen Wirkstoffs „SelG1“ zum Einsatz. Die EMA schätzt die Pharmakodynamik der beiden Wirkstoffe jedoch als vergleichbar ein. Zum neuen Wirkstoff „SEG101“ müssen die finalen Ergebnisse der Studie CSEG101A2202 (SOLACE-adults) und die primären Analyseergebnisse der Phase-III-Studie CSEG101A2301 (STAND) bis Dezember 2025 bei der EMA vorgelegt werden. In der multizentrischen, doppelblinden, placebokontrollierten Phase-III-Studie STAND wird Crizanlizumab in zwei Dosen (5,0 mg/kg oder 7,5 mg/kg) mit oder ohne HU bei 240 Patientinnen und Patienten ab 12 Jahren mit Sichelzellerkrankung und mindestens zwei VOC in den vergangenen 12 Monaten untersucht. Im Studienregister wird für die Studie STAND als Schätzung für die letzte Datenerhebung der/des letzten Studienteilnehmenden in Bezug auf den primären Endpunkt der 30. Mai 2022 angegeben [19,21]. Das Studienende ist voraussichtlich am 6. Dezember 2027.

## 4.2 Design und Methodik der Studie

### Zusammenfassung des Studiendesigns

Die Nutzenbewertung stützt sich auf die abgeschlossene pivotale Studie SUSTAIN, einer Phase-II-Studie zur Bewertung der Wirksamkeit und Sicherheit von 2 Dosierungen des Wirkstoffs Crizanlizumab bei der Prävention wiederkehrender VOC bei Patientinnen und Patienten ab 16 Jahren mit Sichelzellerkrankung. Es wurden die Dosierungen 5,0 und 2,5 mg/kg gegenüber Placebo in drei parallelen Behandlungsarmen untersucht. Die Zuteilung zu den Behandlungsarmen erfolgte anhand einer 1:1:1-Randomisierung mit einer Stratifizierung nach Begleitbehandlung mit HU (ja vs. nein) und der Anzahl an VOC während der vergangenen 12 Monate (2–4 vs. 5–10). Der Anteil der Patientinnen und Patienten mit HU war auf 65 % limitiert. Nach einer 30-tägigen Screeningphase, folgte bei Einschluss eine 52-wöchige Behandlungsphase mit Studienmedikationsgaben an Tag 1 und Tag 15 sowie anschließend in einem 4-wöchigen Rhythmus bis Woche 50 sowie einer Behandlungsabschlussuntersuchung in Woche 52. Den Studienunterlagen war zu entnehmen, dass der Zeitpunkt der Randomisierung i. d. R. an Tag 1 der Behandlung stattgefunden hat. Unklar ist, ob es von dieser Regel Abweichungen gab, da den Unterlagen keine systematische Darstellung der Zeit zwischen Randomisierung und initialer Medikation zu entnehmen war. Die Beobachtungszeit von Randomisierung bis zur Behandlungsabschlussuntersuchung in Woche 52 stellt die Grundlage aller Wirksamkeitsanalysen dar, mit Ausnahme jener der patientenberichteten Endpunkte (BPI-LF und SF-36), welche die Follow-up-Visite umfassten. Die Follow-up-Visite, welche insbesondere in den Sicherheitsauswertungen berücksichtigt wurde, erfolgte in Woche 58. Die Behandlung mit Crizanlizumab erfolgte bis zum Auftreten von UE, die zum Abbruch führten, von inakzeptablen QTc-Intervallerhöhungen, die zum Abbruch führten, von Schwangerschaft, bis zum Widerruf der Einverständniserklärung, bis zum Tod, bis zu einem anderweitig begründeten Studienabbruch oder bis zur Beendigung der 50-wöchigen Behandlung. Den Studienunterlagen konnten keine eindeutig differenzierenden Angaben zum Behandlungs- bzw. Studienabbruchvorgehen entnommen werden. In die Studie wurden 198 Personen in die bewertungsrelevanten Behandlungsarme eingeschlossen, davon 67 in den Crizanlizumab-Arm (5,0 mg/kg) und 65 in den Placebo-Arm.

Im Rahmen der Wirksamkeitsuntersuchung der Studie wurde in erster Linie die Veränderung von VOC untersucht. Als primärer Endpunkt der Studie wurde „Jährliche Rate VOC“ gemäß CRC

definiert. Weitere Endpunkte sind „Jährliche Rate der Krankenhaustage“, „Zeit bis zur ersten VOC“, „Zeit bis zur zweiten VOC“, „Jährliche Rate unkomplizierte VOC“, „Jährliche Rate akute Thoraxsyndrome“, „Jährliche Rate der transfundierten Erythrozytenkonzentrate“, „Jährliche Rate der Schmerzkrisentage“, „Anteil der Personen ohne VOC während der Behandlungsphase“ sowie UE und eine Vielzahl an Laboruntersuchungen. Patientenberichtete Endpunkte wurden mittels BPI-LF und SF-36 erhoben.

A priori war eine Gesamtstudiendauer von 25 Monaten an 70 Studienzentren geplant. Die tatsächliche Gesamtstudiendauer vom ersten Studieneinschluss bis zum Studienabschluss belief sich auf etwa 32 Monate und fand an 60 Studienzentren bis zum Studienabschluss der/des letzten Teilnehmenden am 23 März 2016 statt. Rationale für diese Abweichungen liegen nicht vor. Nach Angaben des pU gegenüber der EMA wurde die Datenbank am 13. Mai 2016 gesperrt [3].

### **Protokolländerungen**

Es liegen 4 Protokollversionen vor, wovon 2 Protokolle nach Studienbeginn in Kraft traten und zu welchen unklar ist, wie viele Patientinnen und Patienten jeweils von den Änderungen betroffen waren. Maßgebliche studienführungsspezifische Änderungen konnten dabei nicht identifiziert werden. Jedoch wurde gemäß Studienbericht mit Protokollversion 3 spezifiziert, dass anhand einer statistischen Analyse ein sekundärer Schlüsselendpunkt identifiziert werden sollte. In Protokollversion 4 wurde dieser sekundäre Schlüsselendpunkt spezifiziert, weshalb unklar ist, ob und unter welchen Bedingungen etwaige Interimsanalysen stattgefunden haben. Protokollversion 4 (21. April 2016) erlangte erst nach Studienabschluss (der/des letzten Teilnehmenden am 23. März 2016) Gültigkeit und bezieht sich somit ausschließlich auf den Analyseplan, welcher im später finalisierten SAP (2. Mai 2016) aufgeführt ist. Die Änderungen umfassten u. a. das Entfernen von Sensitivitätsanalysen mit Imputationen, welches neben dem Einreichen einer Vielzahl von Post-hoc-Sensitivitätsanalysen begründet, warum die EMA eine ergebnisgesteuerte Berichterstattung nicht ausschließen konnte [3]. Daher erfolgte in der vorliegenden Nutzenbewertung eine Auswahl zum Heranziehen von einzelnen Sensitivitätsanalysen basierend auf deren methodischer Nachvollziehbarkeit.

### **Studienpopulation**

Die Studienpopulation der bewertungsrelevanten Studienarme besteht aus Paersonen mit einer bestätigten Diagnose einer Sichelzellerkrankung, die nach Begleitbehandlung mit HU (ja vs. nein) und der Anzahl an VOC während der vergangenen 12 Monate (2–4 vs. 5–10) stratifiziert wurden. Ungefähr zwei Drittel aller Patientinnen und Patienten erhielt eine begleitende HU-Behandlung und sowohl diese als auch die kategoriale Stratifikation der Anzahl an VOC während der vergangenen 12 Monate erscheint zwischen den Behandlungsarmen weitestgehend ausgeglichen. Jedoch mussten gemäß Einschlusskriterien die Studienteilnehmenden zwischen 2 und 10 VOC in den 12 Monaten vor Studieneinschluss aufweisen, wobei insgesamt fünf Personen, drei Personen im Crizanlizumab-Arm und zwei Personen im Placebo-Arm, dieses Einschlusskriterium nicht erfüllten. Eine dieser Personen im Crizanlizumab-Arm wurde im späteren Analyseverlauf für die Sensitivitätsanalysen gemäß PhS-M6b ausgeschlossen.

Hinsichtlich der Genotypen wurden relevante deskriptive Unterschiede ausschließlich hinsichtlich der Thalassämien berichtet, sodass 4,5 % (n = 3) im Crizanlizumab-Arm gegenüber 10,8 % (n = 7) im Placebo-Arm eine HbS $\beta$ 0-Thalassämie bzw. 10,4 % (n = 7) im Crizanlizumab-Arm gegenüber

1,5 % (n = 1) im Placebo-Arm eine HbS $\beta$ +Thalassämie aufwiesen. Für diesbezüglich aussagekräftige Subgruppenanalysen erscheinen die jeweiligen Gruppenumfänge jedoch zu klein. Das mediane Alter betrug zu Studienbeginn 28 Jahre (min; max: 16; 63), wobei aufgrund der protokollbedingt eingeschränkten Rekrutierungsdauer von Patientinnen und Patienten zwischen 16 und 18 Jahren von einer diesbezüglich limitierten Anzahl auszugehen ist [3]. Eindeutige Angaben zu Altersgruppierungen und insbesondere jener der 16- bis 18-Jährigen konnten nicht identifiziert werden. Im Placebo-Arm wurden deskriptiv mehr Frauen (58,5 %) als Männer (41,5 %) eingeschlossen, während die Geschlechter im Crizanlizumab-Arm ungefähr gleichverteilt waren (52,2 % Frauen und 47,8 % Männer). Deskriptive Angaben oder statistische Auswertungen zum Zeitraum seit Diagnosestellung der Sichelzellerkrankung konnten trotz Erfassung im eCRF nicht identifiziert werden. Diese hätten möglicherweise neben der Anzahl der VOC im Jahr vor Studieneinschluss und der medizinischen Vorgeschichte zur Einschätzung der Krankheitsschwere beitragen können, da diesbezüglich keine spezifische Bewertung bzw. Begutachtung zu Studieneinschluss stattgefunden hat. Hinsichtlich der medizinischen Vorgeschichte muss beachtet werden, dass keine Informationen dazu vorliegen, wie umfassend diese erfasst und überprüft wurde. Deskriptiv zeigten sich Auffälligkeiten mit einem Unterschied (von  $\geq 10$  Prozentpunkten bezüglich der betroffenen Personen) zwischen den beiden Behandlungsarmen hinsichtlich der Systemorganklassen „Herzerkrankungen“, „Infektionen und parasitäre Erkrankungen“, „Stoffwechsel- und Ernährungsstörungen“ und „Erkrankungen der Atemwege, des Brustraums und Mediastinums“, die deskriptiv häufiger von Personen im Placebo-Arm berichtet wurden (statistische Analysen wurden nicht berichtet).

