

# Nutzenbewertung



Gemeinsamer  
Bundesausschuss

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

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

**Wirkstoff: Voretigen Neparvovec**

Datum der Veröffentlichung: 1. Juli 2022

## Inhaltsverzeichnis

Tabellenverzeichnis .....	4
Abbildungsverzeichnis.....	6
Abkürzungsverzeichnis.....	7
Hintergrund .....	9
1 Fragestellung .....	10
2 Auswahl und Methodik der relevanten Studien .....	11
2.1 Studienbasis für die Nutzenbewertung .....	11
2.2 Beschreibung der eingeschlossenen Studien .....	12
2.3 Endpunkte.....	19
2.3.1 Mortalität.....	20
2.3.2 Morbidität.....	20
2.3.3 Lebensqualität .....	31
2.3.4 Sicherheit .....	32
2.3.5 Übersicht der Erhebungszeitpunkte.....	35
2.4 Statistische Methoden.....	37
2.5 Verzerrungspotential auf Studien- und Endpunktebene .....	40
3 Ergebnisse der eingeschlossenen Studien.....	43
3.1 Studiencharakteristika und Studienmedikation .....	43
3.2 Mortalität.....	51
3.3 Morbidität.....	51
3.4 Lebensqualität .....	63
3.5 Sicherheit .....	63
4 Diskussion der Methodik und Ergebnisse .....	73
4.1 Zulassungsstatus und Zulassungspopulation von Voretigen Neparvovec .....	73
4.2 Design und Methodik der Studie .....	73
4.3 Mortalität.....	75
4.4 Morbidität.....	76
4.5 Lebensqualität .....	81
4.6 Sicherheit .....	81
5 Anforderung an eine qualitätsgesicherte Anwendung .....	84
6 Zusammenfassung der Nutzenbewertung .....	85
Referenzen .....	90
Anhang .....	93
A 1 Studie 301 und LTFU .....	93
A 2 Beschreibung der ergänzend dargestellten Registerstudie 401.....	96
A 2.1 Endpunkte .....	98
A 2.1.1 Mortalität.....	98
A 2.1.2 Sicherheit .....	99
A 2.2 Statistische Methoden .....	101
A 2.3 Ergebnisse.....	101

A 2.3.1 Studiencharakteristika und Studienmedikation .....	101
A 2.3.2 Mortalität .....	103
A 2.3.3 Morbidität .....	103
A 2.3.4 Lebensqualität.....	103
A 2.3.5 Sicherheit .....	103

## Tabellenverzeichnis

Tabelle 1: Übersicht über die Studienbasis .....	11
Tabelle 2: Charakterisierung der Studie 301 .....	12
Tabelle 3: Charakterisierung der Studie LTFU .....	16
Tabelle 4: Charakterisierung der Intervention in Studie 301 und LTFU.....	18
Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie 301 und der LTFU-Studie..	19
Tabelle 6: Umrechnungstabelle für den Mobilitätstest-Score .....	22
Tabelle 7: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie 301/LTFU.....	36
Tabelle 8: Verzerrungspotential der Studie 301 .....	40
Tabelle 9: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie 301 .....	41
Tabelle 10: Allgemeine Angaben; Studie 301.....	43
Tabelle 11: Allgemeine Angaben; LTFU-Studie; Datenschnitt 30. Juni 2020.....	45
Tabelle 12: Charakterisierung der Studienpopulation in Studie 301 .....	46
Tabelle 13: LTFU-Studie: Charakterisierung der Studienpopulation (mITT).....	48
Tabelle 14: Studie 301: Begleitmedikation mit $\geq 30\%$ in einer der beiden Behandlungsgruppen (Sicherheitspopulation) .....	50
Tabelle 15: Studie 301: Änderung im MLMT bilateral gemessen, Jahr 1B/C im Vergleich zu Baseline (ITT-Population) .....	51
Tabelle 16: LTFU-Studie: Ergebnisse zum Endpunkt MLMT bilateral gemessen im Vergleich zu Baseline vor Injektion bis zum Datenschnitt vom 30. Juni 2020 (mITT-Population der Studie 301).....	52
Tabelle 17: Ergebnisse zum Endpunkt Sehschärfe für beide Sehtafeln (ETDRS, HOTV) zusammen: Jahr 1B/C im Vergleich zu Baseline (ITT-Population).....	54
Tabelle 18: Studie 301: Endpunkt Sehschärfe (ETDRS allein): Ergebnisse der Responderanalysen zur Verbesserung um $\geq 10$ (logMAR $\geq 0,2$ ) bzw. $\geq 15$ Buchstaben (logMAR $\geq 0,3$ ), Jahr 1B/C im Vergleich zu Baseline (ITT-Population) .....	55
Tabelle 19: LTFU-Studie: Endpunkt Sehschärfe (ETDRS- und HOTV-Sehtafel, Holladay Off-Chart), gemittelt über beide Augen, Datenschnitt am 30. Juni 2020 (mITT-Population).....	55
Tabelle 20: Studie 301: Ergebnisse zum Endpunkt FST zur Visite Jahr 1B/C im Vergleich zu Baseline (ITT-Population) .....	56
Tabelle 21: LTFU-Studie: Ergebnisse zum Endpunkt FST gemittelt über beide Augen bis zum Datenschnitt am 30. Juni 2020 (mITT-Population der Studie 301).....	58
Tabelle 22: Studie 301: Gesichtsfeldmessung, gemittelt über beide Augen; Jahr 1B/C im Vergleich zu Baseline (ITT-Population) .....	60
Tabelle 23: LTFU-Studie: Perimetrie nach Goldmann, gemittelt über beide Augen, Datenschnitt am 30. Juni 2020 (mITT-Population der Studie 301).....	61

Tabelle 24: LTFU-Studie: Perimetrie nach Humphrey: gemittelt über beide Augen, Datenschnitt am 30. Juni 2020 (mITT-Population der Studie 301).....	62
Tabelle 25: Studie 301: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE zu Jahr 1 nach Baseline (Sicherheitspopulation) .....	63
Tabelle 26: Studie LTFU: Ergebnisse der Nebenwirkungen <sup>1</sup> – Zusammenfassung der UE bis zum Datenschnitt vom 30. Juni 2020 (Sicherheitspopulation) .....	64
Tabelle 27: Studie 301: UE mit Inzidenz $\geq 15$ % der Sicherheitspopulation innerhalb eines Jahres nach Baseline (Sicherheitspopulation).....	65
Tabelle 28: Studie LTFU: UE <sup>1</sup> mit Inzidenz $\geq 15$ % bis Jahr 1B bzw. Jahr 5 (Sicherheitspopulation) .....	68
Tabelle 29: Studie 301: Alle SUE innerhalb eines Jahres nach Baseline (Sicherheitspopulation) .....	69
Tabelle 30: Studie LTFU: alle SUE <sup>1</sup> bis Jahr 1B bzw. Jahr 5 (Sicherheitspopulation).....	70
Tabelle 31: Studie 301: Alle schweren UE <sup>1</sup> innerhalb eines Jahres nach Baseline (Sicherheitspopulation).....	70
Tabelle 32: Studie LTFU: alle schweren UE <sup>1,2</sup> bis Jahr 1B bzw. Jahr 5 (Sicherheitspopulation) ...	71
Tabelle 33: Zusammenfassende Darstellung der Ergebnisse der Studie 301 (ITT-Population)....	85
Tabelle 34: Zusammenfassende Darstellung der Ergebnisse der LTFU-Studie für die ehemalige Studie 301 (mITT-Population) bis Datenschnitt vom 30. Juni 2020 .....	88
Tabelle 35: Studie 301: alle UE Grad $\geq 3$ <sup>1</sup> innerhalb eines Jahres nach Baseline (Sicherheitspopulation).....	93
Tabelle 36: LTFU-Studie: alle UE Grad $\geq 3$ <sup>1</sup> bis Jahr 1B bzw. Jahr 5 (Sicherheitspopulation) .....	94
Tabelle 37: Charakterisierung der Registerstudie 401.....	96
Tabelle 38: Zusammenfassung der Endpunktbewertung der Studie 401 .....	98
Tabelle 39: Erhebungszeitpunkte der ergänzend dargestellten Endpunkte in der Registerstudie 401 .....	100
Tabelle 40: Registerstudie 401: Allgemeine Angaben; Datenschnitt vom 31. August 2021 .....	102
Tabelle 41: Registerstudie 401: Charakterisierung der Studienpopulation.....	102
Tabelle 42: Registerstudie 401: Ergebnisse der Nebenwirkungen (Datenschnitt: 31. August 2021) .....	104
Tabelle 43: Registerstudie 401: Okuläre UE mit Inzidenz $\geq 10,0$ %, FAS-Population (Datenschnitt: 31. August 2021) .....	104
Tabelle 44: Registerstudie 401: Okuläre UE von besonderem Interesse <sup>1</sup> , FAS-Population; Datenschnitt: 31. August 2021.....	105

## **Abbildungsverzeichnis**

Abbildung 1: Studiendesign der Studie 301/LTFU [24] .....	15
Abbildung 2: Beispiel für einen der 12 Hindernisparcours des MLMT [1] .....	21
Abbildung 3: Verwendete Beleuchtungsstärken beim MLMT der Studie 301 [10] .....	22
Abbildung 4: Beispiel einer ETDRS-Tafel [27] .....	24
Abbildung 5: Beispiel einer HOTV-Tafel [27] .....	25

## Abkürzungsverzeichnis

AESI	Adverse events of special interest
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
CHOP	Children's Hospital of Philadelphia
CSR	Clinical study report
CRF	Case report form
EMA	European Medicines Agency
EMEA	Europe, Middle East, Africa
EPAR	European Public Assessment Report
ETDRS	Early Treatment of Diabetic Retinopathy Study
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
FAS	Full analysis set
FDA	Food and Drug Administration
FST	Full-Light-Sensitivity-Threshold-Test (Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test)
IOWA	University of Iowa
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to Treat
KI	Konfidenzintervall
logMAR	Logarithmus des minimalen Auflösungswinkels
LTFU	Long-Term Follow-up
MedDRA	Medical Dictionary for Regulatory Activities
MID	Minimal important difference
mITT	Modified Intention-to-Treat
MLMT	Multi-Luminanz-Mobilitätstest
MMRM	Mixed model for repeated measures
MOP	Manual operating procedure
MW	Mittelwert
N	Anzahl
NEI-VFQ-25	National Eye Institute Visual Functioning Questionnaire - 25
PRO	Patient reported outcome
PT	Preferred Term
pU	Pharmazeutischer Unternehmer
RPE65	Retinales Pigmentepithel-spezifisches 65 kDa-Protein
SAP	Statistischer Analyseplan
SGB	Sozialgesetzbuch
SMD	Standardisierte Mittelwertdifferenz
SOC	Systemorganklasse
SOP	Standard operating procedure
SST	Submacular Surgery Trials Research Group

SUE	Schwerwiegendes unerwünschtes Ereignis
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des G-BA
Vg	Vektorgenome
WHO	World Health Organization



## Hintergrund

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

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Voretigen Neparvovec in seiner Sitzung am 28. Juni 2022 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 31. März 2022 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

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

Die Nutzenbewertung wird am 01. Juli 2022 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

Voretigen Neparvovec (Luxturna®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der G-BA bestimmt bei Orphan Drugs, die einen Umsatz von 50 Millionen Euro in den letzten 12 Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. Die Bewertung erfolgt gemäß Zulassungsstatus für folgende Indikation [3]:

Behandlung von erwachsenen und pädiatrischen Patientinnen und Patienten mit Sehverlust aufgrund einer erblichen Netzhautdystrophie, die auf nachgewiesenen biallelischen RPE65-Mutationen beruht, und die über ausreichend lebensfähige Netzhautzellen verfügen.

Gemäß Fachinformation beträgt die zugelassene Dosierung  $1,5 \times 10^{11}$  Vg Voretigen Neparvovec pro Auge. Jede Dosis wird in einem Gesamtvolumen von 0,3 ml in den subretinalen Raum verabreicht. Die individuelle Verabreichung pro Auge erfolgt an unterschiedlichen Tagen innerhalb eines kurzen Zeitintervalls, das jedoch mindestens 6 Tage betragen muss [3].

## 2 Auswahl und Methodik der relevanten Studien

### 2.1 Studienbasis für die Nutzenbewertung

Tabelle 1: Übersicht über die Studienbasis

Studiename (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzenbewertung	Ausschlussgrund
<b>Studien zum Wirkstoff</b>				
Studie 101	Ja	Nein	Nein	<ul style="list-style-type: none"> <li>• Einarmige Dosis-Findungsstudie mit einer Studiendauer von 1 Jahr</li> <li>• Nur 3 der 12 Teilnehmenden erhielten die zulassungsrelevante Dosierung in ein Auge.</li> </ul>
Studie 102	Ja	Nein	Nein	<ul style="list-style-type: none"> <li>• Einarmige Nachbeobachtungsstudie der Studie 101</li> <li>• 9 der 12 Teilnehmenden erhielt die zulassungsrelevante Dosierung in das andere Auge zu Beginn der Studie.</li> <li>• Zeitlicher Abstand (1 Jahr) zwischen beiden Injektionen ist nicht zulassungskonform.</li> </ul>
Studie 301 <sup>1)</sup>	Ja	Ja	Ja	
Studie LTFU <sup>1)2)</sup>	Ja	Ja	Ja	
Studie 401	Ja	Ja	Nein <sup>3)</sup>	Auf Basis der medianen Beobachtungsdauer von ca. 0,65 Jahren sind keine Aussagen zu Langzeiteffekten möglich.

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

<sup>2)</sup> Es werden die Ergebnisse der Teilnehmenden der vorherigen Studie 301 in der Nutzenbewertung dargestellt. Der pU reicht keine weiteren Daten zu den vorherigen Studien 101/102 ein.

<sup>3)</sup> Ergänzende Darstellung im Anhang der Nutzenbewertung

Abkürzungen: EPAR: European Public Assessment Report; LTFU: Long-term follow-up; pU: pharmazeutischer Unternehmer

Der pU reicht im aktuellen Dossier auf Grundlage der vorherigen Nutzenbewertung zu Voretigen Neparvovec [4,7] und dem zugehörigen Beschluss [5,6,8,9] für die Langzeitnachbeobachtungsstudie LTFU Ergebnisse für diejenigen Personen, die zuvor an der Vorgängerstudie 301 teilnahmen, ein.

Die Registerstudie 401 war Bestandteil der Befristungsaufgaben zur Neubewertung nach Fristablauf der Langzeiteffekte, insbesondere zur Sicherheit [6,9]. Es handelt sich hierbei um eine nicht-interventionelle, offene, einarmige Beobachtungsstudie mit erwachsenen und pädiatrischen Patientinnen und Patienten im Alter von 3 Jahren oder älter mit einem Sehverlust aufgrund einer erblichen Netzhautdystrophie, die die Intervention bereits erhalten haben oder noch erhalten werden. Es wurden insgesamt 106 Teilnehmende, von denen 103 das Prüfpräparat erhalten haben, in die Studie aufgenommen. Die mediane Beobachtungszeit seit Injektion ins erste Auge bis zum Datenschnitt vom 31. August 2021 beträgt jedoch nur

0,65 Jahre (min; max: 0; 2,3). Im Gegensatz dazu beträgt die mediane Beobachtungszeit in der Studie LTFU seit Injektion ins erste Auge bis Datenschnitt (30. Juni 2020) im ehemaligen Interventionsarm der Studie 301 (N = 20) im Median 2.489 Tage (min; max: 2.276; 2.722) bzw. im Mittel ca. 6 Jahre 10 Monate und im ehemaligen Kontrollarm der Studie 301 (N = 9) im Median 2.197 Tage (1.987; 2.395) bzw. ca. im Mittel 6 Jahre. Aufgrund der geringen Beobachtungszeit der Studie 401 bis zum aktuellen Datenschnitt können darauf basierend keine Aussagen zu Langzeiteffekten getroffen werden. Es erfolgt jedoch eine ergänzende Darstellung der Sicherheitsauswertung der Studie 401 im Anhang der Nutzenbewertung.

### Zur Nutzenbewertung für Voretigen Neparvovec herangezogene Studien und Daten

- Herstellerdossier zu Voretigen Neparvovec [10]
- Nutzenbewertung (15. Juli 2019) + Amendment Teil 1 + 2 (12. September 2019) zu Voretigen Neparvovec, veröffentlicht am 15. Juli 2019 bzw. 15. Oktober 2019 [4,7]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [2]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der pivotalen Studie 301 [17,23,24,25]
- Studienbericht inkl. Studienprotokoll und SAP der Studie LTFU [18,19,21,22]
- Studienbericht inkl. Studienprotokoll und SAP der Registerstudie 401 [12,13,14]

## 2.2 Beschreibung der eingeschlossenen Studien

Die Unterlagen zur Nutzenbewertung für Voretigen Neparvovec basieren auf der Zulassungsstudie 301 und der Nachbeobachtungsstudie LTFU. Die Studien und die Interventionen werden in den Tabellen 2–4 charakterisiert.