Die eingereichten Studienunterlagen zur Begleitmedikation während der Studie lassen aufgrund von umfassenden Unklarheiten eine eindeutige Interpretation und Einschätzung zur Vergleichbarkeit zwischen den Studienarmen nicht zu. Seitens der EMA wurden aufgrund dessen Auswertungen gemäß des Anatomisch-therapeutisch-chemischen Klassifikationssystems (ATC-Klassen) nachgefordert, welche zur Nutzenbewertung nur im Rahmen des EPAR vorlagen [3]. Beim Vergleich der deskriptiven Angaben zeigt sich dabei, dass Medikamente subsummiert gemäß ATC-Klassen insgesamt häufiger im Crizanlizumab-Arm gegenüber Placebo angewendet worden sind. Medikamente die mit  $\geq 10$  Prozentpunkten Unterschied zwischen den Behandlungsarmen verschrieben wurden, gehörten zu den ATC-Klassen „Kardiovaskuläres System“ (n<sub>C</sub> = 52, n<sub>P</sub> = 42) und „Urogenitalsystem und Sexualhormone“ (n<sub>C</sub> = 51, n<sub>P</sub> = 39) [3]. Es ist zu berücksichtigen, dass eine Medikation mehr als einer ATC-Klasse zugeordnet werden konnte.

Von den teilnehmenden Personen brachen 35,85 % im Crizanlizumab-Arm (n = 24) und 36,9 % im Placebo-Arm (n = 24) die Studie vorzeitig ab und 38,6 % aller Patientinnen und Patienten wiesen gemäß pU mindestens eine relevante Protokollverletzung auf. Ähnliche Auswertungen zu weiteren potentiell relevanten Protokollverletzungen, wie bspw. Dosisverzögerung, Aufhebung der Verblindung oder weitere Verletzungen der Ein- und Ausschlusskriterien (außer der historischen VOC) konnten trotz einer Erfassung derselben im CRF nicht identifiziert werden. Während die Behandlungsdauer und Studienmedikationsgabe beider Arme vergleichbar waren, liegen keine näheren Angaben zur jeweiligen Beobachtungsdauer vor. Da bei Studienabbruch keine Nachbeobachtung bis Studienende stattgefunden hat und die Begründungen zu den Studienabbrüchen in vielen Fällen – wie bspw. bei Entzug der Einverständniserklärung von 7 Personen im Crizanlizumab-Arm gegenüber 6 Personen im Placebo-Arm oder bei Angabe von sonstigen Gründen durch 7 Personen im Crizanlizumab-Arm gegenüber 4 Personen im Placebo-Arm – einen potentiellen Zusammenhang mit Wirksamkeits- und Sicherheitsaspekten nicht zweifelsfrei ausschließen, wirkt sich dies negativ auf das Verzerrungspotential der Studie aus. Bezüglich des negativen Ausmaßes

dieser Aspekte wirkt sich limitierend aus, dass jeweils die vorzeitigen Studienabbrüche insgesamt und die mediane Behandlungsdauer zwischen den beiden Behandlungsarmen deskriptiv ähnlicher Ausprägung waren, sodass nicht unbedingt von einem differentiellen Informationsbias ausgegangen werden muss. Statistische Überprüfungen dazu, bspw. zu etwaigen Unterschieden der medianen Behandlungsdauer bzw. auch Angaben zur Beobachtungsdauer, lagen in den Studienunterlagen nicht vor.

### **4.3 Mortalität**

Todesfälle wurden im Rahmen der Sicherheitserfassung dokumentiert. Es wurden pro Arm jeweils 2 Todesfälle deskriptiv berichtet.

#### **Zusammenfassende Einschätzung zur Mortalität**

Aufgrund der hinsichtlich der Indikation geringen Studiendauer und der vielen Studienabbrüche lassen die in der Studie SUSTAIN beobachteten Ergebnisse keine validen Aussagen zum Effekt von Crizanlizumab auf das Gesamtüberleben zu.

### **4.4 Morbidität**

Für die Endpunktkategorie Morbidität reichte der pU Daten der Studie SUSTAIN mit „Jährliche Rate VOC“, „Jährliche Rate unkomplizierte VOC“, „Jährliche Rate akute Thoraxsyndrome“, „Zeit bis zur ersten VOC“, „Zeit bis zur zweiten VOC“, „Jährliche Rate der Schmerzkrisentage“ und „Anteil der Personen ohne VOC während der Behandlungsphase“ Endpunkte ein, die auf der Operationalisierung vasookklusiver Schmerzkrisen basieren. Außerdem legte der pU die Endpunkte „Jährliche Rate der Krankenhaustage“, „Jährliche Rate der Krankenhaustage aufgrund einer VOC“, „Jährliche Rate der transfundierten Erythrozytenkonzentrate“ sowie Auswertungen zum patientenberichteten Endpunkt „BPI-LF“ vor. Die Endpunkte „Jährliche Rate der Krankenhaustage“, „Jährliche Rate der Krankenhaustage aufgrund einer VOC“ und „Jährliche Rate der transfundierten Erythrozytenkonzentrate“ werden als nicht unmittelbar patientenrelevant bewertet, da sich etwaige Vorteile des zu bewertenden Arzneimittels in eindeutig patientenrelevanten Endpunkten, wie z. B. hinsichtlich der Krankheitssymptome, Lebensqualität oder Nebenwirkungen, widerspiegeln sollten. Diese Endpunkte werden daher nicht zur Nutzenbewertung herangezogen. Der Endpunkt „Jährliche Rate der Schmerzkrisentage“ wird aufgrund von umfangreichen Einschränkungen hinsichtlich der Validität, die u. a. die fehlende Standardisierung des Beginns und Endes einer VOC betreffen, nicht dargestellt. Ergebnisse aus dem Endpunkt „Anteil der Personen ohne VOC während der Behandlungsphase“ gingen bereits in die Analyse des Endpunkts „Zeit bis zur ersten VOC“ ein und sind nicht gesondert dargestellt. Der Endpunkt „BPI-LF“ bildet eher den chronisch wahrgenommenen Schmerz ab, das vorliegende Anwendungsgebiet bezieht sich jedoch auf wiederkehrende akute VOC. Zudem liegen die Rücklaufquoten im Interventionsarm bereits zum ersten Messzeitpunkt nach Studienbeginn (Tag 15) und auch im Verlauf unter 70 %. Die Ergebnisse im Studienverlauf werden daher nicht dargestellt.

### **Vasookklusive Schmerzkrisen**

VOC waren in der Studie SUSTAIN operationalisiert als:

- a) Akute Schmerzepisoden
- b) ohne andere Schmerzursache als ein vasookklusives Ereignis,
- c) bei denen die Vorstellung in einer medizinischen Einrichtung und
- d) die orale bzw. parenterale Behandlung mit Opioiden oder die parenterale Behandlung mit NSAID notwendig waren.

Jegliche VOC wurde dabei einer der fünf Subklassifikationen „Jährliche Rate unkomplizierte VOC“, „Jährliche Rate akute Thoraxsyndrome“, „Milzsequestration“, „Lebersequestration“ oder „Priapismus“ zugeordnet. Die ausschließlich deskriptiv durchgeführten Auswertungen zu Milzsequestration und Lebersequestration mit jeweils keinen Ereignissen in beiden Armen sowie Priapismus ( $n_P = 1$  gemäß CRC) zeigten bei marginalem Aufkommen keine bedeutsamen Unterschiede auf.

Die Validität der VOC-Endpunkte erscheint durch die Operationalisierung, insbesondere der allgemeinen Kriterien (a–d)), eingeschränkt. Das Aufsuchen einer medizinischen Einrichtung aufgrund einer VOC sowie die spezifische Behandlung mit Opioiden oder NSAID könnte – abgesehen vom Schweregrad des Ereignisses – jeweils in Abhängigkeit vom Krankenversicherungsstatus (bzw. der ökonomischen Situation), der Erreichbarkeit und Entfernung der nächsten medizinischen Einrichtung, dem gleichzeitigen Besitz von adäquaten Medikamenten und dem Versorgungskontext variieren. Dabei ist hinsichtlich der Einschränkungen der externen Validität zu berücksichtigen, dass die Studie SUSTAIN überwiegend in den USA durchgeführt wurde und alle teilnehmenden Studienzentren außerhalb Europas liegen.

VOC wurden innerhalb der SUSTAIN-Studie durch ärztliches Prüfpersonal erfasst und sollten anschließend durch ein verblindetes CRC adjudiziert werden. Alle präspezifizierten Analysen basierten auf diesen Einschätzungen gemäß CRC. Im Rahmen des EPAR wurden durch GCP-Inspektionen jedoch kritische Aspekte und Unklarheiten bezüglich des klinischen Adjudikationsprozesses und dessen Dokumentation aufgedeckt [3]. Die GCP-Untersucher der EMA empfahlen daher, die Daten des verblindeten CRC nicht zu akzeptieren und stattdessen Analysen basierend auf VOC-Einschätzungen gemäß ärztlichem Prüfpersonal zur Nutzen-Risiko-Bewertung heranzuziehen. In der vorliegenden Nutzenbewertung wird davon ausgegangen, dass – aufgrund der potentiellen Variabilität und einer Nichtberücksichtigung von Studienzentren als Stratifikationsfaktoren – die Validität der VOC sowohl gemäß Einschätzung durch das ärztliche Prüfpersonal als auch die Validität der CRC-adjudizierten VOC aufgrund der kritischen GCP-Bewertungen Einschränkungen unterliegen. Beim Vergleich der Einschätzung zu VOC gemäß ärztlichem Prüfpersonal gegenüber der adjudizierten VOC-Ereignisanzahl gemäß CRC zeigen sich zudem erhebliche deskriptive Unterschiede zwischen den beiden Einschätzungsvarianten, die in beiden Armen, jedoch v. a. im Crizanlizumab-Arm zu beobachten sind. Die vorliegenden Unterlagen zur Bewertung der Konkordanz sind allerdings beschränkt. Angaben im EPAR weisen auf eine 64,6%ige Konkordanz im aktiven Behandlungsarm gegenüber einer Konkordanz von 81,6 % im Kontrollarm hin und stützen somit diese Annahme [3]. Die Konkordanz im Crizanlizumab-Arm konnte gemäß EPAR durch den analysespezifischen Ausschluss einer Person auf 77,6 % gesteigert werden. Auswertungen zu den Adjudizierungsgründen liegen zur Nutzenbewertung nicht vor, sodass Unklarheiten verbleiben.

Der primäre Endpunkt „Jährliche Rate VOC“ sowie die Endpunkte „Jährliche Rate unkomplizierte VOC“ und „Jährliche Rate akute Thoraxsyndrome“ wurden gemäß Präspezifizierung primär

anhand des Wilcoxon-Rangsummen-Tests und mithilfe des Hodges-Lehmann-Schätzers untersucht. Die statistischen Annahmen über das Vorliegen ähnlicher Verteilungsformen der jeweiligen VOC beider Studienarme, die voneinander verschoben sein müssen, erscheint jedoch nicht erfüllt [3], weshalb eine eindeutige Interpretation der Ergebnisse hinsichtlich der Signifikanz des Unterschieds der Mediane und eine verlässliche Schätzung des Medians der Differenzen nicht möglich erscheint. Das heißt, dass bspw. die berichtete Signifikanz des Wilcoxon-Rangsummen-Tests tatsächlich eine Signifikanz der Unterschiede der Verteilungsformen und/oder eine fehlende Lageverschiebung repräsentieren könnten.

Weiterhin wurden die aufgeführten Endpunkte mittels einer Negativ-Binomial-Regression (Log-Link-Funktion) mit dem Logarithmus der Studiendauer als Offset-Variable und der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10), dem Einsatz einer begleitenden HU-Therapie (ja vs. nein) sowie der Behandlung als unabhängige Variablen analysiert. Diese wird als adäquatere Auswertung als die primäre Analyse eingeschätzt, auch wenn im Detail widersprüchliche Informationen zum Annualisierungsvorgehen vorliegen. Die Angaben lassen einmal auf eine simple Annualisierung analog zum Vorgehen beim primären Endpunkt schließen, während der Einbezug des Logarithmus der Studiendauer als Offset-Variable auf eine Art indirekte Annualisierung schließen lässt, indem die Anzahl von VOC anhand der Studiendauer relativiert wurde. Wie genau vorgegangen wurde bleibt demnach unklar. Bei beiden beschriebenen Vorgehensweisen wäre jedoch zu berücksichtigen, dass eine Annahme einer gleichbleibenden Ereignisrate von VOC nach Studienabbruch nicht nachvollziehbar erscheint. Die geschätzten JR VOC von Crizanlizumab- und Placebo-Arm (Rate Ratio: 0,65 (95%-KI: [0,47; 0,9];  $p = 0,008$ )) könnten somit in Verbindung mit der hohen Studienabbruchrate von etwas mehr als einem Drittel hochverzerrt sein. Es ist nicht auszuschließen, dass manche der Studienabbrüche im Zusammenhang mit der Auftretensrate von VOC (unabhängig von der Operationalisierung) der jeweiligen Person standen. Potentiell aufschlussgebende Vergleichsdaten zu Studienabbrechern und regelhaften Studienteilnehmenden lagen nicht vor. Konservativere Post-hoc-Sensitivitätsanalysen, welche verschiedene Imputationsmodelle anwenden, zeigen Effektschätzer in Richtung der Wirksamkeit von Crizanlizumab an, wobei die 95%-KI in zwei von drei Fällen jedoch die 1 einschließen.

In der Sensitivitätsanalyse PhS-M5 erfolgte eine Imputation fehlender Werte von Personen mit Therapieabbruch innerhalb der ersten 6 Monate durch die Anzahl der VOC vor der Randomisierung (Rate Ratio: 0,72 (95%-KI: [0,54; 0,95];  $p = 0,022$ )). Ergebnisangaben zu VOC in den einzelnen Behandlungsarmen fehlen unter diesem Modell. Bei der Interpretation ist u. a. zu berücksichtigen, dass die VOC vor Randomisierung anders erhoben wurden als VOC während des Studienverlaufs. Erstere konnten u. a. auf Erinnerungen der Patientinnen und Patienten und der Einschätzungen der jeweiligen ärztlichen Prüfperson basieren, wurden jedoch nicht vom CRC adjudiziert, wodurch eine Verzerrung nicht ausgeschlossen werden kann. Die vorliegenden Angaben deuten zudem daraufhin, dass die Anzahl der VOC von Personen mit Therapieabbruch innerhalb der ersten 6 Monate nicht allein imputiert, sondern ersetzt wurden [10], sodass eine etwaige höhere vorliegende Anzahl an VOC einer Person während der Studie vor Studienabbruch nicht berücksichtigt worden wäre und potentiell durch eine niedrigere Anzahl ersetzt wurde. Zudem werden fehlende Beobachtungszeiten bei einem Studienabbruch nach 6 Monaten in diesem Modell nicht berücksichtigt, da für Studienabbrecher nach diesem Zeitpunkt nicht imputiert wurde. Aufgrund der dargestellten Einschränkungen und Unklarheiten, kann eine eindeutige Interpretation der Ergebnisse nicht erfolgen.

Bei der Sensitivitätsanalyse PhS-M6a (Rate Ratio: 0,74 (95%-KI: [0,54; 1,03]; p = keine Angabe)) erfolgte hingegen eine Imputation der Anzahl der VOC bei vorzeitigem Behandlungsabbruch für Personen im Crizanlizumab-Arm über eine Jump-to-Reference-Methode basierend auf VOC von Personen aus dem Placebo-Arm. Inwiefern diesbezüglich eine multiple Imputation, wie vom pU in Modul 4 aufgeführt, stattgefunden hat und ob diese im klassischen Sinn (vgl. [24]) durchgeführt wurde, ist unklar, da dazu keine näheren Angaben (wie z. B. zum Multiplen Imputationsmodell, Anzahl der Imputationen bzw. Anzahl der Imputationsdatensätze) vorliegen. Für Personen im Placebo-Arm erfolgte die Ersetzung unter der Missing-at-Random-Annahme basierend auf Daten von Personen desselben Studienarms vor Behandlungsabbruch. Auch hierzu liegen keine näheren Angaben vor. Angaben zu VOC in den einzelnen Behandlungsarmen fehlen wie zu Modell PhS-M5 auch für Modell PhS-M6a. Im Gegensatz zu PhS-M5 wurden bei PhS-M6a jedoch alle Studienabbrüche berücksichtigt und es erfolgte eine Imputation basierend auf VOC, die innerhalb des Studienverlaufs erfasst wurden, weshalb die Ergebnisse unter Berücksichtigung der Unklarheiten im Vergleich zu PhS-M5 als belastbarer erscheinen.

In beiden Modellen (PhS-M5 und PhS-M6a) basieren die Auswertungen der VOC jedoch auf einer Einschätzung des CRC, welche gemäß EMA erhebliche Unklarheiten hinsichtlich der kontinuierlichen Anwendung einer Regel zum Zusammenfassen von VOC aufwies [3]. Die Sensitivitätsanalyse PhS-M6b (Rate Ratio: 0,74 (95%-KI: [0,52; 1,06]; p = keine Angabe)) nimmt demnach eine analoge Imputation zu PhS-M5b vor, jedoch basiert die VOC-Einschätzung auf dem eCRF. Zudem wurde eine Person im Crizanlizumab-Arm ausgeschlossen, die mit 37 VOC in 6 Monaten einen statistischen Ausreißer darstellt und die Einschlusskriterien nicht erfüllte, da sie bereits im Jahr vor der Randomisierung 30 VOC erlitten hatte. Während diese Sensitivitätsanalyse unter den gegebenen Einschränkungen der CRC-Einschätzung angemessener erscheint, bleibt jedoch unklar, warum nur eine einzige Person aus dem Crizanlizumab-Arm aus der Analyse ausgeschlossen wurde, während sowohl im Crizanlizumab-Arm mit einem Range von 1–13 VOC während der 12 Monate vor Randomisierung als auch im Placebo-Arm mit einem Range von 1–24 VOC während der 12 Monate vor Randomisierung weitere Fälle von Personen, welche die VOC-spezifischen Einschlusskriterien mit einem Range von 2–10 im Jahr vor Studieneinschluss nicht erfüllt haben, vorliegen müssen.