*Tabelle 2: Charakterisierung der Studie 301*

Charakteristikum	Beschreibung
<b>Design und Studienablauf</b>	<p>Randomisierte, kontrollierte, offene Phase-III-Studie im Parallelgruppenvergleich, bei der Voretigen Neparvovec als gentherapeutische Intervention in Form einer subretinalen Injektion mit beobachtendem Abwarten in der Kontrollgruppe verglichen wurde. Ein Wechsel von Personen der Kontrollgruppe in den Interventions-Arm war nach mindestens 1 Jahr möglich. Die Beobachtungsdauer nach Injektion der Prüfmedikation ins zweite Auge betrug in beiden Gruppen 1–1,5 Jahre, wobei in der Kontrollgruppe die Möglichkeit der Behandlung mit dem Prüfpräparat erst nach Visite zu Jahr 1 bestand. Die Patientinnen und Patienten konnten im Anschluss in der LTFU-Studie bis zu 15 Jahre weiterbeobachtet werden. Das Studiendesign ist in Abbildung 1 dargestellt.</p> <p>Die Randomisierung erfolgte im Verhältnis 2:1 in Form einer Block-Randomisierung (Blockgröße von 3), stratifiziert nach Alter zum Zeitpunkt der Screening-Visite (<math>\geq 10</math> Jahre vs. <math>&lt; 10</math> Jahre) und Ergebnislevel des Mobilitätstests zum Screening mit dem am schlechtesten im Test abschneidenden Auge (Bestehen bei <math>\geq 125</math> Lux vs. <math>&lt; 125</math> Lux).</p> <p>Die Studie ist abgeschlossen. Es liegt der finale Studienbericht vom 9. August 2017 vor.</p>

Charakteristikum	Beschreibung
<b>Population</b>	<p><b>Wesentliche Einschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Diagnose einer Leberschen kongenitalen Anomalie (LCA) aufgrund von Mutationen im Gen RPE65; Durchführung oder Bestätigung der Molekulardiagnostik erfolgte durch ein CLIA-zertifiziertes Labor (Clinical Laboratory Improvement Amendments of 1988).</li> <li>• Alter <math>\geq 3</math> Jahre</li> <li>• Sehschärfe schlechter als 20/60 (beide Augen) und/oder ein Gesichtsfeld kleiner als <math>20^\circ</math> in jedem Meridian (beide Augen), gemessen mit beispielsweise III4e-Isopter oder äquivalent.</li> <li>• Ausreichend lebensfähige Netzhautzellen, bestimmt durch nicht-invasive Methoden wie optische Kohärenztomographie (OCT) und/oder Ophthalmoskopie. Folgendes muss erfüllt sein: <ul style="list-style-type: none"> <li>○ Vorhandensein eines Bereiches der Netzhaut innerhalb des hinteren Pols von <math>&gt; 100 \mu\text{m}</math> Dicke sichtbar durch OCT</li> <li>oder</li> <li>○ <math>\geq 3</math> Pupillendurchmesser der Netzhaut ohne Atrophie oder Pigmentdegeneration innerhalb des hinteren Pols</li> <li>oder</li> <li>○ ein verbleibendes Gesichtsfeld innerhalb von <math>30^\circ</math> der Fixation, gemessen mit III4e-Isopter oder gleichwertigem.</li> </ul> </li> <li>• Der MLMT muss auswertbar sein (Primärer Endpunkt): <ul style="list-style-type: none"> <li>○ Die Fähigkeit, den MLMT in der definierten Beleuchtungsstärke zu absolvieren, muss vorhanden sein. Personen mussten einen Genauigkeitswert<sup>1)</sup> von <math>\leq 1</math> bei einem Mobilitätstest zum Zeitpunkt des Screenings von 400 Lux oder weniger erreichen. Personen mit einem Genauigkeitswert<sup>1)</sup> von <math>&gt; 1</math> bei allen MLMT zum Zeitpunkt des Screenings wurden, ebenso wie Personen, die die Teilnahme verweigerten, von der Studienteilnahme ausgeschlossen.</li> <li>○ Die teilnehmenden Personen dürfen den MLMT zum Zeitpunkt des Screenings bei einer Beleuchtungsstärke von 1 Lux nicht bestehen, um für die Studie geeignet zu sein; Personen, die einen oder mehrere MLMT zum Zeitpunkt des Screenings bei 1 Lux bestehen, sind von der Studienteilnahme ausgeschlossen.</li> </ul> </li> </ul> <p><b>Wesentliche Ausschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Die Person ist nicht in der Lage oder nicht gewillt, die Anforderungen der Studie zu erfüllen, was auch den Erhalt von bilateralen, subretinalen Injektionen des Vektors beinhaltet.</li> <li>• Jede vorherige Teilnahme an einer Studie, bei der ein gentherapeutischer Vektor verabreicht wurde.</li> <li>• Einsatz von Retinoid-Verbindungen oder deren Vorläufersubstanzen, die möglicherweise eine Interaktion mit der biochemischen Aktivität des Enzyms RPE65 eingehen könnten. Eine Teilnahme an der Studie war möglich, sofern diese Verbindungen seit 18 Monaten nicht mehr angewandt wurden.</li> <li>• Vorangegangene intraokulare Operation innerhalb von 6 Monaten.</li> <li>• Bekannte Empfindlichkeiten gegenüber Medikamenten, die in der perioperativen Phase verwendet werden sollen.</li> <li>• Vorerkrankungen des Auges oder erschwerende systemische Erkrankungen, die die geplante Operation nicht zulassen oder die Interpretation der Studie beeinträchtigen würden. Zu den erschwerenden systemischen Erkrankungen gehören solche, bei denen die Krankheit an sich oder die Behandlung der Erkrankung die Funktion des Auges verändern kann. Beispiele sind</li> </ul>

Charakteristikum	Beschreibung
	<p>Malignome, deren Behandlung die Funktion des zentralen Nervensystems beeinflussen kann (z. B.: Strahlenbehandlung der Augenhöhle; Leukämie mit Beteiligung des zentralen Nervensystems / Sehnervs). Erkrankte mit Diabetes oder einer Sichelzellerkrankung wurden ausgeschlossen, wenn sie eine fortgeschrittene Retinopathie (z. B. Makulaödem oder proliferative Veränderung) hatten. Erkrankte mit einer Immunschwäche (erworben oder angeboren) wurden ebenfalls ausgeschlossen, da eine Prädisposition für opportunistische Infektionen bestehen kann (wie beispielsweise eine CMV-Retinitis).</p>
<b>Intervention und Zahl der Patientinnen und Patienten</b>	<p>Es wurden 36 Personen gescreent, wovon 31 Personen randomisiert wurden.</p> <ul style="list-style-type: none"> <li>• Interventionsgruppe (Voretigen Neparvovec): N = 21</li> <li>• Kontrollgruppe (Beobachtendes Abwarten): N = 10</li> </ul>
<b>Ort und Zeitraum der Durchführung</b>	<p><b>Ort der Durchführung</b> 2 Studienzentren in den USA</p> <ul style="list-style-type: none"> <li>• Children’s Hospital of Philadelphia (CHOP), N = 19</li> <li>• University of Iowa (Iowa), N = 12</li> </ul> <p><b>Zeitraum der Durchführung</b></p> <ul style="list-style-type: none"> <li>• Einschluss der/des ersten Patientin/Patienten: 15. November 2012</li> <li>• Letzte Visite der/des letzten Patientin/Patienten: 06. April 2015</li> </ul> <p><b>Datenschnitte</b></p> <ul style="list-style-type: none"> <li>• Einziger Datenschnitt vom 16. Juli 2015 (Interimsanalysen waren nicht geplant). Zusätzlich wurde im Studienbericht angegeben, dass die Datenbank für die Visite Jahr 1B/C zum 24. August 2015 geschlossen wurde.</li> <li>• Finaler Studienbericht: 13. Dezember 2016, inklusive 3 administrativer Amendments nach Veröffentlichung des Studienberichts (letztes Amendment vom 9. August 2017)</li> </ul>
<b>Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll</b>	<p><b>Primärer Endpunkt</b></p> <ul style="list-style-type: none"> <li>• MLMT: Änderung im MLMT bilateral gemessen nach Jahr 1B (= Jahr 1 nach Injektion ins zweite Auge in der Interventionsgruppe) bzw. Jahr 1C (= Jahr 1 nach der Baseline-Visite der Kontrollgruppe) im Vergleich zu Baseline</li> </ul> <p><b>Sekundäre Endpunkte</b></p> <ul style="list-style-type: none"> <li>• Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test (FST, Full-Light-Sensitivity-Threshold-Test): Veränderung in der Lichtempfindlichkeit (Durchschnitt beider Augen) für weißes Licht nach einem Jahr (Jahr 1B/C) im Vergleich zu Baseline</li> <li>• Monokularer Mobilitätstest<sup>2)</sup>: Veränderung im Mobilitätstest-Score (unilateral, Durchführung mit dem Auge, das zuerst behandelt wurde) nach einem Jahr (Jahr 1B/C) im Vergleich zu Baseline</li> <li>• Sehschärfe: Veränderung der Sehschärfe anhand der ETDRS- oder HOTV-Tafel (Durchschnitt beider Augen) nach einem Jahr (Jahr 1B/C) im Vergleich zu Baseline</li> </ul> <p><b>Weitere Endpunkte</b></p> <ul style="list-style-type: none"> <li>• Seh- und Netzhautfunktionstest <ul style="list-style-type: none"> <li>○ Pupillenreflextest</li> <li>○ Gesichtsfeldtest (Perimetrie nach Goldmann und Humphrey)</li> <li>○ Kontrastempfindlichkeit<sup>2)</sup></li> </ul> </li> <li>• Orientierungs- und Mobilitätsmessungen</li> <li>• Patientenberichteter Endpunkt zur visuellen Lebensqualität</li> </ul>

Charakteristikum	Beschreibung
	<p><b>Sicherheit (gemäß SAP)</b></p> <ul style="list-style-type: none"> <li>• Unerwünschte Ereignisse</li> <li>• Begleitmedikation</li> <li>• Körperliche Untersuchung</li> <li>• Ophthalmologische Untersuchungen</li> <li>• Klinisches Labor</li> <li>• Hämatologische Untersuchungen</li> <li>• Immunologische Untersuchungen</li> <li>• Vektorablösung<sup>3)</sup></li> </ul>
<b>Subgruppenanalysen</b>	Keine vorgesehen und durchgeführt.

- 1) Genauigkeitswert des MLMT: Anzahl der Verstöße gegen die Genauigkeitsregeln dividiert durch die Gesamtzahl der Hindernisse
- 2) Laut statistischem Analyseplan (SAP Version 4.0 vom 19. August 2015) war die monokulare Veränderung im MLMT nach einem Jahr (Jahr 1B/C) ein sekundärer Wirksamkeitsendpunkt (Durchführung mit dem Auge, das zuerst behandelt wurde), während der Endpunkt Kontrastempfindlichkeit als zusätzlicher Wirksamkeitsendpunkt erfasst wurde. Im Studienprotokoll (Version 6.0 vom 20. August 2013) verhält es sich umgekehrt: MLMT (unilateral) wird als zusätzlicher Endpunkt erwähnt, während die Kontrastempfindlichkeit als sekundärer Endpunkt definiert ist. Die Änderung wurde nach Rücksprache mit der FDA im aktuellen SAP dokumentiert.
- 3) Das Vorhandensein des Vektors im peripheren Blut oder Tränensammlung

Abkürzungen: CHOP: Children’s Hospital of Philadelphia; ETDRS: Early Treatment of Diabetic Retinopathy Study; FST: Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test; Jahr 1B: Jahr 1 nach Behandlung des zweiten Auges in der Interventionsgruppe; Jahr 1C: Jahr 1 nach Baseline der Kontrollgruppe; LTFU: Long-term follow up; MLMT: Multi-Luminanz-Mobilitätstest; OCT: optische Kohärenztomographie; SAP: Statistical Analysis Plan

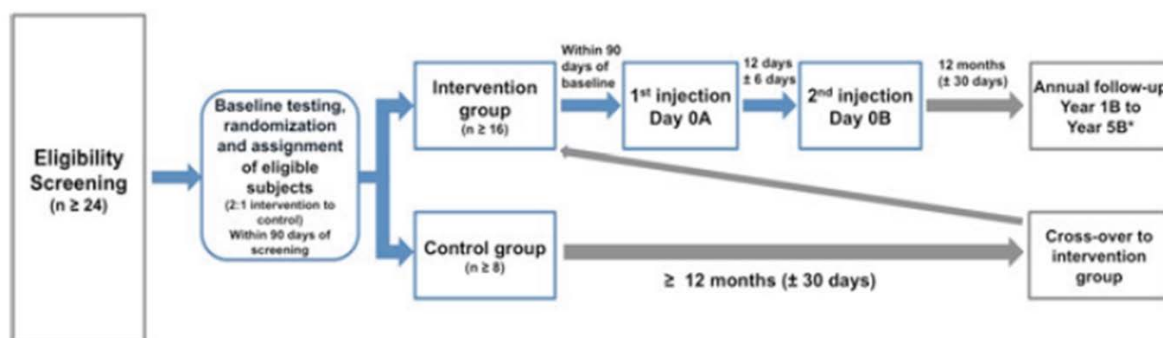


Abbildung 1: Studiendesign der Studie 301/LTFU [24]

Jahr 1B bis 5B: Jahr 1 bis 5 nach Behandlung des zweiten Auges in der ursprünglichen Interventionsgruppe in Studie 301 oder für ehemalige Kontrollgruppe nach Behandlung des zweiten Auges nach Abschluss der Studie 301 und Eintritt in die LTFU-Studie; Jahr 1C: Visite zu Jahr 1 nach Baseline für die Kontrollgruppe der Studie 301; Tag 0A: Applikation des Prüfpräparats ins erste Auge innerhalb von 90 Tagen nach Baseline/Voruntersuchung; Tag 0B: Applikation des Prüfpräparats ins zweite Auge innerhalb von 12 ± 6 Tagen nach Applikation ins erste Auge.

Tabelle 3: Charakterisierung der Studie LTFU

Charakteristikum	Beschreibung
<b>Design und Studienablauf</b>	<p>Diese multizentrische, nichtrandomisierte Beobachtungsstudie soll die Langzeit-Sicherheit und -Wirksamkeit von Voretigen Neparvovec bei Patientinnen und Patienten untersuchen, die eine Injektion bereits in vorherigen Studien erhalten haben. Dadurch können Teilnehmende der Studien 101 bzw. 102 und 301 an dieser Studie teilnehmen. Für die Nutzenbewertung werden nur Ergebnisse der Teilnehmenden, die vorab an der Studie 301 teilgenommen haben, dargestellt. Personen der ursprünglichen Interventionsgruppe der Studie 301 haben das Prüfpräparat bereits zu Beginn der Studie 301 erhalten. Nach Abschluss der Studie 301 konnten Personen der Kontrollgruppe, sofern sie die Ein- und Ausschlusskriterien der Studie 301 zu diesem Zeitpunkt erfüllten, das Prüfpräparat zu Beginn der Studie LTFU/302 erhalten. Es wurden die Ein-/Ausschlusskriterien der Studie 301 bei der Kontrollgruppe erneut überprüft und bei Eignung wurde das Prüfpräparat appliziert. Die Vorgaben zu den Injektionen in beide Augen waren gleich der ursprünglichen Interventionsgruppe.</p> <p>Die Studiendauer soll 15 Jahre betragen; die Studie ist derzeit nicht abgeschlossen.</p>
<b>Population</b>	<p><b>Wesentliche Einschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Patientinnen und Patienten, die bereits vorher in klinischen Studien das Prüfpräparat erhalten haben bzw. an Studien, bei denen das Prüfpräparat untersucht wurde, teilgenommen haben.</li> </ul> <p><b>Wesentliche Ausschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Patientinnen und Patienten, die nicht an der Studie teilnehmen wollen.</li> <li>• Patientinnen und Patienten, bei denen die Prüffärztin / der Prüffarzt eine Teilnahme an der Studie für nicht möglich hält.</li> </ul>
<b>Intervention und Zahl der Patientinnen und Patienten</b>	<p>Alle Personen, die die Studie 301 abgeschlossen haben, nehmen an der Verlängerungsstudie teil. In der Studie 301 wurden 21 bzw. 10 Personen in die Interventions- bzw. Kontrollgruppe randomisiert. Jeweils 1 Person in beiden Behandlungsgruppen ist zu Beginn der Studie 301 ausgeschieden.</p> <ul style="list-style-type: none"> <li>• Interventionsgruppe: N = 20</li> <li>• Ursprüngliche Kontrollgruppe: N = 9</li> </ul>
<b>Ort und Zeitraum der Durchführung</b>	<p><b>Ort der Durchführung</b></p> <ul style="list-style-type: none"> <li>• 2 Studienzentren in den USA</li> </ul> <p><b>Zeitraum der Durchführung</b></p> <ul style="list-style-type: none"> <li>• Datum der Studieninitiierung: 15. November 2012</li> <li>• Erste/r Patient/in, die/der aus der Studie 301 in die Studie LTFU/302 gewechselt ist: Juni 2015</li> <li>• Letzte/r Patient/in, die/der aus der Studie 301 in die Studie LTFU/302 gewechselt ist: 14. Juni 2016</li> </ul> <p><b>Datenschnitte</b></p> <ul style="list-style-type: none"> <li>• 1. Datenschnitt: 16. Juli 2015</li> <li>• Jährliche Datenschnitte vorgesehen</li> <li>• Aktueller Datenschnitt: 30. Juni 2020</li> <li>• Geplantes Studienende: Juni 2030</li> </ul>