Bezüglich der Subgruppenanalysen ist festzuhalten, dass die Studie SUSTAIN nicht ausreichend hohe Fallzahlen aufzuweisen schien, insbesondere in Anbetracht der Studienabbrüche, um etwaige aussagekräftige Unterschiede herauszustellen (vgl. [3]).

Insgesamt liegen umfangreiche Unklarheiten hinsichtlich der VOC-Auswertungen vor und die Konfidenzintervalle der Effekte der als angemessener eingeschätzten Sensitivitätsanalysen PhS-M6a und PhS-M6b umschließen die 1. Jedoch ist ersichtlich, dass die Effektschätzer der Sensitivitätsanalysen konsistent in Richtung der Wirksamkeit von Crizanlizumab gegenüber Placebo zeigen (p-Werte fehlen bei PhS-M6a und PhS-M6b).

#### „Zeit bis zur ersten VOC“ und „Zeit bis zur zweiten VOC“

Time-to-Event-Analysen basieren auf der statistischen Annahme einer nicht-informativen Zensierung, die im vorliegenden Fall durch die hohe Rate von Studienabbrüchen sowie deren Hintergrund (bspw. „Entzug der Einverständniserklärung“ und „Sonstige Gründe“) in beiden Behandlungsarmen nicht gesichert erscheint. Auch liegen keine Angaben zu Beobachtungszeiten sowie nur sehr limitierte Informationen zu Zensierungszeitpunkten und -gründen vor. Letztere konnten nur im

Rahmen anderer Auswertungen indirekt abgeleitet werden und unterliegen demnach Unsicherheiten. Die Ergebnisse der diesbezüglichen Sensitivitätsanalyse mit Imputation PhS-M6b zu „Zeit bis zur ersten VOC“ unterliegt ähnlichen Unklarheiten zu jener des Endpunkts „Jährliche Rate VOC“, wobei vorliegend nur fehlende Informationen zur Imputation der ersten VOC relevant erscheinen. Dabei ist zu beachten, dass die Zusatzanalyse nicht im Rahmen des Dossiers eingereicht wurde und die Darstellung auf Angaben der EMA basierte [3]. Unter Berücksichtigung der umfassenden Unklarheiten und dem hohen Verzerrungspotential scheinen die Ergebnisse insgesamt eine Verlängerung der Zeit bis zur ersten VOC im Crizanlizumab-Arm (PhS-M6a: HR: 0,54 (95%-KI: [0,36; 0,81]; p = keine Angabe) zu stützen.

Informationen zum Endpunkt „Zeit bis zur zweiten VOC“ erscheinen noch limitierter als zu „Zeit bis zur ersten VOC“ und umfassen ebenso keine Ergebnisse einer Sensitivitätsanalyse, weshalb eine Interpretation der Endpunktergebnisse nicht möglich erscheint. Eine Argumentation des pU zur besonderen Bedeutung der Zeit bis zur zweiten VOC bspw. im Kontext anderer VOC konnte den Studienunterlagen zudem nicht entnommen werden.

#### Subklassifikationen: „Jährliche Rate unkomplizierte VOC“ und „Jährliche Rate akute Thoraxsyndrome“

Die primären Analysen der beiden Subklassifikationen sind bereits in Kapitel 2.3.2 (Abschnitt „Vasookklusive Schmerzkrisen“) diskutiert und erscheinen nicht interpretierbar.

Unkomplizierte VOC repräsentierten den Großteil aller erfassten VOC und variierten in der Häufigkeit über beide Arme hinweg je nach Einschätzungsvariante zwischen 89,1 % gemäß CRC und 92,95 % gemäß ärztlichem Prüfpersonal (Bezugspopulation hier Sicherheitspopulation, da ITT-Daten nicht vorliegend). Die Effektrichtung der Sensitivitätsanalyse PhS-M6b zur Wirksamkeit von Crizanlizumab liegt auch hinsichtlich der Subklassifikation „Jährliche Rate unkomplizierte VOC“ vor (Rate Ratio: 0,72 (95%-KI: [0,49; 1,05]; p-Wert: keine Angabe), wobei das Konfidenzintervall wie zur Sensitivitätsanalyse PhS-M6b aller VOC („Jährliche Rate VOC“) die 1 einschließt. Da sich alle VOC gemäß primärem Endpunkt zu etwa 90 % aus der Subklassifikation „Jährliche Rate unkomplizierte VOC“ zusammensetzen, erscheinen weitestgehend übereinstimmende Ergebnisse nachvollziehbar (vgl. PhS-M6b „Jährliche Rate VOC“: Rate Ratio: 0,74 (95%-KI: [0,52; 1,06]; p = keine Angabe). Einschränkungen zur Interpretation verhalten sich analog zu den Beschreibungen zur Analyse PhS-M6b aller VOC.

Deskriptiv waren akute Thoraxsyndrome anteilig an allen VOC etwas häufiger im Crizanlizumab-Arm als im Placebo-Arm zu beobachten, obwohl gemäß medizinischer Vorgeschichte in der Vergangenheit ACS fast doppelt so häufig im Placebo-Arm (21,5 %) als im Crizanlizumab-Arm (11,9 %) auftraten. Gegenüber der EMA berichtete der pU, dass 19,4 % der Patientinnen und Patienten im Crizanlizumab-Arm gegenüber 7,7 % im Placebo-Arm mindestens ein ACS im Jahr vor der Randomisierung aufwiesen und nach Baseline mit 14,9 % (10 von 67 Personen im Crizanlizumab-Arm) bzw. 16,9 % (11 von 65 Personen im Placebo-Arm) sich die Raten angeglichen hätten ([3]; Daten nicht mit Dossier eingereicht). Eine adäquate Analyse der akuten Thoraxsyndrome liegt nicht vor, zumal die Aussagekraft durch eine insgesamt geringe Ereignisrate begrenzt erscheint.

#### **Zusammenfassende Einschätzung zur Wirksamkeit**

Aufgrund umfangreicher Einschränkungen, die sich aus der Operationalisierung, der vielen Studienabbrüche, der fehlenden Werte und Spezifizierungen sowie dem Analyseverfahren und potentieller ergebnisgesteuerter Berichterstattung ergeben, unterliegen die Ergebnisse zur

Morbidität einem hohen Verzerrungspotential. Die als angemessener beurteilten Sensitivitätsanalysen mit konservativen Vorgehensweisen PhS-M6a und PhS-M6b zeigen eine gleichbleibende Effektrichtung an, deren 95%-KI jedoch jeweils die 1 umschließen.

## 4.5 Lebensqualität

Ergebnisse zur Lebensqualität wurden nur ergänzend dargestellt (siehe Anhang).

## 4.6 Sicherheit

Im Rahmen der Sicherheitsauswertung ergaben sich statistisch signifikante Ergebnisse im Rahmen der post hoc erfolgten Analysen ausschließlich zur Systemorganklasse „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ als ein UE mit Inzidenz  $\geq 10\%$  und zuungunsten von Crizanlizumab (RR: 4,70 (95%-KI: [1,07; 20,59];  $p = 0,022$ )). Diesbezügliche Preferred Terms weisen eine hohe Anzahl mit geringen betroffenen Personenmengen auf. Bei der Interpretation der Ergebnisse ist zu berücksichtigen, dass keine Angaben über die (mediane) Beobachtungsdauer der Patientinnen und Patienten vorlagen sowie aufgrund einer hohen Studienabbruchrate von etwas mehr als einem Drittel der Patientinnen und Patienten ein etwaiger Zusammenhang mit der (Wirksamkeit und/oder) Sicherheit in manchen Fällen (bspw. bei „Sonstige Gründe“ oder „Entzug der Einverständniserklärung“) nicht vollständig ausgeschlossen werden kann. Es lagen keine Informationen zu den Charakteristika (wie bspw. prognostische Faktoren, VOC, UE) der Studienabbrecher vor, um eine nähere Einschätzung vorzunehmen. Fehlende Angaben zur Beobachtungsdauer sind im Vergleich zu den Wirksamkeitsanalysen als (noch) limitierender zu bewerten, da die vorliegenden Informationen zur Behandlungsdauer aufgrund der regelhaften Nachbeobachtungszeit von 8 Wochen nach der letzten Behandlung im Rahmen der Sicherheitsauswertung einen weniger präzisen Annäherungswert liefern. Weiterhin seien gemäß EMA in mehreren Fällen Patientinnen und Patienten nach Studienmedikationsgabe nicht für ein präspezifizierte Zeit von einer Stunde nachbeobachtet worden und die Berichterstattung zu UE wurde nicht immer fristgerecht eingehalten [3]. Über den Umfang dieser Protokollverletzungen liegen keine Auswertungen vor.

Wie die EMA ausführt [3], ist die Gesamtrate an UE in der Studie SUSTAIN hoch, jedoch wie die Rate schwerer UE (RR: 0,94 (95%-KI: [0,46; 1,93];  $p = 0,858$ )) und SUE (RR: 0,94 (95%-KI: [0,53; 1,67];  $p = 0,82$ )) ausgeglichen zwischen den beiden Studienarmen. In der laufenden, unkontrollierten Studie SOLACE-adults mit zwei Dosen Crizanlizumab sei die Rate schwerer UE mit 40 % höher. Auffallend sei, dass insgesamt von SUSTAIN zu SOLACE-adults eine Zunahme der Schwere von UE zu beobachten sei, welche jedoch auf unterschiedliche Gradierungssysteme zurückgehen könnte. Diesbezüglich ist innerhalb der SUSTAIN-Studie auffallend, dass mehr SUE als schwere UE in beiden Studienarmen berichtet wurden, was auf die limitierte Reliabilität der Schweregradierung in leicht (keine Beeinträchtigung von Alltagsaktivitäten), moderat (Beeinträchtigung von Alltagsaktivitäten) und schwer (Verhinderung von Alltagsaktivitäten) zurückzuführen sein könnte. Insgesamt scheinen nach Angabe der EMA die vorläufigen Sicherheitsergebnisse der Studie SOLACE-adults jene der SUSTAIN-Studie zu unterstützen, wobei verbleibende Unklarheiten durch die laufende Phase-III-Studie STAND adressiert werden sollen [3].

Das Sicherheitsprofil sei gemäß EMA insbesondere für den langfristigen Gebrauch noch nicht ausreichend spezifizierbar und Ergebnisse der STAND-Studie sollen diesbezüglich Aufklärung

bringen [3]. Auch zum Sicherheitsprofil fehlten für schwangere und stillende Frauen noch Informationen und nur zu 6 Personen unter 18 Jahren lagen in allen von der EMA berücksichtigten Studien Daten vor. Ergebnisse zu älteren Personen werden gemäß EMA jedoch als auf die zuletzt genannte Patientengruppe übertragbar erachtet. Infusionsbedingte Reaktionen (gemäß UE von besonderem Interesse) seien in der Studie SUSTAIN häufiger bei Crizanlizumab (34,8 %) gegenüber Placebo-Behandlung (21 %) zu beobachten und stellen gemäß EMA ein wichtiges identifiziertes Risiko dar. Der Unterschied zwischen den Behandlungsarmen war in der Post-hoc-Analyse im Dossier jedoch nicht signifikant (RR: 1,66 (95%-KI: [0,93; 2,98];  $p = 0,084$ )). Kritische Sicherheitsergebnisse seien zwar in den klinischen Studien nicht vorgefunden worden, diese Sachlage sei jedoch aufgrund der limitierten Datenlage nicht verlässlich, welches die besondere Bedeutung der Ergebnisse der Studie STAND und die Bedeutung der finalen Ergebnisse der Studie SOLACE-adults für das Sicherheitsprofil begründet. Diese Daten werden seitens der EMA bis Ende 2025 vom pU im Kontext der CMA gefordert.

### **Zusammenfassende Einschätzung zur Sicherheit**

Die vorliegenden Sicherheitsdaten deuten unter Limitationen und einem hohen Verzerrungspotential auf ein weitestgehend ausgeglichenes Sicherheitsprofil von Crizanlizumab gegenüber Placebo hin. Es ist insbesondere zu berücksichtigen, dass die Sicherheitsdaten u. a. aufgrund ihres Umfangs noch vorläufiger Art erscheinen und künftige Ergebnisse der Studie STAND verlässlichere Aussagen zum Sicherheitsprofil von Crizanlizumab ermöglichen könnten.

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

Die Angaben der Fachinformation sind zu berücksichtigen. Die Einleitung und Überwachung der Therapie mit Crizanlizumab soll durch in der Therapie von Patientinnen und Patienten mit Sichelzellerkrankung erfahrene Ärztinnen und Ärzte erfolgen.

## 6 Zusammenfassung der Nutzenbewertung

Crizanlizumab (Adakveo®) ist zugelassen zur Prävention wiederkehrender vasookklusiver Krisen bei Patienten ab 16 Jahren mit Sichelzellerkrankung. Die Nutzenbewertung von Crizanlizumab basiert auf der zulassungsbegründenden Studie SUSTAIN. Es handelt sich bei der Studie um eine randomisierte, doppelblinde, placebokontrollierte, multizentrische Phase-II-Studie.

Die Ergebnisse der Studie werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Crizanlizumab 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 25: Zusammenfassende Darstellung der Ergebnisse der Studie SUSTAIN

Endpunkt	Crizanlizumab 5,0 mg/kg	Placebo	Effekt
<b>Mortalität</b>			
Mortalität <sup>1)</sup>	N = 66	N = 62	n. b.
Todesfälle, n (%)	2 (3) <sup>2)</sup>	2 (3,2) <sup>2)</sup>	
<b>Morbidität</b>			
<b>Vasookklusive Schmerzkrise</b>			
<i>Jährliche Rate VOC – Sensitivitätsanalysen<sup>3)</sup></i>			
PS-1: VOC gemäß CRC	N = 67	N = 65	↑
Rate Ratio [95%-KI]; p-Wert	0,65 [0,47; 0,9]; 0,008 <sup>4)</sup>		
PhS-M6a: VOC gemäß CRC <sup>5)</sup>	N = 67	N = 65	↔
Rate Ratio [95%-KI]; p-Wert	0,74 [0,54; 1,03] <sup>4)</sup> ; k. A.		
PhS-M6b: VOC gemäß ärztlichem Prüfpersonal <sup>6)</sup>	N = 66 <sup>7)</sup>	N = 65	↔
Rate Ratio [95%-KI]; p-Wert	0,74 [0,52; 1,06] <sup>4)</sup> ; k. A.		
<i>Zeit bis zur ersten VOC – Sensitivitätsanalyse<sup>8)9)</sup></i>			
PhS-M6b: VOC gemäß ärztlichem Prüfpersonal	N = 66	N = 65	↑
Hazard Ratio [95%-KI]; p-Wert	0,54 [0,36; 0,81] <sup>10)</sup> ; k. A.		
<b>Sicherheit</b>			
<i>Schwere UE<sup>11)</sup></i>			
Patientinnen und Patienten mit ≥ 1 Ereignis, n (%)	12 (18,2)	12 (19,4)	↔
Relatives Risiko [95%-KI]; p-Wert <sup>12)</sup>	0,94 [0,46; 1,93]; 0,858 <sup>13)</sup>		
<i>SUE</i>			
Patientinnen und Patienten mit ≥ 1 Ereignis, n (%)	17 (25,8)	17 (27,4)	↔
Relatives Risiko [95%-KI]; p-Wert <sup>12)</sup>	0,94 [0,53; 1,67]; 0,82 <sup>13)</sup>		
<i>UE, das zum Therapieabbruch führte</i>			
Patientinnen und Patienten mit ≥ 1 Ereignis, n (%)	2 (3,0)	3 (4,8)	↔
Relatives Risiko [95%-KI]; p-Wert <sup>12)</sup>	0,63 [0,11; 3,62]; 0,595 <sup>13)</sup>		

- <sup>1)</sup> Todesfälle wurden im Rahmen der Sicherheitserhebung erfasst.
- <sup>2)</sup> Aufgrund der hinsichtlich der Indikation geringen Studiendauer und der vielen Studienabbrüche von etwas mehr als einem Drittel in beiden Studienarmen lassen die in der Studie SUSTAIN beobachteten Ergebnisse keine validen Aussagen zum Effekt von Crizanlizumab auf das Gesamtüberleben zu.
- <sup>3)</sup> Ergebnisse der präspezifizierten primären Analyse des Endpunkts „Jährliche Rate VOC“ sind basierend auf einer fehlenden eindeutigen Interpretierbarkeit der Analysemethode nicht dargestellt. Aufgeführt sind die Ergebnisse der präspezifizierten Sensitivitätsanalyse (PS-1) und der Post-hoc-Sensitivitätsanalysen PhS-M6a und PhS-M6b, welche alle mittels Negativ-Binomial-Regressionsanalyse ausgewertet wurden.
- <sup>4)</sup> Ratenverhältnisse, 95%-KI sowie p-Wert geschätzt mithilfe eines Negativ-Binomial-Regressionsmodells mit dem Logarithmus der Studiendauer als Offset-Variable und der Anzahl VOC im vorherigen Jahr, dem Einsatz einer begleitenden HU-Therapie sowie der Behandlung als unabhängige Variablen.
- <sup>5)</sup> Eine Imputation der Anzahl der VOC bei vorzeitigem Behandlungsabbruch für Personen im Crizanlizumab-Arm über eine Jump-to-Reference-Methode basierend auf VOC von Personen aus dem Placebo-Arm. Für Personen im Placebo-Arm erfolgte die Ersetzung unter der Missing-at-Random-Annahme basierend auf Daten von Personen desselben Studienarms vor Behandlungsabbruch.
- <sup>6)</sup> Das Vorgehen entspricht methodisch PhS-M6a, erfolgt aber basierend auf Daten zu VOC gemäß ärztlichem Prüfpersonal und unter Ausschluss einer Person im Crizanlizumab-Arm.
- <sup>7)</sup> Bei der von der Analyse ausgeschlossenen Person handelt es sich um einen statistischen Ausreißer mit 37 VOC in 6 Monaten, die zudem die Einschlusskriterien nicht erfüllte, da sie bereits im Jahr vor der Randomisierung 30 VOC erlitten hatte.
- <sup>8)</sup> Diese Analyse wurde nicht mit dem Dossier vorgelegt. Die vorliegenden Angaben stammen aus dem EPAR [3].
- <sup>9)</sup> Aufgrund umfangreicher Unklarheiten hinsichtlich der primären Analyse erfolgt ausschließlich eine Darstellung der Post-hoc-Sensitivitätsanalyse PhS-M6b.
- <sup>10)</sup> Schätzung des Hazard Ratios mithilfe eines Cox-Regressionsmodells mit der Anzahl der VOC im vorherigen Jahr, dem Einsatz einer begleitenden HU-Therapie und der Behandlung als Kovariaten.
- <sup>11)</sup> Die Bewertung des Schweregrads erfolgte nach den Einstufungen leicht (keine Beeinträchtigung von Alltagsaktivitäten), moderat (Beeinträchtigung von Alltagsaktivitäten) und schwer (Verhinderung von Alltagsaktivitäten).
- <sup>12)</sup> Effektschätzer und p-Werte wurden post hoc berechnet.
- <sup>13)</sup> Berechnung des p-Werts mithilfe eines Cochran-Mantel-Haenszel-Tests stratifiziert nach der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10) und dem Einsatz einer begleitenden HU-Therapie (ja vs. nein).