Charakteristikum	Beschreibung
<b>Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll</b>	<p><b>Sicherheit</b></p> <ul style="list-style-type: none"> <li>• Sammeln von Daten zu spät auftretenden UE bei Personen, die das Prüfpräparat erhalten haben</li> <li>• Überprüfung eines möglichen Zusammenhangs zwischen UE und Prüfpräparat, um langfristige Konsequenzen für die Abgabe des Prüfpräparats festzulegen</li> <li>• Untersuchung der Expressionsdauer des RPE65-Gens nach Erhalt des Prüfpräparats durch Untersuchung der retinalen bzw. Sehfunktion</li> </ul> <p>Folgende Wirksamkeitsendpunkte konnten erhoben werden:</p> <ul style="list-style-type: none"> <li>• MLMT <ul style="list-style-type: none"> <li>○ Änderung im MLMT bilateral gemessen zu jeder Visite im Vergleich zu Baseline</li> <li>○ Änderungen im MLMT-Summenscore (erstes, zweites Auge und bilateral), gemessen zu jeder Visite im Vergleich zu Baseline</li> </ul> </li> <li>• Lichtempfindlichkeit, gemessen mit dem FST <ul style="list-style-type: none"> <li>○ Änderung in der Lichtempfindlichkeit (Durchschnitt beider Augen) für weißes, blaues und rotes Licht zu jeder Visite im Vergleich zu Baseline</li> </ul> </li> <li>• Sehschärfetest <ul style="list-style-type: none"> <li>○ Änderung der Sehschärfe anhand der ETDRS- oder HOTV-Sehtafel (Durchschnitt beider Augen) bzw. Off-Chart zu jeder Visite im Vergleich zu Baseline</li> </ul> </li> <li>• Gesichtsfeldmessung: Perimetrie nach Goldmann bzw. Humphrey <ul style="list-style-type: none"> <li>○ Änderung im Gesichtsfeld zu jeder Visite im Vergleich zu Baseline</li> </ul> </li> <li>• Fragebogen zur visuellen Lebensqualität</li> <li>• Pupillenreflexionstest</li> <li>• Kontrastempfindlichkeit</li> <li>• Orientierungs- und Mobilitätsmessungen</li> </ul>
<b>Subgruppenanalysen</b>	<p>A priori definiert: keine</p> <p>Post hoc gemäß Studienbericht:</p> <p>Im Anhang des Studienberichts werden Subgruppenanalysen für UE post hoc deskriptiv nach Alter (&lt; 18; ≥ 18 Jahre Jahre) und Geschlecht beschrieben.</p>

Abkürzungen: ETDRS: Early Treatment of Diabetic Retinopathy Study; FST: Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test; LTFU: Long-Term Follow-up; MLMT: Multi-Luminanz-Mobilitätstest; UE: Unerwünschte/s Ereignis/se.

## Protokolländerungen

### Studie 301

Es gab für das Studienprotokoll 5 Änderungen zwischen Version 1 (19. Oktober 2009) und Version 6 (20. August 2013). Version 6 war die Protokollversion, die nach dem Einschluss des/der ersten Patienten/Patientin (15. November 2012) in Kraft getreten ist. Die Änderungen beinhalteten Klarstellungen bzw. administrative Änderungen. Wie viele Studienteilnehmende bei Inkrafttreten des Protokolls 6 bereits in die Studie eingeschlossen waren, ist nicht bekannt.

### Studie LTFU

Das Protokoll zur Studie LTFU wurde als Addendum zum Studienprotokoll der Studie 301 erstmalig 2014 als Version 1 (22. Oktober 2014), im Anschluss als Version 2 (25. April 2016) herausgegeben. Die Protokollversion 2 vom 25. April 2016 wird in einem eigenständigen Dokument als Amendment 1 genannt. Als Amendment 2 wird eine Protokollversion vom 15. Juni 2018 bezeichnet, jedoch wurden für diese Version keine Angaben zu den Änderungen identifiziert. Es folgte ein drittes Amendment (30. Mai 2019). Als relevante Änderung wurde eine Straffung der Beschreibung, welche UE berichtet bzw. nicht berichtet werden sollen, vorgenommen.

## Charakterisierung der Intervention

Tabelle 4: Charakterisierung der Intervention in Studie 301 und LTFU

Intervention	Kontrolle
<p><u>Studie 301</u> Nicht-simultane Injektionen von je <math>1,5 \times 10^{11}</math> Vg Voretigen Neparvec (Luxturna®) mit einem Gesamtvolumen von 300 µl, in den subretinalen Bereich beider Augen.</p> <p>Die Teilnehmenden der Interventionsgruppe konnten die Intervention nicht gleichzeitig in jedes Auge erhalten. Nach der Applikation ins erste Auge (Tag 0A; höchstens 90 Tage nach Baseline) erfolgte innerhalb von 12 Tagen <math>\pm</math> 6 Tage die Injektion in das zweite Auge (Tag 0B).</p> <p>Der empfohlene Applikationsort lag im post-äquatorialen Bereich im hinteren Augenabschnitt. Hierfür wurde 1/3 bis 1/4 der Gesamtfläche der Retina, inklusive eines Teils der Makula, die als Zielort des visuellen Sehens und der Empfindlichkeit angesehen wird, gesetzt. Die Prüfmedikation durfte nicht in die unmittelbare Umgebung der Fovea centralis verabreicht werden, um diese nicht zu beschädigen.</p> <p><u>Studie LTFU</u> Die Vorgaben zu den Injektionen in beide Augen der ehemaligen Kontrollgruppe der Studie 301 entsprachen denen der Studie 301 für die ursprüngliche Interventionsgruppe.</p>	<p><u>Studie 301</u> Beobachtendes Abwarten für mindestens 1 Jahr nach Baseline.</p> <p>Die Kontrollgruppe erhielt keine Injektion (Tag 0C = Baseline). Nach der Visite zu Jahr 1C konnten die Teilnehmenden, sofern die Einschlusskriterien noch erfüllt waren, die Prüfmedikation erhalten. Die Vorgaben zu den Injektionen in beide Augen waren ähnlich der Interventionsgruppe. Die Weiterbeobachtung erfolgte in der Studie LTFU.</p> <p>Im Protokoll konnten keine Kriterien zur Festlegung des ersten bzw. zweiten Auges identifiziert werden.</p>
<p><b>Nicht erlaubte Begleitmedikation</b></p> <p><u>Studie 301</u></p> <p>Andere Studienwirkstoffe</p> <ul style="list-style-type: none"> <li>• Hochdosiertes Vitamin A (&gt; 7.500 Retinol-Äquivalente oder &gt; 3.300 IU am Tag)</li> <li>• Tretinoinhaltige Hautcreme (z. B. Retin-A®)</li> <li>• Isotretinoin</li> <li>• Viagra® (Sildenafil) oder ähnliche Substanzen zur Behandlung der erektilen Dysfunktion</li> <li>• Hydroxychloroquin, Chloroquin, Thioridazin oder ähnliche retinotoxische Substanzen</li> </ul> <p><b>Erlaubte Begleitmedikation</b></p> <p><u>Studie 301</u> Studienteilnehmende im Interventionsarm erhielten systemische Glukokortikoide vor der ersten Vektorinjektion (Tag -3A) und kontinuierlich für die nächsten 8 Tage bis zur zweiten Injektion (Tag 8B). Je nach Zeitpunkt der zweiten Applikation konnten sich die Studienteilnehmenden zwischen 18 und 30 Tagen unter Glukokortikoid-Therapie befinden.</p> <p><u>Vorbehandlung:</u> Zur ersten Injektion von Voretigen Neparvec:</p> <ul style="list-style-type: none"> <li>• Prednison zur oralen Anwendung: <ul style="list-style-type: none"> <li>○ 1 mg/kg/Tag (max. 40 mg/Tag) für 7 Tage ab dem 3. Tag vor der ersten Injektion</li> <li>○ 0,5 mg/kg/Tag (max. 20 mg/Tag) für weitere 5 Tage</li> <li>○ 0,5 mg/kg/QOD (max. 20 mg/Tag) bis 3 Tage vor der zweiten Injektion</li> </ul> </li> </ul>	

Intervention	Kontrolle
Zur zweiten Injektion von Voretigen Neparvovec: <ul style="list-style-type: none"> <li>• Prednison zur oralen Anwendung:               <ul style="list-style-type: none"> <li>○ 1 mg/kg/Tag (max. 40 mg/Tag) für 7 Tage ab dem 3. Tag vor der zweiten Injektion</li> <li>○ 0,5 mg/kg/Tag (max. 20 mg/Tag) bis 8 Tage nach der zweiten Injektion</li> </ul> </li> </ul>	

Abkürzungen: Jahr 1C: Jahr 1 nach Baseline für die Kontrollgruppe der Studie 301; LTFU: Long-Term Follow-up; QOD: every other day; Tag 0A: Tag der Injektion ins erste Auge; Tag 0B: Tag der Injektion ins zweite Auge; Tag 8B: Tag 8 nach Injektion ins zweite Auge; Tag 0C: Baseline für die Kontrollgruppe der Studie 301

## 2.3 Endpunkte

In diesem Kapitel werden die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. Alle Endpunkte der eingeschlossenen Studien (siehe Tabelle 2-3) wurden einer Prüfung unterzogen. Endpunkte, die in der nachfolgenden Tabelle 5 nicht gelistet sind, wurden weder vom pU noch im Rahmen der Nutzenbewertung als patientenrelevant bzw. bewertungsrelevant eingestuft. Tabelle 5 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 5: Zusammenfassung der Endpunktbewertung der Studie 301 und der LTFU-Studie

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU		Berücksichtigung in der Nutzenbewertung
		301	LTFU	
Mortalität <sup>1)</sup>	Mortalität	Ja	Ja	Ja
Änderungen im MLMT („Mobilitätstest“) bilateral gemessen <sup>2)</sup>	Morbidität	Ja	Ja	Ja
Vollfeld-Lichtempfindlichkeits-Schwellenwert-Test („FST“)		Ja	Ja	Ja
Perimetrie nach Humphrey		Ja	Ja	Ja
Perimetrie nach Goldmann		Ja	Ja	Ja
Sehschärfe anhand der ETDRS- oder HOTV-Sehtafel		Ja	Ja	Ja
PRO-Instrument zur Erfassung der visuellen Lebensqualität <sup>3)</sup>	Lebensqualität	Ja	Ja	Nein <sup>4)</sup>
Unerwünschte Ereignisse <sup>5)</sup>	Sicherheit	Ja	Ja	Ja

<sup>1)</sup> Über die Sicherheit erfasst

<sup>2)</sup> Primärer Endpunkt der Studie 301

<sup>3)</sup> Vom pU entwickelter Fragebogen in Orientierung an den National Eye Institute Visual Functioning Questionnaire – 25 (NEI-VFQ-25).

<sup>4)</sup> Validität nicht nachgewiesen.

<sup>5)</sup> Primärer Endpunkt der Studie LTFU

Abkürzungen: ETDRS: Early Treatment of Diabetic Retinopathy Study; FST: Full-Light-Sensitivity-Test; LTFU: Long-Term Follow-up; MLMT: Multi-Luminanz-Mobilitätstest; PRO: Patient-Reported Outcome; pU: pharmazeutischer Unternehmer

### 2.3.1 Mortalität

#### **Mortalität**

Der Endpunkt Mortalität wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

In den Studien 301 und LTFU wurden Todesfälle über die Sicherheit erhoben.

#### **Studie 301**

Der Endpunkt wurde in der Interventionsgruppe ab Tag 0A (= erster Tag der ersten Injektion) bzw. ab dem ersten Tag der Baseline-Untersuchungen in der Kontrollgruppe gemäß Modul 4 dargestellt. Falls die Person vorab aus der Studie ausschied, wurde der Endpunkt als UE, sofern dieser behandlungsbedingt war, bis zu 30 Tage nachverfolgt.

#### **LTFU**

Informationen zu Todesfällen wurden im Rahmen der (S)UE-Dokumentation erhoben.

##### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Für den Interventionsarm wurden UE erst ab Injektion ins erste Auge und nicht, anders als in der Kontrollgruppe, bereits ab Baseline berichtet.

In der Studie LTFU ist die Operationalisierung nachvollziehbar. Todesfälle wurden über SUE erfasst, die gemäß Studienprotokoll vollständig erhoben werden sollten.

#### 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

In beiden Studien wurden Todesfälle über die Sicherheit erfasst. Kritisch anzumerken ist der ungleiche Erfassungsbeginn der UE in der Interventionsgruppe, die nicht ab Baseline, wie in der Kontrollgruppe, sondern ab dem Tag der ersten Injektion gezählt wurden. Die mediane Beobachtungsdauer zwischen Baseline und erster Injektion betrug 22 Tage (min; max: 6; 78). Es wird jedoch davon ausgegangen, dass Todesfälle bei Einschluss in die Studie erfasst wurden und Limitationen sich auf andere Sicherheitsendpunkte auswirken.

### 2.3.2 Morbidität

#### **Multi-Luminanz-Mobilitätstest (MLMT)**

Der Endpunkt MLMT wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

#### **Studie 301/LTFU**

Der Multi-Luminanz-Mobilitätstest (MLMT) dient dazu, Veränderungen im funktionalen Sehen zu messen, insbesondere die Fähigkeit, sich unter verschiedenen Lichtverhältnissen in einem Hindernisparcours zu orientieren und selbstständig bewegen zu können.

### Aufbau

Die Gesamtfläche des Hindernisparcours betrug 7 x 12 Fuß, wobei der Hindernisparcours auf einer 5 x 10 Fuß großen Fläche aufgestellt wurde und er zusätzlich eine 1 Fuß große, an allen Ecken angrenzende freie Fläche besitzt. Für den Test wurden 7 standardisierte Beleuchtungsstärken von 1 Lux bis 400 Lux verwendet, die sich mit kalibrierten Lichtmessgeräten an 5 verschiedenen Positionen auf dem Hindernisparcours befanden. Für die Durchführung des Mobilitätstests wurden mit Hilfe von dreidimensionalen Hindernissen, mit hohem Kontrast zum Boden und zu den Stirnwänden, mit Farbunterschieden sowie mit Positionierung in unterschiedlichen Höhen, 12 Hindernisparcours mit vergleichbarem Schwierigkeitsgrad entwickelt (Abbildung 2).

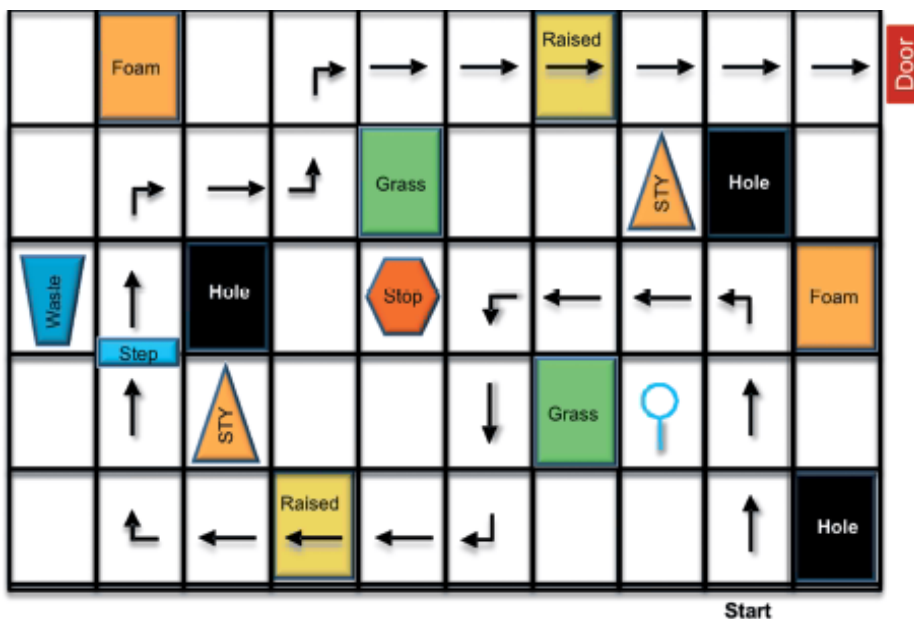


Abbildung 2: Beispiel für einen der 12 Hindernisparcours des MLMT [1]

### Durchführung des Tests

Vor den Baseline-/Studienvisiten wurden Trainingsdurchläufe in allen 3 Testsituationen (bilateral und unilateral mit jedem Auge einzeln) bei einer Beleuchtungsstärke von 250 Lux (und ggf. niedriger) mit einem hierfür extra eingesetzten Übungsparcours durchgeführt. Sehhilfen konnten ggf. bei den Übungsläufen eingesetzt werden. Die Teilnehmenden erhielten wesentliche Instruktionen und Anweisungen zur Durchführung des Tests.

Bei der eigentlichen Durchführung des Tests während der Studie sollten sich die Patientinnen und Patienten in einer 40-minütigen Phase bei niedriger Beleuchtung an die Dunkelheit gewöhnen. Danach wurde zunächst mit dem ersten, dann mit dem zweiten Auge und zuletzt mit beiden Augen unter mindestens 2 unterschiedlichen Lichtverhältnissen ein Hindernisparcours durchlaufen. Sofern es eine starke Abweichung im Testergebnis zwischen beiden Augen gab, konnte ein weiterer Testdurchlauf stattfinden. Gemäß dem Modul 4 des Nutzendossiers waren keine Hilfsmittel (Gehstock, Sehhilfe, etc.) während der Durchführung des Tests für beide Behandlungsgruppen erlaubt.

Die unterschiedlichen Beleuchtungsstärken sind laut Studienunterlagen alltäglichen Lichtverhältnissen nachempfunden (Abbildung 3).

Lux	Leuchtdichte (cd/m <sup>2</sup> )	Entsprechende Umgebung
1	0,32 Mesopisches Sehen	Mondlose Sommernacht; Innenraum-Nachtbeleuchtung
4	1,3 Mesopisches Sehen	Wolkenlose Sommernacht mit Halbmond; Parkplatz bei Nacht
10	3,2 Mesopisches Sehen	60 min nach Sonnenuntergang in städtischer Umgebung; Bushaltestelle bei Nacht
50	15,9 Photopisches Sehen	Außenbereich eines Bahnhofs bei Nacht; schwach beleuchtetes Treppenhaus
125	39,8 Photopisches Sehen	30 min vor einem wolkenlosen Sonnenaufgang; Innenraum eines Einkaufszentrums, Zuges oder Busses bei Nacht
250	79,6 Photopisches Sehen	Innenraum eines Aufzuges, Büroflurs oder einer Bibliothek
400	127,3 Photopisches Sehen	Büroräume

cd/m<sup>2</sup>: Candela pro Quadratmeter; min: Minuten

Abbildung 3: Verwendete Beleuchtungsstärken beim MLMT der Studie 301 [10]

Die Tests sollten alle am gleichen Tag stattfinden, allerdings bestand die Möglichkeit, Durchläufe am Folgetag durchzuführen. Es gab vor der Visite eine randomisierte Auswahl an Hindernisparcours für jede Augen- und Lichtkombination, um den Lerneffekt zu minimieren. Es wurden keine weiteren verbalen oder physischen Hinweise gegeben.

Jeder Testdurchlauf wurde von 2 Personen (bei Dissens 3 Personen) auf Basis von Videoaufnahmen bewertet. Im Manual sind zusätzlich Vorgaben zur Verwendung der Kamera angegeben. Das Ergebnis des Tests wurde verblindet bewertet. Informationen zur Gruppenzuteilung oder zu anderen Ergebnissen von retinalen/visuellen Funktionstests lagen den bewertenden Personen nicht vor. Zur standardisierten Bewertung der Tests fanden Schulungen statt.

Maßgeblich für die Bewertung der Mobilität war die Beleuchtungsstärke, bei der die Patientin oder der Patient den Mobilitätstest noch bestehen konnte (niedrigster Lichtsensitivitätsgrenzwert). Dieser Lux-Wert wurde dann für jede Testsituation in einen Mobilitätstest-Score zwischen 6 und -1 umgewandelt, wobei ein höherer Score einer besseren Mobilität entspricht (Tabelle 6). Um den Durchlauf bei einer Beleuchtungsstärke zu bestehen, mussten sowohl vorab definierte Genauigkeits- als auch die Zeitkriterien bestanden werden. Dies trifft auf alle Testsituationen (jeweils ein Auge bzw. mit beiden Augen unverdeckt) zu.

Tabelle 6: Umrechnungstabelle für den Mobilitätstest-Score

Lux	1	4	10	50	125	250	400	> 400
Mobilitätsscore	6	5	4	3	2	1	0	-1

#### Qualitätssicherung der Bewertung der Tests während der Studie 301

Während der Studie war geplant, die Interrater- und die Intrarater-Reliabilität als auch die Konsentierung zwischen den Bewertern bzw. Bewerberinnen zu untersuchen, indem sowohl für die Kriterien zum Bestehen einer Testsituation (Genauigkeits- und Zeitkriterien) als auch des finalen Bewertungskriteriums „bestanden / nicht bestanden“ Intra-Class-Korrelationen, Kappa-Werte sowie Anzahl an Übereinstimmungen berechnet wurden. Alle 3 Monate wurden 10 % der Videos vom vorherigen Quartal zufällig für eine erneute Beurteilung ausgewählt und zu den neu zu bewertenden Videos gemischt, ohne dass es die Beurteiler bzw. Beurteilerinnen wussten.

Für die Validierungsstudie zum MLMT als auch für die Studie 301 wurden im Rahmen der Qualitätssicherung 3.651 Videos von mindestens 2 Bewertern/Bewerberinnen beurteilt. Die Übereinstimmung für das Bestehen des Tests bei einer Leuchtstärke zwischen den Bewerter/innen betrug  $k = 0,95$ .

### **Studie LTFU**

Gemäß Modul 4 war die Operationalisierung vergleichbar mit Studie 301.

Gemäß Studienprotokoll wurde für die Wirksamkeitsendpunkte ein Bewertungsparadigma (im Studienprotokoll als „evaluation paradigm“ bezeichnet), bezogen auf die Auswahl der Tests und deren Durchführung, individuell an die Patientinnen und Patienten angepasst, abhängig von ihren individuellen Zielen, dem Grad der Sehfunktion, dem Grad des Verständnisses der Durchführungen, der Ausprägung des Nystagmus und der Kooperationsfähigkeit und Geduld bei längeren Testsitzungen.

#### *Bewertung*

Die Operationalisierung ist für die Studie 301 nachvollziehbar. Positiv anzumerken ist die verblindete Beurteilung der Testdurchläufe durch 2 (bzw. optional 3) Bewerter bzw. Bewerterinnen bei einem offenen Studiendesign.

Die Operationalisierung ist für die Studie LTFU überwiegend nachvollziehbar. Es bleibt unklar, inwieweit eine Anpassung der Testdurchführung an die individuelle Situation der Teilnehmenden vorgenommen wurde und dies die Standardisierung und Vergleichbarkeit der Erhebung des Endpunkts während der Studie und zwischen beiden Studien beeinflusste.

#### Patientenrelevanz

Die Orientierung und Mobilität unter unterschiedlichen Lichtverhältnissen wird als patientenrelevant angesehen, da diese direkt wahrnehmbar ist. Dabei wird die Patientenrelevanz der bilateralen Testsituation gegenüber anderen Testszenarien mit verdeckten Augen als übergeordnet bewertet, da diese der Alltagssituation von Erkrankten am nächsten ist.

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“.

#### Validität

An der vom pU eingereichten Validierungsstudie mit einer Studiendauer von 1 Jahr und insgesamt 4 Visiten nahmen 60 Probanden teil, davon 31 Personen mit einer Sehbeeinträchtigung (u. a. n = 20 mit Leberscher kongenitaler Amaurose, n = 4 mit Retinitis Pigmentosa) und 29 normalsichtig. Das Durchschnittsalter betrug 15 Jahre (SD 9) und lag zwischen 4 und 40 Jahren [1,20].

Die Durchführung der Validierungsstudie ähnelte der Durchführung und Bewertung der Studie 301. Es wurden 12 Durchläufe für jede Visite zufällig ausgewählt. Je Visite konnte ein Teilnehmender nicht mehr als 12 Durchläufe durchführen (d. h. maximal 4 Beleuchtungsstärken für jedes der 3 Augenszenarien) und es erfolgte ein Test mit jedem Auge einzeln oder mit beiden Augen unverdeckt. Anders als in der Studie 301 wurde in der Validierungsstudie für die Beleuchtungsstärken 100 und 150 Lux der gleiche Mobilitätsscore von 2 (Studie 301: Mobilitätsscore = 2; Beleuchtungsstärke 125 Lux) sowie für die Beleuchtungsstärken 200 und 250 Lux der Mobilitätsscore von 1 codiert (Studie 301: Mobilitätsscore = 1; Beleuchtungsstärke 250 Lux) [1,20].

#### *Ergebnisse der Validierungsstudie*

Zur Entwicklung des Tests wurden keine Daten vorgelegt. Aus den Protokollversionen der Studie 301 geht hervor, dass die Beleuchtungsstärken in Abhängigkeit von Ergebnissen einer vorherigen Auswertung der Test-Retest-Reliabilität ausgewählt wurden.

Alle normalsichtigen Personen (n = 29) konnten den Test bei 1 Lux, bei der niedrigsten Beleuchtungsstufe, bestehen. Die Anzahl an verwendeten Beleuchtungsstärken bei sehbeeinträchtigten Personen lässt auf eine schiefe Verteilung schließen: Personen wurden

überwiegend bei 1 Lux (n = 22) bzw. bei 4 und 10 Lux (jeweils n = 13) getestet und nur wenige Personen bei 50 Lux (n = 9), 100 Lux (n = 6), 150 Lux (n = 3) und bei 400 Lux (n = 3), keine Person bei 200 bzw. 250 Lux [20]. Es kann für die Beleuchtungsstärken 200 Lux bzw. 250 Lux, da keine Person unter diesen Beleuchtungsstärken den Test absolviert hat, keine abschließende Validitäts- und Reliabilitätsbeurteilung durchgeführt werden.

In der Validierungsstudie deuten die Ergebnisse auf eine gute Reliabilität, eine ausreichende Konstruktvalidität für die Genauigkeits- und Zeitkriterien zum Bestehen des Tests und Änderungssensitivität hin. Der Test konnte zwischen gesunden Probanden und Teilnehmenden mit Seheinschränkungen unterscheiden. Eine abschließende Beurteilung des Tests bei den beiden Beleuchtungsstärken 200 und 250 Lux, die später zu 250 Lux zusammengefasst wurden, konnte nicht durchgeführt werden. Deckeneffekte sind aufgrund der hohen Anzahl an durchgeführten Tests bei hohem Mobilitätsscore (bzw. niedriger Beleuchtungsstärke) zu beobachten [20].

In der Publikation von Chung et al. [1] wurde als eine der Einschränkungen des Fragebogens angegeben, dass es keine logarithmische Abstufung in den Leuchtstufen gab. Die Abstände orientierten sich stattdessen an Beleuchtungsstärken aus dem Alltag.

### *Zusammenfassung*

Die Testdurchführung wurde anhand der SOP umfassend beschrieben und lässt auf ein standardisiertes Vorgehen schließen, das auch die Teilnahme von Kindern berücksichtigt. Auch ist die verblindete Bewertung bei offenem Studiendesign zu begrüßen. Zudem wurden Maßnahmen zur Qualitätssicherung während der Studie bei der verblindeten Bewertung der Testdurchläufe und eine Randomisierung der Hindernisparcours zur Reduktion von Lerneffekten ergriffen. In der Validierungsstudie konnte die Reliabilität, Validität und Änderungssensitivität nachgewiesen werden, wenn auch diese mit Unsicherheiten in einigen Faktoren verbunden ist (u. a. Entwicklung des Parcours, Deckeneffekte, unzureichende Untersuchung bei einzelnen Beleuchtungsstärken/Mobilitätsscores).

### **Sehschärfe (ETDRS-, HOTV-Sehtafel)**

Der Endpunkt Sehschärfe wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

Für die Untersuchung der Sehschärfe wurden die Sehtafel der Early Treatment of Diabetic Retinopathy Study (ETDRS) und der HOTV-Test angewandt.

#### ETDRS-Sehtafel

Die ETDRS-Sehtafel ist aus 14 Sehzeichenreihen mit jeweils 5 Buchstaben in einer Reihe aufgebaut (insgesamt 70 Buchstaben), wobei die Größe der Buchstaben mit jeder Reihe abnimmt (Abbildung 4). Die Sehschärfe wurde für den Vergleich zwischen den Behandlungsgruppen in den Logarithmus des minimalen Auflösungswinkels (logMAR) umgerechnet. Ein abnehmender logMAR-Wert ist somit mit einer besseren Sehschärfe gleichzusetzen [27].

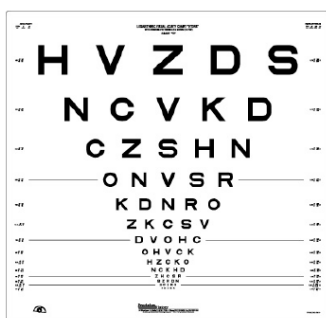


Abbildung 4: Beispiel einer ETDRS-Tafel [27]



### HOTV-Test

Beim HOTV-Test werden die Buchstaben H, O, T, V verwendet, die um eine vertikale Achse zentriert sind und somit auch von Kindern leichter identifiziert werden können. Die Größe der Buchstaben nimmt mit jeder Reihe ab. Die Sehschärfe kann ebenfalls in logMAR-Werte umgewandelt werden [27]. In der SOP wird als Beispiel eine HOTV-Sehtafel beschrieben, bei der jeweils nur 2 Buchstaben pro Zeile aufgeführt werden (Abbildung 5) [27].



Abbildung 5: Beispiel einer HOTV-Tafel [27]

### *Beschreibung*

#### **Studie 301**

Es sollte die best-korrigierte Sehkraft untersucht werden. Der Testaufbau erfolgte standardisiert nach Empfehlung der SST Kapitel 9.3 [30]. Der Test wurde im nicht-dilatierten Zustand durchgeführt. Es wurden abhängig von den kognitiven Fähigkeiten des Kindes die ETDRS- oder die HOTV-Sehtafel eingesetzt. Eine feste Alterseinstufung war nicht vorgesehen. Konnte das Kind den Test nicht mit der ETDRS-Sehtafel durchführen, wurde die HOTV-Sehtafel analog verwendet. Die entsprechende Sehtafel wurde im weiteren Studienverlauf angewandt. Falls bei einem Kind, bei dem die HOTV-Sehtafel ab Baseline verwendet wurde, im Studienverlauf die ETDRS-Sehtafel angewandt werden konnte, wurden beide Sehtafeln eingesetzt. Angaben, inwieweit die Auswertung dadurch beeinflusst wurde, konnten nicht identifiziert werden [27].

Bei der Durchführung und Auswertung wurde nicht zwischen den Sehtafeln differenziert. Alle Teilnehmenden der Studie begannen den Test mit ihrer Sehtafel in einem Abstand von 4 m und reduzierten dann den Abstand auf 2 m. Sofern nicht mindestens 3 der 5 Optotypen (Sehzeichen) bei diesen Entfernungen erkannt werden konnten, wurde die Entfernung auf 0,5 m reduziert. Falls die Sehzeichen bei dieser Entfernung ebenfalls nicht erkannt werden konnten, wurde ein Test auf niedrige Sehschärfe durchgeführt. Sollte die Testperson auch hier das Sehzeichen nicht identifizieren können, wurde ein Test auf Handbewegung durchgeführt. Falls die Person diesen Test ebenfalls nicht bestehen konnte, wurde überprüft, ob Hell- und Dunkelunterschiede wahrgenommen werden konnten [27].

Falls die Teilnehmenden den Test aus anderen Gründen als der Sehkraft nicht durchführen konnten, wurde der Test als nicht reliabel bzw. das Testergebnis als nicht zuverlässig eingestuft und nicht bewertet. Ein Score wurde berechnet, sofern die Patientin / der Patient mindestens 3 Optotypen korrekt erkennen konnte. Falls die Person bei keiner Distanz mindestens 3 der 5 Optotypen erkennen konnte, wurde die Distanz verwendet, bei der ein Optotyp identifiziert werden konnte [27].

Der Test sollte, um reliable Ergebnisse zu erzielen, zu Baseline und den anderen Visiten mehrfach durchgeführt werden. Eine einheitliche Reihenfolge zur Durchführung der Funktionstests (u. a. FST, Perimetrie, Sehschärfe) war nicht vorgesehen.

### **Studie LTFU**

Gemäß Modul 4 war die Operationalisierung vergleichbar mit Studie 301. Gemäß Studienprotokoll wurde für die Wirksamkeitsendpunkte ein Bewertungsparadigma (im Studienprotokoll als „evaluation paradigm“ bezeichnet), bezogen auf die Auswahl der Tests und deren Durchführung, individuell an die Patientinnen und Patienten angepasst, abhängig von ihren individuellen Zielen, dem Grad der Sehfunktion, dem Grad des Verständnisses der Durchführungen, der Ausprägung des Nystagmus und der Kooperationsfähigkeit und Geduld bei längeren Testsitzungen.

#### *Bewertung*

Die Operationalisierung des Endpunkts ist überwiegend nachvollziehbar.

Kritisch anzumerken ist die nicht-verblindete Durchführung und Bewertung der Testergebnisse durch die Untersucherin / den Untersucher bei offenem Studiendesign. In der SOP wurde als Beispiel für die HOTV-Sehtafel eine Tafel mit jeweils 2 Optotypen pro Zeile gezeigt, obwohl bei der Durchführung und Auswertung gemäß der SOP nicht zwischen den beiden Sehtafeln (5 Buchstaben je Zeile bei der ETDRS) unterschieden wurde.