Abkürzungen: CRC: Crisis Review Committee; EPAR: European Public Assessment Report; HU: Hydroxyurea/Hydroxycarbamid; k. A.: keine Angabe; KI: Konfidenzintervall; M: Imputationsmethode; n. b.: nicht berechnet oder nicht berechenbar; PhS: Post-hoc-Sensitivitätsanalyse; PS: Präspezifizierte Sensitivitätsanalyse; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis; VOC: Vasookklusive Schmerzkrise.

## Referenzen

1. **Cleeland CS, MD Anderson Cancer Cencter.** The Brief Pain Inventory (BPI); User Guide [online]. Houston (USA): University of Texas; 1991. [Zugriff: 16.02.2021]. URL: [https://www.mdanderson.org/documents/Departments-and-Divisions/Symptom-Research/BPI\\_UserGuide.pdf](https://www.mdanderson.org/documents/Departments-and-Divisions/Symptom-Research/BPI_UserGuide.pdf).
2. **Daut RL, Cleeland CS, Flanery RC.** Development of the Wisconsin Brief Pain Questionnaire to assess pain in cancer and other diseases. Pain 1983;17(2):197-210.
3. **European Medicines Agency (EMA).** Adakveo: European public assessment report EMEA/H/C/004874/0000 [online]. 24.11.2020. Amsterdam (NED): EMA. [Zugriff: 08.12.2020]. URL: [https://www.ema.europa.eu/documents/assessment-report/adakveo-epar-public-assessment-report\\_en.pdf](https://www.ema.europa.eu/documents/assessment-report/adakveo-epar-public-assessment-report_en.pdf).
4. **Gesellschaft für Pädiatrische Onkologie und Hämatologie (GPOH).** Sichelzellkrankheit; S2k-Leitlinie, 2. Auflage [online]. AWMF-Registernummer 025-016. Berlin (GER): AWMF; 02.07.2020. [Zugriff: 16.02.2021]. URL: [https://www.sichelzellkrankheit.info/app/download/9999280585/AWMF-Leitlinie+SCD+V2\\_0.pdf?t=1594220957](https://www.sichelzellkrankheit.info/app/download/9999280585/AWMF-Leitlinie+SCD+V2_0.pdf?t=1594220957).
5. **Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG).** Allgemeine Methoden; Vers. 6.0 [online]. 05.11.2020. Köln (GER): IQWiG; 2017. [Zugriff: 16.02.2021]. URL: <https://www.iqwig.de/methoden/allgemeine-methoden-version-6-0.pdf?rev=144030>.
6. **Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG).** Enzalutamid (neues Anwendungsgebiet) - Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung [online]. IQWiG-Berichte - Nr. 291. Köln (GER): IQWiG; 2015. [Zugriff: 16.02.2021]. (IQWiG-Berichte). URL: [https://www.g-ba.de/downloads/92-975-750/Nutzenbewertung%20IQWiG\\_Enzalutamid\\_nAWG.pdf](https://www.g-ba.de/downloads/92-975-750/Nutzenbewertung%20IQWiG_Enzalutamid_nAWG.pdf).
7. **Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG).** Filgotinib (rheumatoide Arthritis) - Nutzenbewertung gemäß § 35a SGB V; Dossierbewertung; Auftrag A20-90 [online]. Köln (GER): IQWiG; 2021. [Zugriff: 16.02.2021]. (IQWiG-Berichte; Band 1023). URL: [https://www.iqwig.de/download/G14-01\\_Macitentan\\_Bewertung-35a-Abs1-Satz10-SGB-V.pdf](https://www.iqwig.de/download/G14-01_Macitentan_Bewertung-35a-Abs1-Satz10-SGB-V.pdf).
8. **Maruish M.** User's Manual for the SF-36v2 Health Survey; 3rd. Edition. Lincoln (USA): Quality Metric; 2011.
9. **Novartis Pharma.** Adakveo; Fachinformation, Stand: Oktober 2020 [online]. 2020. [Zugriff: 16.02.2021]. URL: [www.fachinfo.de](http://www.fachinfo.de).
10. **Novartis Pharma.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4A: Adakveo (crizanlizumab): Sichelzellkrankheit [unveröffentlicht]. 25.11.2020.

11. **Novartis Pharma.** A phase 2, multicenter, open-label study to assess appropriate dosing and to evaluate safety of crizanlizumab, with or without hydroxyurea/ hydroxycarbamide, in sequential, descending age groups of pediatric sickle cell disease patients with vaso-occlusive crisis (SUSTAIN); CSEG101B2201. Oncology Clinical Protocol [online]. In: ClinicalTrials.gov. [Zugriff: 16.02.2021]. URL: <https://clinicaltrials.gov/ct2/show/record/NCT03938454>.
12. **Novartis Pharma.** A phase 2, multicenter, open-label study to assess appropriate dosing and to evaluate safety of crizanlizumab, with or without hydroxyurea/ hydroxycarbamide, in sequential, descending age groups of pediatric sickle cell disease patients with vaso-occlusive crisis (SUSTAIN); CSEG101B2201. Oncology Clinical Protocol, Vers. 03 [online]. In: ClinicalTrials.gov. 22.10.2020. [Zugriff: 16.02.2021]. URL: <https://clinicaltrials.gov/ct2/show/record/NCT03938454>.
13. **Novartis Pharma.** A Phase 2, multicenter, open-label study to assess PK/PD of SEG101 (crizanlizumab), with or without hydroxyurea/hydroxycarbamide, in sickle cell patients with vaso-occlusive crisis; CSEG101A2202. Clinical study report for interim analysis [unveröffentlicht]. 25.02.2019.
14. **Novartis Pharma.** A phase 2, multicenter, open-label study to assess appropriate dosing and to evaluate safety of crizanlizumab, with or without hydroxyurea/ hydroxycarbamide, in sequential, descending age groups of pediatric sickle cell disease patients with vaso-occlusive crisis. (SOLACE-kids); CSEG101B2201. Clinical Trial Protocol [unveröffentlicht]. 28.07.2020.
15. **Novartis Pharma.** A Phase I single dose, randomized, open-label, parallel group, single-center trial to assess the comparability of crizanlizumab pharmacokinetics and pharmacodynamics administered as SEG101 in comparison to SelG1 in healthy subjects; CSEG101A2102. Clinical study report [unveröffentlicht]. 28.06.2018.
16. **Novartis Pharma.** A Phase II multicenter, randomized, placebo-controlled, double-blind, 12-month study to assess safety and efficacy of SelG1 with or without hydroxy urea therapy in sickle cell disease patients with sickle cell-related pain crises; SelG1-00005 (SUSTAIN). Zusatzanalysen [unveröffentlicht]. 2020.
17. **Novartis Pharma.** A phase II, multicenter, randomized, open label two arm study comparing the effect of crizanlizumab + standard of care to standard of care alone on renal function in sickle cell disease patients  $\geq 16$  years with chronic kidney disease due to sickle cell nephropathy (STEADFAST); CSEG101A2203 [online]. In: ClinicalTrials.gov. 19.10.2020. [Zugriff: 16.02.2021]. URL: <https://clinicaltrials.gov/ct2/show/record/NCT04053764>.
18. **Novartis Pharma.** A phase II, multicenter, randomized, placebo-controlled, double-blind, 12-month study to assess safety and efficacy of SELG1 with or without hydroxy urea therapy in sickle cell disease patients with sickle cell-related pain crises. Statistical analysis plan [unveröffentlicht]. 02.05.2016.

19. **Novartis Pharma.** A phase III, multicenter, randomized, double-blind study to assess efficacy and safety of two doses of crizanlizumab versus placebo, with or without Hydroxyurea/Hydroxycarbamide Therapy, in adolescent and adult sickle cell disease patients with vaso-occlusive crises (STAND); CSEG101A2301. Clinical Trial Protocol, Amended Protocol, Version 03 [unveröffentlicht]. 28.07.2020.
20. **Novartis Pharma.** A prospective phase II, open-label, single-arm, multicenter, study to assess efficacy and safety of SEG101 (Crizanlizumab), in sickle cell disease patients with priapism (SPARTAN); CSEG101AUS05 [online]. In: ClinicalTrials.gov. 22.10.2020. [Zugriff: 16.02.2021]. URL: <https://clinicaltrials.gov/ct2/show/record/NCT03938454>.
21. **Novartis Pharma.** Study of two doses of crizanlizumab versus placebo in adolescent and adult sickle cell disease patients (STAND); CSEG101A2203 [online]. In: ClinicalTrials.gov. 11.02.2021. [Zugriff: 16.02.2021]. URL: <https://clinicaltrials.gov/ct2/show/NCT03814746>.
22. **Poquet N, Lin C.** The Brief Pain Inventory (BPI). J Physiother 2016;62(1):52.
23. **PTJA10 Authoring Team.** Crizanlizumab for the prevention of recurrent vaso-occlusive crises in sickle cell disease patients aged 16 years and older. Joint Assessment Report [online]. Report No.: PTJA10. Diemen (NED): EUnetHTA; 2020. [Zugriff: 16.02.2021]. (EUnetHTA Joint Action 3 WP4). URL: <https://www.eunetha.eu>.
24. **Rubin D.** Multiple imputation for nonresponse in surveys. Wiley Series in Probability and Statistics ed: Wiley Classics Library; 2008.
25. **Selexys Pharmaceuticals.** A phase II multicenter, randomized, placebo-controlled, double-blind, 12-month study to assess safety and efficacy of selg1 with or without hydroxyurea therapy in sickle cell disease patients with sickle cell-related pain crises; (SUSTAIN); SelG1-00005. Clinical study protocol, Amendment 4. 21.04.2016.
26. **Selexys Pharmaceuticals.** A Phase II multicenter, randomized, placebo-controlled, double-blind, 12-month study to assess safety and efficacy of selg1 with or without hydroxyurea therapy in sickle cell disease patients with sickle cell-related pain crises; SelG1-00005 (SUSTAIN). Integrated clinical and statistical report. 16.10.2016.
27. **Selexys Pharmaceuticals.** Placebo-controlled, double-blind, first-in-human, ascending single dose and multiple dose study of intravenous-administered SelG1 in healthy subjects; SelG1-00003. Clinical study report [unveröffentlicht]. 16.05.2012.