Version 3 der SOP wurde am 13. Januar 2014 erstellt, mehr als ein Jahr nachdem die erste Person am 15. November 2012 in die Studie eingeschlossen wurde. Änderungen in den Versionen der SOP können nicht nachvollzogen werden, da hierzu keine Daten eingereicht wurden.

#### Patientenrelevanz

Die Sehschärfe wird als unmittelbar patientenrelevant angesehen.

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“.

#### Validität

Die ETDRS-Sehtafel ist ein etabliertes Instrument und wird als valide und geeignet zur Erhebung und Untersuchung der Sehschärfe angesehen. Der HOTV-Test ist ein verbreitetes Instrument zur Messung der Sehschärfe bei Kindern. Es bleibt unklar, inwieweit beide Tafeln austauschbar sind und somit eine gemeinsame Auswertung als adäquat angesehen werden kann.

Es werden Angaben zur Schulung des Studienpersonals, das mit der Durchführung des Tests beauftragt war, als auch zur standardisierten Durchführung gemacht. Kritisch anzumerken ist die Durchführung und Bewertung des Testergebnisses bei offenem Studiendesign. Eine verblindete Auswertung war nicht vorgesehen.

### **Lichtempfindlichkeit: Vollfeld-Lichtempfindlichkeits-Schwellenwert-Test (FST)**

Der Endpunkt Lichtempfindlichkeit wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

### **Studie 301/LTFU**

Der Vollfeld-Lichtempfindlichkeits-Schwellenwert-Test (Full-light Sensitivity Test, FST) hat das Ziel, die subjektive Lichtempfindlichkeit des gesamten Gesichtsfeldes, bei der die Testperson noch sehen kann, zu erfassen.

Der Test an beiden Augen sollte zunächst viermal mit weißem Licht wiederholt werden und sofern ein Testergebnis von der Untersucherin / dem Untersucher als zuverlässig (in der SOP

als reliabel definiert) eingestuft wurde, wurde der Test zweimal mit rotem und blauem Licht durchgeführt [29]. In der Studie 301 wurden weiße, rote und blaue Lichtreize für jedes Auge einzeln getestet. Die Reihenfolge der Lichtintensität wurde durch das Zufallsprinzip festgelegt. Es wurde ein Ganzfeld-Elektroretinogramm (ERG) angewandt, mit dem die Leuchtdichte eines Lichtblitzes, den die Testperson noch sehen kann, erfasst wird. Die Augen wurden zunächst erweitert. Der Test startete nach einer 40-minütigen Eingewöhnungsphase an die Dunkelheit. Es wurde jedes Auge einzeln getestet, indem das kontralaterale Auge verdeckt wurde. Zu Beginn der Prüfung blinkte eine Lampe in der gesamten Kuppel, begleitet von einem Piepton. Jedes Mal, wenn ein Piepton ertönte, musste die/der Teilnehmende durch Drücken einer „Ja“- oder „Nein“-Taste angeben, ob das Licht gesehen wurde oder nicht. Die Lichtblitze wurden mit unterschiedlichen Intensitäten (bis zu 80 db Differenz) fortgesetzt. Ein Algorithmus identifizierte die minimale Leuchtdichte (Helligkeit), bei der die Testperson zuverlässig Licht wahrnahm. Die Leuchtdichte wurde in einen logarithmischen Wert umgewandelt. Für  $\log_{10}$  (candela (cd)  $s/m^2$ ) entspricht ein negativeres Ergebnis einem niedrigeren Schwellenwert und damit einer verbesserten Lichtempfindlichkeit, was auf eine verbesserte Fotorezeptorfunktion hinweist [29].

Die Untersucherin / der Untersucher konnte weitere Tests veranlassen, sofern sie/er der Meinung war, dass die Patientin / der Patient in der Lage war, ein besseres Testergebnis zu erzielen. Die Ermittlung eines finalen Scores wurde in der SOP beschrieben. Es gingen nur Testergebnisse in die Auswertung ein, die einen Reliabilitätsscore  $> 0$  zur Bewertung der Zuverlässigkeit des Messergebnisses aufwiesen [29].

Sowohl die Software als auch die Untersucherin / der Untersucher konnten einen Wert für den Testdurchlauf ermitteln. Falls kein reliabler Wert durch die Software ermittelt werden konnte, sollte der Untersucher bzw. die Untersucherin im Anschluss den Score manuell über eine graphische Zeichnung ermitteln. Zur Bestimmung des Grenzwertes wurde in diesem Fall eine Kurve gezeichnet und anschließend im Graphen der Schwellenwert abgelesen. Der Reliabilitätsscore diente der Bewertung der Zuverlässigkeit bzw. Reliabilität der Messung. Hierbei erfolgte zunächst eine graphische Überprüfung der Entfernung der ermittelten Datenpunkte um den Schwellenwert zur berechneten Kurve. Der Reliabilitätsscore konnte zwischen 0 und 4 liegen, wobei ein hoher Wert eine geringe Entfernung der ermittelten Datenpunkte zur berechneten Kurve bedeutet und somit mit einer hohen Reliabilität assoziiert war. Der Untersucher bzw. die Untersucherin konnte weitere Tests veranlassen, sofern er bzw. sie der Meinung war, dass der Patient bzw. die Patientin in der Lage war, ein besseres Testergebnis zu erzielen. Ein besseres Testergebnis war assoziiert mit einer höheren Reliabilität der Messung.

Die Untersucherin / der Untersucher wurden gemäß GCP und der SOP geschult [29].

Gesonderte Maßnahmen zur Durchführung bei Kindern waren nicht vorgesehen. Auch erfolgten die Durchführung und die Auswertung nicht verblindet.

Eine einheitliche Reihenfolge zur Durchführung der Funktionstests (u. a. FST, Perimetrie, Sehschärfe) war nicht vorgesehen.

Der pU berücksichtigt in Modul 4 nur die Ergebnisse des FST mit weißem Licht, nicht jedoch mit rotem und blauem Licht. Diese werden in den Studienunterlagen dargestellt.

### **Studie LTFU**

Gemäß Modul 4 war die Operationalisierung vergleichbar mit Studie 301. Gemäß Studienprotokoll wurde für die Wirksamkeitsendpunkte ein Bewertungsparadigma (im Studienprotokoll als „evaluation paradigm“ bezeichnet), bezogen auf die Auswahl der Tests und deren Durchführung, individuell an die Patientinnen und Patienten angepasst, abhängig von ihren individuellen Zielen, dem Grad der Sehfunktion, dem Grad des Verständnisses der

Durchführungen, der Ausprägung des Nystagmus und der Kooperationsfähigkeit und Geduld bei längeren Testsitzungen.

#### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Es liegen Informationen zur Schulung der Untersucherinnen/Untersucher, Überprüfung der Zuverlässigkeit der Messung und der konkreten Testdurchführung vor.

Kritisch anzumerken ist die unverblindete Durchführung und Bewertung der Testergebnisse durch die Untersucherin / den Untersucher bei offenem Studiendesign, zumal der Test abhängig von der Einschätzung durch die Untersucherin / den Untersucher wiederholt werden konnte, bis akzeptable Reliabilitätswerte erreicht wurden. In der SOP sind Kriterien zur Reliabilitätsbewertung beschrieben. Bei der Entscheidung zur wiederholten Messung zur Erreichung eines zuverlässigeren Testergebnisses können neben den in der SOP angegebenen Kriterien zur Reliabilitätsbeurteilung auch das subjektive Ermessen des Untersuchers / der Untersucherin einen Einfluss ausüben.

Version 3 der SOP wurde am 13.01.2014 erstellt, mehr als ein Jahr nachdem die erste Person am 15.11.2012 in die Studie eingeschlossen wurde. Änderungen in den Versionen der SOP können nicht nachvollzogen werden, da hierzu keine Daten eingereicht wurden.

#### Patientenrelevanz

Die Lichtempfindlichkeit ist ein patientenrelevanter Endpunkt.

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“.

#### Validität

Nach Angaben des pU ist der Test geeignet zur Untersuchung der Lichtempfindlichkeit bei Personen mit Nystagmus, kleinem Gesichtsfeld und schlechter Sehschärfe. Untersuchungen zur Reliabilität und Konstruktvalidität liegen für Erwachsene, nicht jedoch für Kinder vor. Informationen zur Änderungssensitivität konnten nicht identifiziert werden. Eine abschließende Beurteilung unter Berücksichtigung der pädiatrischen Patientenpopulation ist daher derzeit nicht möglich.

Die Ergebnisse der Messung mit weißem, rotem und blauem Licht werden trotz der oben genannten Limitationen dargestellt. Es werden Angaben zur Schulung des Studienpersonals, das mit der Durchführung des Tests beauftragt war, als auch zur standardisierten Durchführung gemacht. Kritisch anzumerken ist die Durchführung und Bewertung der Testergebnisse bei offenem Studiendesign. Eine verblindete Auswertung war nicht vorgesehen.

Wenn auch die Veranlassung einer wiederholten Messung mit einer Verbesserung der Reliabilität assoziiert war und Kriterien zur Reliabilitätsbewertung beschrieben sind, konnte die Entscheidung zur Wiederholung des Tests auch nach subjektivem Ermessen des Untersuchers bzw. der Untersucherin erfolgen, weshalb dieses Vorgehen in Anbetracht des offenen Studiendesigns als kritisch gesehen wird.

## **Perimetrie nach Goldmann bzw. Humphrey**

Die Endpunkte Perimetrie nach Goldmann und Humphrey werden in der Nutzenbewertung berücksichtigt.

### Operationalisierung

#### *Beschreibung*

#### **Studie 301**

Zur Messung des Gesichtsfeldes wurden statische (nach Humphrey) und kinetische (nach Goldmann) Untersuchungsmethoden eingesetzt. Die Goldmann-Perimetrie erfasst das gesamte Gesichtsfeld, während die Humphrey-Perimetrie sich auf bestimmte Regionen im Gesichtsfeld fokussiert.

Die Durchführung von nur einem Perimetrie-Test (nach Goldmann oder Humphrey) je Person war ausreichend. Der Test sollte zu allen Visiten durchgeführt werden, wobei die Durchführung beider Perimetrie-Tests während der gesamten Studie zu bevorzugen war [16].

Beide Tests sollten an nicht-dilatierten Augen und unter Ausgleich der Refraktion an beiden Augen durchgeführt werden.

Die Untersucherin / der Untersucher waren in der Durchführung der beiden Perimetrie-Arten hinsichtlich der SOP geschult [3]. Gemäß Modul 4 wurde für die beiden Perimetrie-Arten ein spezifisches Bewertungsparadigma, bezogen auf die Auswahl der Tests und deren Durchführung (z. B. Einsetzen des Stimulus V4e oder III4e bei der kinetischen Perimetrie nach Goldmann), individuell an die Patientinnen und Patienten angepasst, abhängig von ihren individuellen Zielen, dem Grad der Sehfunktion, dem Grad des Verständnisses der Durchführungen, der Ausprägung des Nystagmus und der Kooperationsfähigkeit und Geduld bei längeren Testsitzungen.

#### Perimetrie nach Goldmann

Bei den kinetischen Untersuchungsmethoden nach Goldmann wird ein Ziellicht als visueller Reiz auf die Innenfläche einer Halbkugel projiziert. Die Intensität des Ziellichts bleibt konstant, das Ziellicht ist jedoch beweglich. Das Ziellicht wird von außerhalb der Gesichtsfeldgrenze in das vermutete Gesichtsfeld hineinbewegt und der Ort der erstmaligen Wahrnehmung wird dokumentiert [26].

Laut Modul 4 des Nutzendossiers wurden die auf der Halbkugel gewonnenen Messwerte in Form eines Isopters (Linie der optischen Wahrnehmung von Punkten bei gleicher visueller Empfindlichkeit) auf eine ebene Fläche projiziert. Diese ebene Fläche bildet graphisch das Gesichtsfeld ab und ist in 24 Meridiane eingeteilt. Für jeden Meridian wird die Gradzahl der ermittelten Grenzen des Gesichtsfeldes bestimmt und schließlich die Gesamtsumme ermittelt.

Als Stimuli wurden der Stimulus V4e (Größe: 64 mm<sup>2</sup>, Leuchtdichte: 315 cd/m<sup>2</sup>) und III4e (Größe: 4 mm<sup>2</sup>, Leuchtdichte: 315 cd/m<sup>2</sup>) verwendet. Zu Baseline wurde zunächst mit beiden Stimuli gemessen. Bei den Folgevisiten wurde der Test mit dem Stimulus III4e (1/16 kleinere Fläche als der Stimulus V4e) begonnen. Wenn dieser geeignet war, eindeutige bzw. zuverlässige Antworten zu produzieren, wurden weiterhin die Tests mit diesem Stimulus durchgeführt. Der Stimulus V4e wurde daraufhin nicht verwendet [26].

Die Zuverlässigkeit des Tests wurde durch die Untersucherin / den Untersucher bewertet. Als ein zuverlässiges Testergebnis war in der SOP definiert, wenn die Person den Teststimulus erkannte oder wenn die Person nicht in der Lage war, das Testobjekt zu sehen, und es dadurch nicht möglich war, den Isopter zu definieren. Ein Testergebnis galt als nicht zuverlässig, wenn die Person nicht in der Lage war, den Test aus anderen Gründen als der Sehkraft durchzuführen [26].

### Perimetrie nach Humphrey

In Modul 4 des Nutzendossiers wurde zur Perimetrie nach Humphrey aufgeführt, dass der Aufbau des Humphrey-Perimeters dem des Goldmann-Perimeters entspricht. Im Gegensatz zur Goldmann-Perimetrie wurde die Humphrey-Perimetrie automatisiert durchgeführt. Gemäß MOP wurde stets der Stimulus V verwendet.

Die Reize befanden sich an einer festgelegten Position in dem zu untersuchenden Gesichtsfeld und die Lichtintensität der Reize wurde durch einen Computeralgorithmus variiert. Die Testperson drückte einen Knopf, wenn der Stimulus wahrgenommen wurde, während die entsprechende Leuchtdichte notiert wurde.

Es wurde zunächst der Test im Bereich der Fovea und anschließend im Bereich der Makula durchgeführt [8]. Da die Prüfmedikation in eine bestimmte Region (Makula) des Auges appliziert wurde, wurden bei der Humphrey-Perimetrie Unterschiede in der Funktion in dieser Region vor und nach Applikation der Prüfmedikation untersucht. Die Applikation in die Fovea-Region sollte vermieden werden, weshalb ebenfalls eine Untersuchung in dieser Region vorgenommen wurde.

Sofern die Person nicht in der Lage war, den Test durchzuführen, aus Gründen, die nicht in Zusammenhang mit der Sehkraft standen, oder das Testergebnis als nicht zuverlässig angesehen werden konnte, wurde der Test nicht bewertet. Als ein nicht zuverlässiges Testergebnis wurde angesehen, wenn das verwendete Instrument für den Testdurchlauf Fehler dokumentierte (beispielsweise fehlende Reaktionsfähigkeit) [28].

Eine einheitliche Reihenfolge zur Durchführung der Funktionstests (u. a. FST, Perimetrie, Sehschärfe) war nicht vorgesehen. Angaben zu einer möglichen Anzahl an Wiederholungen des perimetrischen Tests, ähnlich wie bei den anderen Funktionstests (wie beispielsweise FST), konnten ebenfalls nicht identifiziert werden.

### **Studie LTFU**

Gemäß Modul 4 war die Operationalisierung vergleichbar mit Studie 301. Gemäß Studienprotokoll wurde für die Wirksamkeitsendpunkte ein Bewertungsparadigma (im Studienprotokoll als „evaluation paradigm“ bezeichnet), bezogen auf die Auswahl der Tests und deren Durchführung, individuell an die Patientinnen und Patienten angepasst, abhängig von ihren individuellen Zielen, dem Grad der Sehfunktion, dem Grad des Verständnisses der Durchführungen, der Ausprägung des Nystagmus und der Kooperationsfähigkeit und Geduld bei längeren Testsitzungen.

### *Bewertung*

Die Operationalisierung der Endpunkte ist überwiegend nachvollziehbar. Es wurden Informationen zur Überprüfung der Zuverlässigkeit einer Messung und der konkreten Testdurchführung gemacht. Kritisch anzumerken ist die Durchführung und Bewertung des Testergebnisses bei offenem Studiendesign. Eine verblindete Auswertung war nicht vorgesehen.

Version 3 (Perimetrie nach Humphrey) und Version 2 (Perimetrie nach Goldmann) der beiden MOP wurden am 13.01.2014 erstellt, nachdem die erste Person am 15.11.2012 in die Studie eingeschlossen wurde. Änderungen in den Versionen der MOP können nicht nachvollzogen werden, da hierzu keine Daten vorgelegt wurden.

### Patientenrelevanz

Die Gesichtsfeldmessung ist ein patientenrelevanter Endpunkt.

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“.

### Validität

In einer orientierenden Recherche konnten Publikationen zur Reliabilität beider Perimetrie-Arten bei Kindern identifiziert werden. In der Studie von Patel et al., 2011 [15] wurde eine gute Durchführbarkeit für beide Perimetrie-Verfahren bei Kindern beschrieben. Die erhobenen Reliabilitätskriterien bei der Perimetrie nach Humphrey zeigen jedoch Einschränkungen in der Zuverlässigkeit der Messungen bei Kindern an.

Die Ergebnisse beider Verfahren werden trotz der oben genannten Limitationen dargestellt. Es werden Angaben zur Schulung des Studienpersonals, das mit der Durchführung des Tests beauftragt war, als auch zur standardisierten Durchführung gemacht. In Anbetracht des offenen Studiendesigns wäre zumindest eine verblindete Auswertung der Testergebnisse wünschenswert. Es liegen nur eingeschränkte Informationen zur Reliabilität vor. Angaben zu Validität und Änderungssensitivität liegen seitens des pU nicht vor, jedoch handelt es sich um weit verbreitete Methoden zur Untersuchung des Gesichtsfeldes.

### **2.3.3 Lebensqualität**

#### **PRO-Instrument zur Erfassung der visuellen Lebensqualität**

Der Endpunkt modifizierter VFQ-25 wird in der Nutzenbewertung nicht berücksichtigt aufgrund nicht gegebener Validität.

#### Operationalisierung

##### *Beschreibung*

#### **Studie 301/LTFU**

Die visuelle Lebensqualität wird mit Hilfe eines Fragebogens untersucht. Der pU gibt an, dass eine Neuentwicklung auf Basis des Fragebogens National Eye Institute Visual Function Questionnaire 25 (NEI-VFQ-25) für die Studie 301 durchgeführt wurde, um eine auch für Kinder altersgerechte Erfassung der Lebensqualität zu ermöglichen, da sich manche Fragen des NEI-VFQ-25 an erwachsenen Patienten und Patientinnen orientieren (u. a. Fragen zum Autofahren und zur Arbeitsfähigkeit). Bei dem NEI-VFQ-25 handelt es sich um ein etabliertes, reliables und validiertes Instrument zur Erfassung der visuellen Lebensqualität. Er besteht aus 25 Fragen sowie aus 11 optionalen Zusatzfragen. Die Antwortkategorien variieren zwischen 5–6 Antwortoptionen. Diese werden anschließend umgewandelt auf eine Skala von 0–100, wobei 100 keine Einschränkungen und ein normales Sehvermögen darstellt.

Der vom pU adaptierte Fragebogen besteht ebenfalls aus 25 Fragen, die von den Patienten bzw. Patientinnen und ihrem Elternteil/Erziehungsberechtigten in einem Bezugszeitraum von 1 Monat beantwortet werden sollen. Sofern notwendig konnten die Fragen laut vorgelesen werden, wobei Vorgaben zur Vorleseart definiert wurden. Die Fragebögen wurden während der Studienvisite ausgefüllt. Es gab keine zeitliche Begrenzung zur Beantwortung der Fragebögen, wobei dies bis zum Ende der Studienvisite erfolgen sollte.

Die Unterschiede zwischen den beiden Fragebögen gemäß dem Dossier Modul 4 sind zum einen, dass im vom pU entwickelten Fragebogen keine Unterteilung mehr zwischen den Abschnitten „General Health and Vision“, „Difficulties with Activities“ und „Responses to Vision Problems“ gibt. Bei der adaptierten Version wird gemäß dem pU der Fokus auf die Domäne „Difficulties with Activities“ gelegt. Zudem wird im vorherigen Fragebogen eine Skala zur Beantwortung der Fragen von 5 bzw. 6 Punkten herangezogen. Bei dem für die Studie 301 verwendeten Fragebogen wird für Items eine Skala von 0–10 eingesetzt und anschließend der Durchschnitt der Werte gebildet. Der Gesamtwert des Patienten bzw. der Patientin bzw. des Elternteils/Erziehungsberechtigten werden nicht gemittelt, sondern separat über die Zeit ermittelt.

### *Bewertung*

Die Operationalisierung ist nachvollziehbar.

### Patientenrelevanz

Die visuelle Lebensqualität aus Patientensicht wird als patientenrelevant angesehen.

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“.

### Validität

Die Fragen der beiden Fragebögen unterscheiden sich erheblich in ihrer Formulierung. Es bleibt unklar, nach welchen Kriterien die Fragen im vom pU entwickelten Fragebogen ausgewählt bzw. entwickelt wurden. Zusätzlich wurden die Antwortoptionen und die Struktur des Fragebogens geändert. Vom pU angeführte Validierungsstudien beziehen sich auf den bereits etablierten NEI-VFQ-25. Studien zum vom pU entwickelten Fragebogen konnten in den eingereichten Unterlagen nicht identifiziert werden. Stattdessen gibt der pU an, die Ergebnisse der Validierungsstudien des NEI-VFQ-25 auf den neu entwickelten Fragebogen übertragen zu wollen.

Aufgrund der erheblichen Unterschiede im Aufbau des Fragebogens, der Formulierung der Fragen scheint eine Übertragbarkeit der psychometrischen Eigenschaften vom NEI-VFQ-25 auf den vom pU entwickelten Fragebogen nicht möglich.

## **2.3.4 Sicherheit**

### **Sicherheit**

Der Endpunkt Sicherheit wird in der Nutzenbewertung berücksichtigt.

### Operationalisierung

#### *Beschreibung*

#### **Studie 301**

UE wurden durch Befragung oder durch den Patienten bzw. die Patientin selbst berichtet bzw. durch Beobachtung und/oder körperliche Untersuchung erfasst. Die UE wurden gemäß MedDRA-Kodierung (Version 14.0 oder höher) nach SOC und PT kodiert. Als UE wurden alle unvorhergesehenen oder nachteiligen Ereignisse definiert, unabhängig ob sie in Zusammenhang mit der Prüfmedikation standen oder nicht. Alle unerwünschten Ereignisse wurden von der medizinischen Aufsicht und dem DSMB überprüft.

In der Studie 301 werden UE für die Interventionsgruppe ab der Injektion ins erste Auge und für die Kontrollgruppe ab dem ersten Tag der Baseline-Untersuchung bis 30 Tage nach Ausscheiden aus der Studie berichtet.

Als schwerwiegendes unerwünschtes Ereignis (SUE) waren UE definiert, die

- zum Tod
- zu einem lebensbedrohlichen Ereignis
- zu einem stationären Krankenhausaufenthalt oder zu einem verlängerten Krankenhausaufenthalt
- zu einer langanhaltenden oder signifikanten Unfähigkeit oder erheblichen Beeinträchtigung der Ausführung von alltäglichen Aktivitäten oder
- zu einer angeborenen Anomalie bzw. einem Geburtsfehler führten.



Ereignisse, die den Teilnehmenden gefährden oder medizinische, chirurgische Interventionen, die zur Prävention eines oben aufgeführten Ereignisses durchgeführt werden, können auf Grundlage einer medizinischen Einschätzung als SUE angesehen werden.

Der pU gibt im Studienprotokoll folgende studienindividuelle Einteilung der Schweregrade an:

- Mild: Symptome sind für die/den Teilnehmende/n kaum spürbar und haben keinen Einfluss auf ihre/seine Leistung oder Funktion.
- Moderat: Symptome mit ausreichendem Schweregrad, sodass Teilnehmende Beschwerden haben bzw. in der Verrichtung von täglichen Aktivitäten eingeschränkt sind. Eine Behandlung der Symptome könnte notwendig sein.
- Schwer: Symptome verursachen schwere Beschwerden. Eine Behandlung der Symptome und/oder Hospitalisierung kann erfolgen.
- Lebensgefährlich: Das Auftreten eines Ereignisses kann zu unmittelbarer Todesgefahr führen.
- Tod: Das Auftreten eines Ereignisses führt zum Tod.

Neben der weiter oben beschriebenen studienindividuellen Einteilung nach Schweregraden sollte, falls anwendbar, eine Schweregradeinteilung (Grad 1–4) gemäß der WHO-Toxizitätsskala erfolgen. Vom pU wurden ergänzend Schweregrade für ophthalmologische UE in Orientierung an die Toxizitätsskala vorgenommen.

Im Dossier Modul 4 werden neben den a priori beschriebenen schweren UE gemäß Studienprotokoll weitere Daten zu UE ab Grad  $\geq 3$  berichtet. Im SAP findet sich nur die Information, dass eine Auswertung und Darstellung der UE Grad  $\geq 3$  möglich ist, sofern dies notwendig sei. Es konnten jedoch keine weiteren Informationen identifiziert werden. Es bleibt unklar, inwieweit die Darstellung der UE Grad  $\geq 3$  mit der studienindividuellen Schweregradeinteilung aus den Studienunterlagen durch den pU bzw. der WHO-Toxizitätsskala zusammenhängt oder ob es sich um eine neue Schweregradeinteilung handelt.

In der Studie 301 wurden auffällige Laborparameter nur als UE berichtet, sofern sie Grad 2 oder höher eingestuft wurden oder falls diese nach erneutem Testen bestehen blieben. Sofern Leukozytose und Hyperglykämie während der Gabe von Glukokortikoiden (nach Absetzen der Glukokortikoide) auftraten, wurden diese nur als UE berichtet, sofern die UE als Grad 3 oder höher eingestuft wurden. Bereits vor Baseline vorhandene Beschwerden wurden nur als UE bei einer Verschlechterung und/oder bei häufigerem Auftreten berichtet.

Es konnten im Studienprotokoll als auch im SAP keine a priori definierten UE von besonderem Interesse identifiziert werden.

#### **Unterschiede in der LTFU-Studie**

Es sollen alle klinischen UE überwacht werden. Jedoch soll der Fokus der Dokumentation gemäß aktuellem Studienprotokoll V3.0 (30. Mai 2019) auf folgende UE gelegt werden:

- SUE
- UE, die möglicherweise oder wahrscheinlich in Verbindung mit der Administration des Prüfpräparats stehen
- Neue oder sich verschlechternde UE in einer der 4 Kategorien (onkologische Ereignisse, hämatologische Ereignisse, neurologische Ereignisse und/oder Autoimmunerkrankungen). Diese werden post hoc im Interimsstudienbericht als UE von besonderem Interesse (AESI) bezeichnet.

Die oben genannten UE sollen im CRF dokumentiert werden. Jegliche andere nicht schwerwiegenden UE werden in den Patientenunterlagen (u. a. Aufzeichnungen des Prüfpersonals zu den Visiten sowie Labor- und retinale/Sehfunktionstests) erfasst.

Gemäß SAP V1.0. (22. Juni 2016) war zunächst eine vollständige Überwachung der UE bis Jahr 1B (ein Jahr nach Injektion ins zweite Auge) für beide Behandlungsgruppen vorgesehen. Erst im Anschluss nach Jahr 1B war nur die Erhebung der oben genannten UE beschrieben.

Die Definition eines SUE entsprach der Definition der Studie 301.

Gemäß Studienprotokoll wird der Schweregrad der UE wie folgt klassifiziert:

- Mild: Symptome sind für die/den Teilnehmende/n kaum spürbar und haben keinen Einfluss auf ihre/seine Leistung oder Funktion.
- Moderat: Symptome mit ausreichendem Schweregrad, sodass Teilnehmende Beschwerden haben bzw. in der Verrichtung von täglichen Aktivitäten eingeschränkt sind.
- Schwer: Symptome verursachen schwere Beschwerden. Eine Behandlung der Symptome und/oder Hospitalisierung kann erfolgen.

Die Schweregradeinteilung beinhaltet gemäß Modul 4 eine Einteilung wie in Studie 301. Die Beurteilung erfolgte studienspezifisch durch das Prüfpersonal.

Um zwischen den ursprünglichen Behandlungsarmen vergleichbare Beobachtungszeiträume zugrunde zu legen, wurden gemäß Modul 4 die Sicherheitsendpunkte für die ursprünglichen Studienarme der Studie 301 jeweils zwischen der ersten Injektion und dem Datum der zweiten Injektion + 365 Tage („Jahr 1“) sowie zwischen erster Injektion und dem Datum der zweiten Injektion + 1.825 Tage („Jahr 5“) ausgewertet. Im Studienbericht erfolgt eine Darstellung der UE von Tag 0A (Tag erster Injektion) bis Jahr 1B (Jahr 1 nach zweiter Injektion), von Jahr 1B bis Datenschnitt und Tag 0A bis Datenschnitt.

### *Bewertung*

Die Operationalisierung ist für beide Studien bedingt nachvollziehbar

In der Studie 301 werden in der Interventionsgruppe – abweichend von der Kontrollgruppe und dem üblichen Vorgehen einer zeitgleichen Erfassung – UE nicht ab Baseline, sondern erst ab der ersten Injektion berichtet. Hierdurch werden UE in der Interventionsgruppe zwischen Baseline und Injektion ins erste Auge (Median: 22 Tage (min; max: 6; 78)) nicht erfasst. Kritisch anzumerken ist die in der Studie 301 vorgesehene Erfassung einzelner UE (Leukozytosen, Hyperglykämien) erst ab einem Schweregrad  $\geq 3$  während der Einnahme von Glukokortikoiden.

In der Studie LTFU wird zwar die Erfassung von UE in Patientendaten beschrieben, der Fokus der Dokumentation im CRF liegt jedoch auf bestimmten UE. Es bleibt unklar, inwieweit im Studienbericht / Dossier Modul 4 nur bestimmte UE beschrieben wurden bzw. ab welchem Zeitraum.

### Patientenrelevanz

Das Auftreten von unerwünschten Ereignissen wird als patientenrelevant angesehen.

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“.

### Validität

#### **Studie 301**

Die Erhebung der Sicherheitsendpunkte wird in der Studie 301 als nur zum Teil valide angesehen. Kritisch anzumerken ist der unterschiedliche Beginn der UE-Erfassung in beiden Behandlungsgruppen. In der Studie 301 erfolgte in der Interventionsgruppe die Erfassung, anders als in der Kontrollgruppe (ab Baseline), erst ab der Injektion ins erste Auge. Der

Zeitraum zwischen Baseline und Injektion betrug im Median 22 Tage (min; max: 6; 78). In diesem Zeitraum wurden UE in der Interventionsgruppe, anders als in der Kontrollgruppe, nicht berichtet.

Es werden im Modul 4 zwei Darstellungen der schweren UE berichtet. Es erfolgte neben der a priori vorgesehenen studienindividuellen Schweregradeinteilung im Modul 4 eine Darstellung der Ergebnisse von UE Grad  $\geq 3$ . Im SAP wird nur angegeben, dass, sofern notwendig, eine Darstellung der UE Grad  $\geq 3$  vorgenommen werden kann. Es bleibt jedoch unklar, inwieweit die studienindividuelle Schweregradeinteilung und die WHO-Toxizitätsskala mit der Auswertung der UE Grad  $\geq 3$  zusammenhängen oder ob eine neue Klassifizierung vorgenommen wurde. Die Ergebnisse wurden nur im Dossier Modul 4 berichtet.

Für die studienindividuelle a priori beschriebene Schweregradeinteilung konnten keine Angaben zur Begründung der Auswahl des Klassifizierungssystems identifiziert werden.

Die WHO-Toxizitätsskala sollte, sofern möglich, ebenfalls angewandt werden. Der pU gibt zudem an, dass in Orientierung an diese auch eine Beurteilung der ophthalmologischen UE geplant war. Inwieweit diese valide ist, kann nicht abschließend beurteilt werden, da Kriterien für die Schweregradeinteilung der ophthalmologischen UE nicht identifiziert werden konnten.

Zusätzlich kann es durch die Erfassung von zwei UE (Leukozytose und Hyperglykämie) ab einem bestimmten Schweregrad für den Zeitraum der Glukokortikoid-Einnahme in der Interventionsgruppe zu einer im geringen Maße unvollständigen Erhebung der UE kommen.

In der Nutzenbewertung werden die Ergebnisse zu allen UE, SUE, schweren UE beschrieben. Aufgrund der unklaren Klassifizierung der UE Grad  $\geq 3$  werden diese ergänzend im Anhang dargestellt.

#### **Studie LTFU**

In der Studie LTFU sollten alle UE überwacht werden, es war jedoch keine vollständige Dokumentation der unerwünschten Ereignisse im CRF für den gesamten Beobachtungszeitraum vorgesehen. Gemäß SAP V1.0 (22. Juni 2016) sollten ab Jahr 1B nur bestimmte UE im CRF dokumentiert werden. Im Studienprotokoll ab Version 2.0 (15. Juni 2018) konnten diese Angaben nicht mehr identifiziert werden. Es sollte die Dokumentation bestimmter UE im CRF erfolgen. Alle anderen UE wurden in Patientenakten dokumentiert. Sofern eine Auswertung der Sicherheitsdaten nur auf Grundlage der Daten im CRF erfolgte, kann nicht von einer vollständigen UE-Erfassung ausgegangen werden. Zudem bleibt unklar, inwieweit UE in der ehemaligen Kontrollgruppe zwischen dem Zeitpunkt der Injektion ins erste Auge (Tag 0A) bis zur Visite Jahr nach der zweiten Injektion (Jahr 1B), ähnlich wie in der ehemaligen Interventionsgruppe zu Beginn der Studie 301, vollständig erhoben wurden.