## Anhang

### Übersicht zu Sensitivitätsanalysen des Endpunkts VOC

Tabelle 26: Übersicht zu Sensitivitätsanalysen des primären Endpunkts

<b>Sensitivitätsanalysen des primären Endpunkts</b>	
<b>Präspezifizierte Sensitivitätsanalysen (gemäß Protokoll 3)</b>	
1	Auswertung basierend auf: a) Per Protocol <sup>1)2)3)4)5)</sup> b) Einer Population mit ausschließlich jenen Personen, die mindestens 6 Behandlungsmonate abgeschlossen haben <sup>1)2)3)4)5)</sup>
2	Auswertung mithilfe eines Negativ-Binomial-Regressionsmodells (Log-Link-Funktion) mit dem Logarithmus der Studiendauer als Offset-Variable und der Anzahl der VOC im vorherigen Jahr (2–4 vs. 5–10), dem Einsatz einer begleitenden HU-Therapie (ja vs. nein) sowie der Behandlung als unabhängige Variablen <sup>1)2)3)4)</sup>
M1	Imputationen fehlender Daten für alle Personen, die nicht mindestens 6 Behandlungsmonate abgeschlossen haben, unter Verwendung der vorliegenden VOC-Daten der jeweiligen Person vor Randomisierung <sup>1)</sup>
M2	Multiple Imputation basierend auf Studienabbruchgründen, welche zweifach kategorisiert werden (aufgrund des Fehlens eines Behandlungseffektes vs. der restlichen Dropouts) <sup>1)</sup>
3	Explorative Untersuchung des Einflusses einer HU-Therapieinitiation oder -intensivierung während der Studie auf VOC <sup>1)</sup>
4	Explorative Analyse des Dosis-Responsezusammenhangs <sup>1)2)3)4)</sup>
<b>Zusätzliche nach Studienende spezifizierte Sensitivitätsanalysen</b>	
1	Auswertung basierend auf der mITT-Population <sup>2)3)4)5)</sup>
2	Wiederholung der primären Analyse in Form eines stratifizierten Rangsummentests basierend auf CRF-Daten über HU-Therapie und kategorisierter VOC-Vorgeschichte anstelle der Daten aus dem IXRS-System <sup>2)3)4)</sup>
<b>Post-hoc-Sensitivitätsanalysen</b>	
1	Durchführung wie die primäre Analyse „Jährliche Rate VOC“, jedoch gemäß der Bewertung durch das ärztliche Prüfpersonal (statt mit den CRC adjudizierten Krisenereignissen) unter Ausschluss einer Person im Crizanlizumab-Arm <sup>5)7)</sup>
<u>Imputationen mit Auswertungen basierend auf der primären Analyse (HL-Schätzer)<sup>6)</sup></u>	
M1	„Jährliche Rate VOC“ von Therapieabrechern wird durch die mittlere jährliche Rate von Personen derselben Behandlungsgruppe und desselben Stratum ersetzt, welche die Behandlung wie vorgesehen beendet hatten, insofern diese größer war als die für die jeweilige Person berechnete Rate (ITT-Population) <sup>5)</sup>
M2	Wie (PhS-)M1 unter Ausschluss von Personen, die die Behandlung vor Auftreten der ersten Schmerzkrisis unter der Studienbehandlung abgebrochen hatten <sup>5)</sup>
<u>Imputationen mit Auswertungen basierend auf der Negativ-Binomial-Regression (ITT-Population)<sup>6)</sup></u>	
M3	Imputation der Anzahl der VOC für Personen, welche die Behandlung vor Auftreten der ersten VOC unter Studienbehandlung abgebrochen hatten, mittels Multipler Imputation basierend auf der Missing-at-Random-Annahme unter Berücksichtigung der Behandlungsgruppe, der Behandlung mit HU, der Anzahl der VOC im vorherigen Jahr und der Dauer der Studienbehandlung. <sup>5)</sup>
M4	Imputation der Anzahl der VOC bei vorzeitigem Behandlungsabbruch für Personen im Crizanlizumab-Arm, die die Behandlung aufgrund von möglicherweise im Zusammenhang mit der Behandlung stehenden Gründen abgebrochen hatten („UE“, „Tod“ oder „Fehlende Wirksamkeit“), über eine Jump-to-Reference-Methode basierend auf Daten von Personen aus dem Placebo-Arm. Für Personen im Crizanlizumab-Arm, welche die Behandlung aufgrund von vermutlich nicht im Zusammenhang mit der Behandlung stehenden Gründen abgebrochen hatten („Lost to Follow-up“, „Fehlende Compliance“, „Entscheidung des Arztes bzw. Patienten“ oder „Sonstige Gründe“), sowie für Personen im Placebo-Arm erfolgte die Ersetzung unter der Missing-at-Random-Annahme basierend auf Daten von Personen desselben Studienarmes vor Behandlungsabbruch <sup>5)</sup>

<b>Sensitivitätsanalysen des primären Endpunkts</b>	
M5	Imputation der Werte von Personen mit Therapieabbruch innerhalb der ersten 6 Monate durch die Anzahl der Krisenereignisse vor der Randomisierung <sup>5)</sup>
M6a	Imputation der Anzahl der VOC bei vorzeitigem Behandlungsabbruch für Personen im Crizanlizumab-Arm über eine Jump-to-Reference-Methode basierend auf Daten von Personen aus dem Placebo-Arm. Für Personen im Placebo-Arm erfolgte die Ersetzung unter der Missing-at-Random-Annahme basierend auf Daten von Personen desselben Studienarmes vor Behandlungsabbruch <sup>5)</sup>
M6b	Wie M6a, aber basierend auf der Bewertung durch das ärztliche Prüfpersonal und unter Ausschluss einer Person im Crizanlizumab-Arm <sup>5)7)</sup>

<sup>1)</sup> Konnte in Protokoll 3 identifiziert werden. [19]

<sup>2)</sup> Konnte in Protokoll 4 identifiziert werden. [25]

<sup>3)</sup> Konnte im SAP identifiziert werden. [18]

<sup>4)</sup> Konnte im Studienbericht identifiziert werden. [26]

<sup>5)</sup> Konnte im Dossier / in Modul 4 des Dossiers identifiziert werden. [10]

<sup>6)</sup> Die dargestellten Imputationen wurden der EMA im Rahmen des Zulassungsprozesses vorgelegt. [3]

<sup>7)</sup> Bei der von der Analyse ausgeschlossenen Person handelt es sich um einen statistischen Ausreißer mit 37 VOC in 6 Monaten, die zudem die Einschlusskriterien nicht erfüllte, da sie bereits im Jahr vor der Randomisierung 30 VOC erlitten hatte.

Abkürzungen: CRC: Crisis Review Committee; CRF: Case Report Form; EMA: European Medicines Agency; HU: Hydroxyurea/ Hydroxycarbamid; ITT: Intention-to-Treat; IXRS: Interactive Voice/Web Response System; M: Imputationsmethode; mITT: modifizierte Intention-to-Treat; PhS: Post-hoc-Sensitivitätsanalysen; SAP: Statistischer Analyseplan; UE: Unerwünschte Ereignis; VOC: Vasookklusive Schmerzkrise.

## Ergebnisse: Lebensqualität – SF-36

Daten zur gesundheitsbezogenen Lebensqualität wurden mithilfe des SF-36 erhoben.

Tabelle 27: Rücklaufquoten des SF-36 in der Studie SUSTAIN – ITT-Population

Zeitpunkt	Crizanlizumab 5,0 mg/kg N = 67		Placebo N = 65	
	Zahl der ausgefüllten Fragebogen (n) <sup>1)</sup>	Rücklauf- quote (%)	Zahl der ausgefüllten Fragebogen (n) <sup>1)</sup>	Rücklauf- quote (%)
Baseline	63	94	65	100
Tag 15	61	91	59	90,8
Woche 14	55	82,1	48	73,8
Woche 26	48	71,6	46	70,8
Woche 38	46	68,7	42	64,6
Woche 52	36	53,7	34	52,3
Follow-up-Visite <sup>2)</sup>	47	70,1	46	70,8

<sup>1)</sup> Die Angaben beziehen sich jeweils auf die Teilnehmenden, die in die Auswertung der beiden Summenscores PCS und MCS eingingen.

<sup>2)</sup> Die Angaben umfassen jeweils die Teilnehmenden, welche entweder an der regulären Follow-up-Visite in Woche 58 oder der Follow-up-Visite mindestens 6 Wochen nach Behandlungsabbruch teilgenommen haben.

Abkürzungen: ITT: Intention-to-Treat; SF-36: 36-item Short Form Health Survey.

Die anfangs hohen Rücklaufquoten fallen zu den Messzeitpunkten Woche 38 und Woche 52 unter 70 %. Die Ergebnisse werden zu diesen beiden Zeitpunkten nicht dargestellt. Da es sich bei den Angaben zur Follow-up-Visite um Werte handelt, die zu unterschiedlichen Zeitpunkten, d. h. regelhaft in Woche 58 oder 6 Wochen nach Studienabbruch, erhoben wurden und der Anteil der Studienabbrüche in beiden Studienarmen über einem Drittel liegt, muss dieses bei der Interpretation der Ergebnisse berücksichtigt werden.