Es werden die Ergebnisse für die post hoc durchgeführten Analysen aus Dossier Modul 4 (UE für den Zeitraum Injektion ins erste Auge bis Jahr 1B und Injektion ins erste Auge bis Jahr 5) dargestellt. Bis Jahr 5 war in beiden Behandlungsgruppen der Anteil an Personen  $> 70\%$ . Für die ehemalige Interventionsgruppe gehen die Ergebnisse der Sicherheitsdaten aus der Studie 301 mit in die Auswertung und Darstellung ein.

### **2.3.5 Übersicht der Erhebungszeitpunkte**

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

Tabelle 7: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie 301/LTFU

Studienvisite Endpunkt	Studie 301/LTFU										
	Screen- ing <sup>1)</sup>	Base line <sup>2)</sup>	Tag 0A/ 0B <sup>2)3)4)</sup>	Tag 1A/ 1B <sup>4)</sup>	Tag 3A/ 3B <sup>4)</sup>	Tag 14B <sup>4)</sup>	Tag 30 B/ C <sup>5)</sup>	Tag 90 B/ C <sup>5)</sup>	Tag 180 B/ C <sup>5)</sup>	Jahr 1B/ C <sup>5)6)</sup>	Jahr 2-15 <sup>7)</sup>
MLMT	x	x		-		-	x	x	x	X <sup>8)</sup>	x
Lichtempfindlichkeit (FST)		x					x	x	x	x	x
Perimetrie nach Goldmann	x	x					x	x	x	x	x
Perimetrie nach Humphrey	x	x					x	x	x	x	x
Sehschärfe (ETDRS-, HOTV-Sehtafel)	x	x		x <sup>9)</sup>	x <sup>9)</sup>	x <sup>9)</sup>	x	x	x	x	x
Unerwünschte Ereignisse <sup>10)</sup>		x	x	x	x	x	x	x	x	x	x

<sup>1)</sup> Abschluss der Screeningvisite  $\leq 90$  Tage vor Baseline. Die Untersuchungen konnten an mehreren Tagen stattfinden. Die Dauer der Untersuchungen zur Überprüfung der Ein-/Ausschlusskriterien und Festlegung der Randomisierung betrug 6–8 Stunden.

<sup>2)</sup> Die Dauer aller Untersuchungen zu Baseline betrug 8–12 Stunden und konnte an 2–4 Tagen erfolgen. Untersuchungen zur visuellen Funktion konnten an mehreren Tagen, dafür mit einer kürzeren Testdauer erfolgen. Subjektive Tests, wie Sehschärfe oder Gesichtsfeldmessung, sollten mehrmals wiederholt werden, um reliable Baseline-Werte zu erhalten.

<sup>3)</sup> Tag 0A: Applikation des Prüfpräparats ins erste Auge innerhalb von 90 Tagen nach Baseline. Tag 0B: Applikation des Prüfpräparats ins zweite Auge innerhalb von  $12 \pm 6$  Tagen nach Applikation ins erste Auge. Für die Interventionsgruppe der Studie 301 bezieht sich der Startpunkt ab Applikation ins erste Auge zu Beginn der Studie 301. Für die ursprüngliche Kontrollgruppe der Studie 301 bezieht sich der Startpunkt nach Ende der Studie 301 und Beginn der Studie LTFU.

<sup>4)</sup> A: Injektion ins erste Auge; B: Injektion ins zweite Auge

<sup>5)</sup> Tag xB / Jahr xB: Visite zu Tag x bzw. Jahr x nach Behandlung des zweiten Auges in der Interventionsgruppe der Studie 301 bzw. ehemaligen Kontrollgruppe nach Ende der Studie 301 und Erhalt des Prüfpräparats; Tag xC / Jahr xC: Visite zu Tag x bzw. Jahr x nach Baseline in der Kontrollgruppe (Studie 301)

<sup>6)</sup> Nach einem Jahr kann, sofern die Ein-/Ausschlusskriterien der Studie 301 für Teilnehmende der Kontrollgruppe weiterhin zutreffen, das Prüfpräparat appliziert werden. Für diese Personen konnte die Visite zu Jahr 1C als Baseline-Visite dienen. Ähnlich wie zu Beginn der Studie 301 konnte der Zeitraum zwischen Baseline-Visite und Injektion ins erste Auge (Tag 0A)  $\leq 90$  Tage betragen.

<sup>7)</sup> Jahr 2–15: Einmal jährliche Erhebung einer oder mehrerer Wirksamkeitsendpunkte und der Sicherheit nach Injektion ins zweite Auge. Ursprünglich war gemäß Studienbericht der Studie 301 die Erhebung ab Jahr 6 optional per Telefon geplant. In den aktuelleren Studienunterlagen der Studie LTFU ist dieser Passus nicht mehr zu finden.

<sup>8)</sup> Studie 301: Primärer Endpunkt: Untersuchung der Änderung im MLMT-Score bilateral gemessen zu Jahr 1B/C im Vergleich zu Baseline zwischen den Behandlungsgruppen.

<sup>9)</sup> Überwachung der Genesung nach operativem Eingriff (Applikation des Prüfpräparats)

<sup>10)</sup> Erfassung der UE war für die ehemalige Interventionsgruppe in der Studie 301 ab Injektion ins erste Auge vorgesehen.

Abkürzungen: ETDRS: Early Treatment of Diabetic Retinopathy Study; FST: Full-Light-Sensitivity-Test; Jahr xB: Jahr x nach Behandlung des zweiten Auges; Jahr 1C: Visite zu Jahr 1 nach Baseline für die Kontrollgruppe der Studie 301; LTFU: Long-Term Follow-up; MLMT: Multi-Luminanz-Mobilitäts-Test; Tag 0A: Applikation des Prüfpräparats ins erste Auge innerhalb von 90 Tagen nach Baseline/Voruntersuchung; Tag 0B: Applikation des Prüfpräparats ins zweite Auge innerhalb von  $12 \pm 6$  Tagen nach Applikation ins erste Auge; UE: Unerwünschtes Ereignis

## 2.4 Statistische Methoden

### Relevante Änderungen in den Versionen des statistischen Analyseplans

#### Studie 301

Es existieren 4 Versionen des SAP (Version 4: 19. August 2015, Version 3: 30. Juni 2015, Version 2: 10. März 2014, Version 1: 25. Januar 2013 bzw. gemäß CSR der LTFU-Studie 12. April 2013). Die Versionen wurden nach Einschluss der ersten Person (15. November 2012) erstellt und stellen umfangreiche Änderungen des SAP dar, die nach Rücksprache mit der EMA bzw. FDA erfolgten. Version 4 wurde erst nach der Datenbanksperrung vom 24. August 2015 erstellt auf Grundlage eines Schreibens der FDA vom 21. Juli 2015.

Relevante Änderungen für die in der Nutzenbewertung dargestellten Endpunkte in Version 4 (19. August 2015): Es sollten die statistischen Analysen auf Grundlage aller randomisierten Personen erfolgen.

#### Studie LTFU

Es wurde eine Version des statistischen Analyseplan V1.0 (22. Juni 2016) als Anhang des Studienberichts Addendum 2016 veröffentlicht.

### Analysepopulationen

#### Studie 301

- ITT (Intention-to-treat) Population: Alle Patientinnen und Patienten, die randomisiert und in die Studie aufgenommen wurden.
- mITT (modified ITT) Population: Alle randomisierten Patientinnen und Patienten, die nicht ausgeschieden sind, bevor die folgenden Personen die Behandlungszuteilung wussten: Patient/in, Elternteil, primäre/r Prüfärzt/in, medizinischer Monitor.
- Sicherheitspopulation: In der Sicherheitspopulation sind alle Personen der Interventionsgruppe, die die Prüfmedikation in beide Augen erhalten haben. Für die Kontrollgruppe wird ein ähnlicher Passus beschrieben wie für die mITT-Population: alle Personen der Kontrollgruppe, die nicht ausgeschieden sind, bevor die folgenden Personen die Behandlungszuteilung wussten: Patient/in, Elternteil, primärer Prüfärzt/in, medizinische Aufsicht.

Gemäß dem Dossier Modul 4 umfassten die mITT- und Sicherheitspopulation die gleiche Population.

#### Studie LTFU

Es wird die mITT-Population bzw. Sicherheitspopulation der Studie 301 als Analysepopulation definiert.

### Datenschnitte

In der Studie 301 war nur ein Datenschnitt (16. Juli 2015) vorgesehen.

In der Studie LTFU waren jährliche Datenschnitte vorgesehen:

- Studienbericht Addendum 2016 vom 12. Januar 2017 (Datenschnitt: 18. Mai 2016)
- Studienbericht Addendum 2017 vom 15. Februar 2018 (Datenschnitt: 05. Mai 2017)
- Studienbericht Addendum 2018 vom 15. Februar 2019 (Datenschnitt: 02. Juli 2018)
- Studienbericht Addendum 2019 vom 01. Mai 2020 (Datenschnitt: 05. Juni 2019)
- Studienbericht Addendum 2020 vom 30. Juli 2021 (Datenschnitt: 30. Juni 2020)

Für die Nutzenbewertung wird der Datenschnitt vom 30. Juni 2020 herangezogen.

## Präspezifizierung geplanter und durchgeführter Analysen

### Studie 301

#### *Definition Baseline*

Für die Wirksamkeitsanalysen in der Behandlungsgruppe wurde für die Berechnung des zeitlichen Abstands zwischen einer Studienvsitede und Applikation der Tag der Injektion ins zweite Auge (Tag 0B) herangezogen. Für die Sicherheitsanalysen der Interventionsgruppe wurde der Tag der ersten Injektion (Tag 0A) herangezogen. Für Analysen der Kontrollgruppe wurde stets der erste Tag der Baseline-Untersuchungen (Tag 0C) verwendet.

Für die Wirksamkeitsendpunkte wurden Analysen bilateral gemessen gegenüber monokular erhobenen Ergebnissen bevorzugt.

#### *Endpunkte MLMT und Perimetrie*

Für den Endpunkt MLMT wurde die mittlere Differenz der beobachteten Veränderungen zu Jahr 1B/C im Vergleich zu Baseline zwischen den Behandlungsgruppen untersucht. Es wurde, wie im SAP a priori definiert, der exakte zweiseitige p-Wert mittels Wilcoxon-Rangsummentest berechnet und zusätzlich mittels Permutationstest verifiziert.

Im Dossier wurde post hoc die Differenz der mittleren Veränderungen zu Baseline und das zugehörige Konfidenzintervall mittels eines gemischten Modells mit Behandlung und Studienvsitede als Terme berechnet. Es fehlen weitere Angaben zum Modell, wie beispielsweise Angaben zur Kovarianzstruktur. Die standardisierte Mittelwertdifferenz (SMD) (+ 95%-KI) wurde nach Hedges' g berechnet. Zudem werden Responderanalysen für eine nicht-validierte Responseschwelle von 1 Punkt auf dem Mobilitätsscore vorgelegt.

Für die Endpunkte Perimetrie nach Goldmann und Humphrey erfolgte gemäß Dossier Modul 4 bzw. Studienbericht post hoc die Auswertung analog zum MLMT. Für die Perimetrie nach Humphrey wurden Responderanalysen für eine nicht-validierte Responseschwelle von 3 bzw. 5 dB für eine Verbesserung/Verschlechterung vorgelegt.

#### *Endpunkte Lichtempfindlichkeit und Sehschärfe*

Für beide Endpunkte wurde präspezifiziert folgendes Modell verwendet: Für die mittlere Veränderung, gemittelt über beide Augen, zu Jahr 1B/C im Vergleich zu Baseline wurde ein Mixed Model for Repeated Measures (MMRM) verwendet, das die Terme Behandlung, Studienvsitede und Behandlung\*Studienvsitede einschloss. Es wurden alle Studienvsiteden inklusive der 5 Messungen pro Visite (Baseline, Tage 30B/C, 90 B/C, 180 B/C und Jahr 1 B/C) für das Modell berücksichtigt. Werte zu Tag 30 B/C, 90 B/C, 180 B/C wurden im Modell nur zur Verbesserung der Präzision verwendet. Als „within subject“-Korrelation wurde eine unstrukturierte Kovarianzstruktur im Wiederholungsterm („repeated“) verwendet.

Für den Endpunkt Lichtempfindlichkeit, gemessen mit FST mit weißem Licht, wurde zusätzlich post hoc in Modul 4 eine standardisierte Mittelwertdifferenz (+ 95%-KI) nach Hedges' g berechnet.

In Modul 4 wurden für die Sehschärfe post hoc Responderanalysen für beide Sehtafeln (ETDRS und HOTV) zusammen und separat, gemittelt über beide Augen, zu Jahr 1B/C im Vergleich zu Baseline berichtet. Für die Beurteilung der Wirksamkeit der Behandlung wurde die Anzahl der Personen erfasst, die zu Jahr 1B/C eine Verbesserung um  $\geq 10$  und  $\geq 15$  Buchstaben (Abnahme logMAR um  $\geq 0,2$  und  $\geq 0,3$  Einheiten) und eine Verschlechterung um  $\geq 10$  bzw.  $\geq 15$  Buchstaben in der Sehschärfe erhoben mittels ETDRS bzw. HOTV (Zunahme logMAR um  $\geq 0,2$  und  $\geq 0,3$  Einheiten) im Vergleich zu Baseline erreicht hatten. Für die separate Darstellung beider Sehtafeln wurden Responderanalysen für eine Verbesserung um  $\geq 10$  bzw.

≥ 15 Buchstaben eingereicht. Für den Vergleich der beiden Gruppen wurde das relative Risiko mit den dazugehörigen 95%-KI berechnet; die 95%-KI mittels exakter Methode. Für den Vergleich der Interventions- und Kontrollgruppe wurde der exakte Test nach Fischer zum Signifikanzniveau 0,05 durchgeführt.

#### *Analyse Sicherheit*

Gemäß SAP war eine deskriptive Analyse geplant. Es wurden im Dossier post hoc Effektschätzer (Odds Ratio, relatives Risiko und Risikodifferenz) mittels Vier-Felder-Tafel sowie p-Werte mittels dem exakten Test nach Fischer berechnet. Das KI wurde mit einer exakten Methode (Tail-Methode) berechnet durch Umkehren zweier separater einseitiger Tests, bei der die Größe des Tests höchstens  $\alpha/2$  und das Konfidenzintervall mindestens  $(1-\alpha)$  ist.

#### Studie LTFU

##### *Definition Baseline*

Gemäß Addendum 2016 des 301-Studienberichts wird der Begriff „Injektions-Baseline“ eingeführt:

- Für Patientinnen und Patienten der ehemaligen Interventionsgruppe der Studie 301 gilt jeweils der Wert/Score der Baseline-Visite der Studie 301 als Baseline.
- Für die ehemalige Kontrollgruppe der Studie 301 gilt folgende Baseline-Definition:
  - Injektion ins erste Auge (Tag 0A) zu Beginn der LTFU: Untersuchungen (Vitalzeichen, Laboruntersuchungen für Urin und Blut). Bei mehreren Tests innerhalb des Visitenfensters (Tag 0A) wird der letzte Test vor Injektion als Baseline definiert.
  - Jahr 1C (= Jahr 1 nach der Baseline-Visite der Kontrollgruppe in der Studie 301): Wirksamkeitsuntersuchungen ohne Daten zu Tag 0A

Gemäß Modul 4 werden als „Studien-Baseline“ die Baseline-Werte zu Beginn der Studie 301 definiert.

##### *Auswertung der Wirksamkeitsanalysen und der Sicherheit*

Für Teilnehmende werden die Ergebnisse des MLMT, des FST, der Perimetrieverfahren und des Sehschärfetests zu Jahr 1, Jahr 2 und fortfolgend bis Jahr 7 im Vergleich zu Baseline ausgewertet und deskriptiv dargestellt.

Die Ergebnisse zur Sicherheit wurden ebenfalls nur deskriptiv ausgewertet.

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

##### Studie 301

MLMT-bilateral: Bei fehlenden Werten war ein konservativer Umgang bzw. waren Imputationen vorgesehen.

- Bei Personen, die aus der Studie ausschieden, wurde angenommen, dass sich diese nicht verbesserten. Diesen Personen wurde in der Analyse eine Änderung von 0 zur Visite Jahr 1B/C zugewiesen.

FST, Sehschärfe: Für die primäre Auswertung waren keine Imputationen vorgesehen.

- Sofern der Wert zu Baseline oder zu Jahr 1B/C fehlte, konnte die mittlere Änderung der jeweiligen Behandlungsgruppe zwischen Jahr 1 und Baseline eingesetzt werden.
- Falls beide Werte (Baseline und Jahr 1B/C) für einen der beiden Endpunkte fehlte, konnte für den fehlenden Änderungswert die mittlere Änderung in der Kontrollgruppe eingesetzt werden.

- Sofern der Baseline-Wert beim Endpunkt Sehschärfe fehlte, konnte der Screening-Wert eingesetzt werden.

Perimetrie nach Goldmann, Humphrey:

- Es konnten keine Imputationsregeln im SAP identifiziert werden.

### Studie LTFU

Nur für den MLMT war bei fehlenden Werten das Vorgehen analog zur Studie 301 vorgesehen.