Im Folgenden werden die Ergebnisse der MMRM-Analyse für die beiden Summenscores in Tabelle 28 sowie für die Subskalen in Tabelle 29 ergänzend dargestellt.

Tabelle 28: Ergebnisse zu den Summenscores des SF-36 in der Studie SUSTAIN – ITT-Population

Studie SUSTAIN	Crizanlizumab 5,0 mg/kg N = 67	Placebo N = 65	Mittelwertdifferenz <sup>1)</sup> [95%-KI], p-Wert
<b>SF-36 – Körperlicher Summenscore (PCS)<sup>2)</sup></b>			
<i>Behandlungsbeginn</i> n (%) <sup>3)</sup> MW (SD)	63 (94) 39,37 (10,11)	65 (100) 40,19 (8,67)	--
<i>Veränderung Behandlungsbeginn zu Tag 15</i> n (%) <sup>3)</sup> MW (SD)	59 (88,1) 0,14 (4,98)	59 (90,8) -1,46 (6,5)	1,73 [-0,3; 3,76]; 0,094
<i>Veränderung Behandlungsbeginn zu Woche 14</i> n (%) <sup>3)</sup> MW (SD)	53 (79,1) -1,17 (5,98)	48 (73,8) -1,09 (7,32)	-0,27 [-3,14; 2,59]; 0,852
<i>Veränderung Behandlungsbeginn zu Woche 26</i> n (%) <sup>3)</sup> MW (SD)	46 (68,7) 0,74 (7,88)	46 (70,8) 0,26 (7,46)	0,51 [-2,33; 3,35]; 0,723
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>4)</sup></i> n (%) <sup>3)</sup> MW (SD)	47 (70,2) 1,06 (8,38)	46 (70,8) 0,43 (8,26)	0,67 [-2,55; 3,90]; 0,680
<b>SF-36 – Psychischer Summenscore (MCS)<sup>2)</sup></b>			
<i>Behandlungsbeginn</i> n (%) <sup>3)</sup> MW (SD)	63 (94) 46,46 (11,1)	65 (100) 43,94 (14,81)	--
<i>Veränderung Behandlungsbeginn zu Tag 15</i> n (%) <sup>3)</sup> MW (SD)	59 (88,1) 2,21 (9,1)	59 (90,8) 2,63 (11,08)	0,49 [-2,59; 3,57]; 0,754
<i>Veränderung Behandlungsbeginn zu Woche 14</i> n (%) <sup>3)</sup> MW (SD)	53 (79,1) 1,23 (12,15)	48 (73,8) 2,28 (9,88)	-1,32 [-5,27; 2,63]; 0,511
<i>Veränderung Behandlungsbeginn zu Woche 26</i> n (%) <sup>3)</sup> MW (SD)	46 (68,7) 1,27 (12,41)	46 (70,8) 2,44 (11,87)	-0,83 [-5; 3,34]; 0,694
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>4)</sup></i> n (%) <sup>3)</sup> MW (SD)	47 (70,2) -0,425 (10,47)	46 (70,8) 1 (11,36)	-1,22 [-5,28; 2,85]; 0,556

<sup>1)</sup> MMRM-Analyse für die Veränderung von Baseline bis zur Follow-up-Visite. Mittelwertdifferenz bezieht sich auf Kleinste-Quadrate-Mittelwertdifferenz des Modells.

<sup>2)</sup> Skala: 0–100. Höhere Werte auf den Summenskalen entsprechen einer besseren gesundheitsbezogenen Lebensqualität.

<sup>3)</sup> Anzahl der Patientinnen und Patienten in der Auswertung (Rücklaufquote des Fragebogens) bezogen auf die ITT-Population. Bei Veränderungsmessungen sind bezüglich MW und SD alle Personen mit Baseline-Erhebung und der jeweiligen Veränderungsmessung dargestellt (Werte können sich von den oben genannten Rücklaufquoten unterscheiden).

<sup>4)</sup> Die Angaben umfassen jeweils die Teilnehmenden, welche entweder an der regulären Follow-up-Visite in Woche 58 oder der Follow-up-Visite mindestens 6 Wochen nach Behandlungsabbruch teilgenommen haben. Da der Anteil der Studienabbrüche in beiden Studienarmen über einem Drittel liegt, muss dieses bei der Interpretation der Ergebnisse berücksichtigt werden.

Abkürzungen: ITT: Intention-to-Treat; KI: Konfidenzintervall; MMRM: Gemischtes Modell mit wiederholten Messungen; MW: Mittelwert; SF-36: 36-item Short Form Health Survey; SD: Standardabweichung.

Tabelle 29: Ergebnisse zu den Subskalen des SF-36 in der Studie SUSTAIN – ITT-Population

Studie SUSTAIN	Crizanlizumab 5,0 mg/kg N = 67	Placebo N = 65
<b>SF-36 – Subskala körperliche Funktionsfähigkeit<sup>1)</sup></b>		
<i>Behandlungsbeginn</i> n (%) <sup>2)</sup> MW (SD)	63 (94) 64,67 (26,04)	65 (100) 65,43 (23,13)
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>3)</sup></i> n (%) <sup>2)</sup> MW (SD)	47 (70,1) 1,35 (18,56)	46 (70,8) -1,58 (21,94)
<b>SF-36 – Subskala körperliche Rollenfunktion<sup>1)</sup></b>		
<i>Behandlungsbeginn</i> n (%) <sup>2)</sup> MW (SD)	63 (94) 59,23 (26,96)	65 (100) 59,42 (29,94)
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>3)</sup></i> n (%) <sup>2)</sup> MW (SD)	49 (73,1) 3,53 (29,57)	46 (70,8) 1,77 (32,29)
<b>SF-36 – Subskala körperlicher Schmerz<sup>1)</sup></b>		
<i>Behandlungsbeginn</i> n (%) <sup>2)</sup> MW (SD)	63 (94) 52,81 (27,31)	65 (100) 51,65 (28,87)
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>3)</sup></i> n (%) <sup>2)</sup> MW (SD)	48 (71,6) 0,75 (27,62)	46 (70,8) 3,89 (27,5)
<b>SF-36 – Subskala allgemeine Gesundheitswahrnehmung<sup>1)</sup></b>		
<i>Behandlungsbeginn</i> n (%) <sup>2)</sup> MW (SD)	63 (94) 41,49 (21,24)	65 (100) 40,18 (19,69)
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>3)</sup></i> n (%) <sup>2)</sup> MW (SD)	48 (71,6) 1,94 (16,96)	46 (70,8) 2,42 (14,69)
<b>SF-36 – Subskala Vitalität<sup>1)</sup></b>		
<i>Behandlungsbeginn</i> n (%) <sup>2)</sup> MW (SD)	63 (94) 47,35 (19,06)	65 (100) 47,76 (26,19)
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>3)</sup></i> n (%) <sup>2)</sup> MW (SD)	49 (73,1) 2,47 (19,03)	46 (70,8) -0,77 (14,75)
<b>SF-36 – Subskala soziale Funktionsfähigkeit<sup>1)</sup></b>		
<i>Behandlungsbeginn</i> n (%) <sup>2)</sup> MW (SD)	63 (94) 69,64 (24,25)	65 (100) 65,96 (27,02)
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>3)</sup></i> n (%) <sup>2)</sup> MW (SD)	49 (73,1) 2,04 (21,85)	46 (70,8) 4,35 (24,47)

Studie SUSTAIN	Crizanlizumab 5,0 mg/kg N = 67	Placebo N = 65
<b>SF-36 – Subskala emotionale Rollenfunktion<sup>1)</sup></b>		
<i>Behandlungsbeginn</i> n (%) <sup>2)</sup> MW (SD)	63 (94) 72,49 (24,58)	65 (100) 66,15 (33,04)
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>3)</sup></i> n (%) <sup>2)</sup> MW (SD)	49 (73,1) -4,25 (23,39)	46 (70,8) 1,09 (31,01)
<b>SF-36 – Subskala psychisches Wohlbefinden<sup>1)</sup></b>		
<i>Behandlungsbeginn</i> n (%) <sup>2)</sup> MW (SD)	63 (94) 70,54 (21,28)	65 (100) 67,31 (25,31)
<i>Veränderung Behandlungsbeginn zur Follow-up-Visite<sup>3)</sup></i> n (%) <sup>2)</sup> MW (SD)	49 (73,1) 2,51 (20)	46 (70,8) 1,77 (22,62)

<sup>1)</sup> Skala: 0–100. Höhere Werte auf den Summenskalen entsprechen einer besseren gesundheitsbezogenen Lebensqualität.

<sup>2)</sup> Anzahl der Patientinnen und Patienten in der Auswertung (Rücklaufquote des Fragebogens) bezogen auf die ITT-Population. Bei Veränderungsmessungen sind bezüglich MW und SD alle Personen mit Baseline-Erhebung und der jeweiligen Veränderungsmessung dargestellt (Werte können sich von den oben genannten Rücklaufquoten unterscheiden).

<sup>3)</sup> Die Angaben umfassen jeweils die Teilnehmenden, welche entweder an der regulären Follow-up-Visite in Woche 58 oder der Follow-up-Visite mindestens 6 Wochen nach Behandlungsabbruch teilgenommen haben. Da der Anteil der Studienabbrüche in beiden Studienarmen über einem Drittel liegt, muss dieses bei der Interpretation der Ergebnisse berücksichtigt werden.

Abkürzungen: ITT: Intention-to-Treat; KI: Konfidenzintervall; MMRM: Gemischtes Modell mit wiederholten Messungen; MW: Mittelwert; SF-36: 36-item Short Form Health Survey; SD: Standardabweichung.