### **Einschätzung der statistischen Auswertungen**

Für die Studie 301 werden die a priori geplanten Analysen für den MLMT und die beiden Perimetrie-Verfahren mittels Wilcoxon-Rangsummen-Test zur Berechnung des p-Wertes und eines gemischten Modells zur Berechnung der mittleren Differenz inklusive 95%-KI sowie für die Lichtempfindlichkeit und Sehschärfe (ETDRS- und HOTV-Sehtafel) das MMRM-Modell als geeignet angesehen.

Die Responderanalysen zur Verbesserung oder Verschlechterung um  $\geq 10$  bzw.  $\geq 15$  Buchstaben werden zur Beurteilung der Sehschärfe, ermittelt mit ETDRS-Tafel, als geeignet angesehen. Für die HOTV-Tafel ist die Eignung einer klinischen Relevanzschwelle von 10 bzw. 15 Buchstaben unklar. Angaben zur Validität einer MID von  $\geq 10$  oder  $\geq 15$  Buchstaben für die HOTV-Tafel konnten nicht identifiziert werden. Für die Auswertung mit beiden Sehtafeln zusammen wird daher die Auswertung mittels MMRM herangezogen, für die ETDRS-Tafel Responderanalysen für eine Verbesserung oder Verschlechterung um  $\geq 10$  bzw.  $\geq 15$  Buchstaben.

Die Randomisierung erfolgte stratifiziert nach Alter ( $\geq 10$  Jahre versus  $< 10$  Jahre) und Ergebnislevel des MLMT (Bestehen bei  $\geq 125$  Lux versus  $< 125$  Lux), während die statistischen Analysen gemäß dem pU aufgrund der geringen Fallzahl (N = 31) und damit verbundener möglicher Konvergenzproblemen nicht stratifiziert durchgeführt wurden.

## **2.5 Verzerrungspotential auf Studien- und Endpunktebene**

*Tabelle 8: Verzerrungspotential der Studie 301*

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenteilung	Verblindung Patientin/Patient	Verblindung Behandlungsperson	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
Studie 301	Ja	Ja	Nein <sup>1)</sup>	Nein <sup>1)</sup>	Nein	Unklar <sup>2)3)</sup>	Hoch

<sup>1)</sup> Offenes Studiendesign

<sup>2)</sup> Bei folgender Begleitmedikation war der Anteil an Personen in der Interventionsgruppe höher als in der Kontrollgruppe (Unterschied mind. 40 %): Schmerzmittel, Propionsäure-Derivate, Anticholinergika, Kortikosteroide + Antiinfektiva in Kombination (siehe Ausführungen Tabelle 12). Da dies auf die Applikation des Prüfpräparats in der Interventionsgruppe im Gegensatz zum beobachtenden Abwarten in der Kontrollgruppe zurückzuführen ist, ergeben sich keine Hinweise auf ein erhöhtes Verzerrungspotential.

<sup>3)</sup> Es finden sich Hinweise auf Unterschiede in der Studiendurchführung zwischen den Zentren, die in der Gesamtschau zu einer Verzerrung der Ergebnisse führen könnten: Es konnten mehr Abweichungen von den SOP/MOP im Studienzentrum in Iowa im Vergleich zum Studienzentrum CHOP beobachtet werden. Dies



spiegelt sich auch beim höheren Anteil an Personen mit jeglichen Protokollabweichungen/-verletzungen in der Interventionsgruppe (81 %) im Vergleich zur Kontrollgruppe (50 %) wider, die größtenteils auf Abweichungen in der Vorgehensweise (SOP/MOP) zurückzuführen waren (Unterschied zwischen den Gruppen: 47 %). Diese Unterschiede können auch durch die ungleiche Verteilung der Kontrollpopulation (N = 10) auf die Studienzentren bedingt sein (CHOP: n = 8 vs. Iowa: n = 2), während diese in der Interventionsgruppe ähnlich war (CHOP: n = 11 vs. Iowa: n = 10), die sich aufgrund der Blockrandomisierung ergeben könnte.

Abkürzungen: CHOP: Children's Hospital of Philadelphia; SD: Standardabweichung; MOP: Manual Operating Procedures; SOP: Standard Operating Procedure

Das Verzerrungspotential auf Studienebene wird als hoch eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 9 dargestellt.

*Tabelle 9: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie 301*

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
MLMT	Nein <sup>1)2)</sup>	Ja	Nein	Unklar <sup>3)4)</sup>	Hoch
Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test (FST)	Nein <sup>1)5)</sup>	Nein <sup>6)</sup>	Nein	Ja <sup>4)7)8)</sup>	Hoch
Sehschärfe (ETDRS-, HOTV-Sehtafel)	Nein <sup>1)5)</sup>	Ja <sup>9)</sup>	Nein	Ja <sup>4)7)10)</sup>	Hoch
Perimetrie nach Goldmann	Nein <sup>1)5)</sup>	Ja <sup>9)</sup>	Nein	Ja <sup>4)7)11)</sup>	Hoch
Perimetrie nach Humphrey	Nein <sup>1)5)</sup>	Ja <sup>9)</sup>	Nein	Ja <sup>4)7)11)</sup>	Hoch
Sicherheit	Nein <sup>1)</sup>	Ja	Nein	Ja <sup>4)12)13)14)</sup>	Hoch

<sup>1)</sup> Offenes Studiendesign

<sup>2)</sup> Die Beurteilenden des Tests waren verblindet.

<sup>3)</sup> Gemäß dem Dossier traten prozedurbedingte Abweichungen vom Verfahrenshandbuch beim MLMT am häufigsten im Interventionsarm auf (Baseline vs. Jahr 1: 3 vs. 2). Im Kontrollarm trat nur zu Baseline eine prozedurbedingte Abweichung auf. Angaben zur Anzahl an Personen, bei denen die genannten Abweichungen auftraten, konnten nicht identifiziert werden.

<sup>4)</sup> Es war gemäß SAP keine Analyse mit Berücksichtigung der Randomisierungsstrata aufgrund der geringen Fallzahl vorgesehen. Es wird im vorliegenden Fall von keinem erhöhten Verzerrungspotential ausgegangen.

<sup>5)</sup> Die Auswertung erfolgte nicht verblindet.

<sup>6)</sup> Gemäß pU war eine Analyse anhand der ITT-Population (Intervention: N = 21; Kontrolle: N = 10) für den Endpunkt FST gemessen mit rotem und blauem Licht vorgesehen. Zu Jahr 1 waren Daten von 17 bzw. 9 Personen in der Interventions- bzw. Kontrollgruppe vorhanden. Informationen, weshalb Ergebnisse von 3 bzw. 1 Person in der Interventions- bzw. Kontrollgruppe nicht berichtet wurden, konnten nicht identifiziert werden.

<sup>7)</sup> Version 3 (bzw. Version 2 bei der Perimetrie nach Goldmann) der SOP wurde am 13.01.2014 erstellt, nachdem die erste Person am 15.11.2012 in die Studie eingeschlossen wurde. Änderungen in den SOP-Versionen können nicht nachvollzogen werden, da hierzu keine Daten vorgelegt wurden.

<sup>8)</sup> Der Test konnte abhängig von der Einschätzung durch die Untersucherin / den Untersucher wiederholt werden, bis akzeptable Reliabilitätswerte erreicht wurden. In der SOP sind Kriterien zur Reliabilitätsbewertung beschrieben. Bei der Entscheidung zur wiederholten Messung zum Erreichen eines zuverlässigeren Testergebnisses können neben den in der SOP angegebenen Kriterien zur

Reliabilitätsbeurteilung auch das subjektive Ermessen des/der Untersuchers/Untersucherin einen Einfluss ausüben. Der Einfluss auf das Verzerrungspotential ist unklar.

- <sup>9)</sup> Gemäß pU war eine Analyse anhand der ITT-Population vorgesehen (Intervention: N = 21; Kontrolle: N = 10). Es wurden jeweils von nur einer Person in beiden Behandlungsgruppen keine Daten für die statistische Analyse vorgelegt. Es waren Daten von  $\geq 90\%$  der randomisierten Personen vorhanden. Insgesamt wird im vorliegenden Fall in der Abweichung von der ITT-Population kein erhöhtes Verzerrungspotential gesehen.
- <sup>10)</sup> Die HOTV-Sehtafel wurde bei  $< 15\%$  der Personen (Intervention: n = 3; Kontrolle: n = 1) eingesetzt. Diese Information lag für das aktuelle Dossier nicht vor und wurde aus der Nutzenbewertung zum vorherigen Verfahren entnommen [4]. In der SOP bzw. dem SAP wurde bei der Beschreibung der Durchführung und Auswertung nicht zwischen den beiden Sehtafeln ETDRS und HOTV (und der Off-Chart-Bestimmung) unterschieden. Der pU legt keine Informationen bzw. Studien zur Austauschbarkeit beider Sehtafeln vor. Trotz der Unklarheiten werden Auswertungen beider Sehtafeln (und Off-Chart-Messungen) zusammen anhand eines gemischten Modells dargestellt. Insgesamt wird im vorliegenden Fall nicht von einem erhöhten Verzerrungspotential ausgegangen.
- <sup>11)</sup> Im SAP bzw. Studienprotokoll wurden keine a priori definierten statistischen Analyseverfahren beschrieben. Konkrete Angaben konnten erst post hoc im Studienbericht bzw. Modul 4 identifiziert werden. Es wird im vorliegenden Fall von keinem erhöhten Verzerrungspotential ausgegangen.
- <sup>12)</sup> Gemäß Studienprotokoll war a priori definiert, UE in der Interventionsgruppe erst ab der ersten Injektion zu erfassen, während in der Kontrollgruppe diese bereits ab Baseline/Randomisierung erfasst werden sollten. Die erste Injektion in der Interventionsgruppe fand im Median 22 Tage (min; max: 6; 78) nach Baseline statt. Hierdurch werden in der Interventionsgruppe keine UE zwischen Baseline und erster Injektion berichtet. Die Auswirkung auf das Verzerrungspotential ist unklar.
- <sup>13)</sup> Für das Dossier wurden keine Angaben zu Beobachtungszeiten vom Tag der ersten Injektion bzw. Baseline bis Jahr 1B/C eingereicht. Bezugnehmend auf das vorherige Verfahren beträgt die durchschnittliche Beobachtungszeit in der Interventionsgruppe 406,6 Tage (SD: 20,4), während diese in der Kontrollgruppe bei durchschnittlich 354,8 Tagen (SD: 20,6) lag [4]. Der Unterschied zwischen beiden Behandlungsgruppen betrug ca. 50 Tage. In der vorliegenden Konstellation wird davon ausgegangen, dass die vom pU im Dossier Modul 4 post hoc berechneten Effektschätzer für die Sicherheitsendpunkte angewendet werden können.
- <sup>14)</sup> Erfassung von bestimmten UE nur ab einem bestimmten Schweregrad im Interventionsarm: Sofern Leukozytose und Hyperglykämie während der Gabe von Glukokortikoiden (18–30 Tage, je nach Zeitpunkt der zweiten Injektion) auftraten, wurden diese nur als UE berichtet, falls sie als Grad 3 oder höher eingestuft wurden. Dies kann zu einer Verzerrung der Ergebnisse zugunsten der Intervention führen, da nur ab einem bestimmten Schweregrad die oben genannten UE berichtet werden.

Abkürzungsverzeichnis: ETDRS: Early Treatment of Diabetic Retinopathy Study; FST: Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test; ITT: Intention-to-treat; MLMT: Multi-Luminanz-Mobilitätstest; pU: pharmazeutischer Unternehmer; SAP: Statistischer Analyseplan; SD: Standardabweichung; SOP: Standard Operating Procedure; UE: Unerwünschte Ereignisse

### 3 Ergebnisse der eingeschlossenen Studien

#### 3.1 Studiencharakteristika und Studienmedikation

##### Studie 301

Für die Studie 301 wurden 36 Personen gescreent und anschließend 31 Personen (ITT-Population, Interventionsgruppe: N = 21 und Kontrollgruppe: N = 10) in die Studie aufgenommen (Tabelle 10).

Angaben zu Kriterien zur Auswahl des zuerst behandelten Auges finden sich erst im Studienbericht. Das erste Auge, das behandelt werden sollte, wurde in Abhängigkeit der Sehschärfe und/oder nach Patientenpräferenz ausgewählt. Bis auf eine Person, bei der die Sehschärfe zwischen den Augen ähnlich war, wurde das Auge als Erstes operiert, das die schlechtere Sehschärfe aufwies.

Es werden Angaben zu den Beobachtungszeiten von Screening bis Studienende im Dossier Modul 4 vorgegeben. Zudem soll die Erfassung der UE ab Screening geplant gewesen sein, obwohl in anderen Studienunterlagen die Erfassung erst ab Baseline beginnen sollte. Angaben zur medianen Beobachtungsdauer ab Baseline bzw. Injektion ins erste Auge konnten für beide Behandlungsgruppen nicht identifiziert werden.

Der Anteil an Personen mit einer Protokollverletzung war in der Interventionsgruppe mit 81 % höher als in der Kontrollgruppe (50 %). Die häufigsten Protokollverletzungen betrafen prozedurbedingte Abweichungen (SOP/MOP), die auf die Prozedur zur Durchführung der Seh- und Netzhautfunktionstests, nicht auf Abweichungen vom klinischen Protokoll zurückzuführen waren. Der Anteil war in der Interventionsgruppe mit 67 % deutlich höher als in der Kontrollgruppe mit 20 % (Tabelle 40). Im Studienbericht wurde berichtet, dass es vermehrt zu Abweichungen vom SOP/MOP im Studienzentrum in Iowa im Vergleich zum Studienzentrum CHOP kam. Gemäß Studienbericht wurde aufgrund dessen ein intensives Training/Management durch den Sponsor durchgeführt. Gemäß dem Dossier traten prozedurbedingte Abweichungen vom Verfahrenshandbuch beim MLMT am häufigsten in der Interventionsgruppe auf (Baseline vs. 1 Jahr: 3 Abweichungen vs. 2 Abweichungen). Im Kontrollarm trat nur zu Baseline eine prozedurbedingte Abweichung auf. Folgende Gründe wurden für die höhere Anzahl an Abweichungen genannt: Durchführung des MLMT bei bestimmten Beleuchtungsstärken nur mit jeweils einem Auge, jedoch nicht bilateral, Dopplung eines Tests sowie die mangelnde Verwertbarkeit einer Videoaufzeichnung, wodurch ein Test wiederholt werden musste.

Tabelle 10: Allgemeine Angaben; Studie 301

<b>Studie 301</b>	<b>Intervention: Voretigen Neparvovec</b>	<b>Kontrolle: Beobachtende s Abwarten</b>
ITT-Population <sup>1)</sup> , n (%)	21 (100,0)	10 (100,0)
mITT-Population <sup>2)</sup> , n (%)	20 (95,2)	9 (90,0)
Sicherheitspopulation <sup>3)</sup> , n (%)	20 (95,2)	9 (90,0)
Studie zu Jahr 1 nach Baseline beendet, n (%)	20 (95,2)	9 (90,0)
Abbruch der Studie, n (%)	1 (4,8) <sup>4)</sup>	1 (10,0) <sup>4)</sup>
Aufgrund von:		
Entscheidung Ärztin/Arzt	1 (4,8) <sup>5)</sup>	0 (0,0)
Entzug der Einverständniserklärung	0 (0,0)	1 (10,0)

<b>Studie 301</b>	<b>Intervention: Voretigen Neparvovec</b>	<b>Kontrolle: Beobachtende s Abwarten</b>
Jegliche Protokollabweichungen/-verletzungen, n (%)	17 (81,0)	5 (50,0)
Abweichungen von der Vorgehensweise (SOP/MOP) <sup>6)</sup>	14 (66,7)	2 (20,0)
Untersuchungen nicht im geforderten Zeitraum durchgeführt	5 (23,8)	2 (20,0)
Abweichungen im Zusammenhang mit der Medikation <sup>7)</sup>	5 (23,8)	1 (10,0)
Eine Untersuchung wurde nicht durchgeführt	4 (19,0)	0 (0,0)
Einschlusskriterien nicht erfüllt	1 (4,8)	0 (0,0)
Studiendauer in Tagen (Screening bis Studienende <sup>8)</sup> )		
Mittelwert (SD)	496 (58,3)	447 (67,0)
Median (min; max)	475 (412; 593)	420 (378; 568)
Beobachtungsdauer (Baseline/Randomisierung bis Jahr 1B/C)		
Mittelwert (SD)	k. A. <sup>9)</sup>	k. A. <sup>9)</sup>
Median (min; max)		
Zeit zwischen Screening und Baseline-Visite		
Mittelwert (SD)	k. A.	k. A.
Median (min; max)		
Zeit von Randomisierung/Baseline bis zur ersten Injektion in Tagen, n (%)		
Mittelwert (SD)	20 (95,2)	n. a.
Median (min; max)	34,3 (27,9) 22 (6; 78)	
Zeit zwischen erster und zweiter Injektion in Tagen, n (%)		
Mittelwert (SD)	20 (95,2)	n. a.
Median (min; max)	8,8 (2,6) 7 (7; 14)	

<sup>1)</sup> Alle Personen, die in eine der beiden Behandlungsgruppen randomisiert wurden.

<sup>2)</sup> Alle randomisierten Personen, die nicht ausgeschieden sind, bevor die folgenden Personen die Behandlungszuteilung wussten: Patient/in, Elternteil, primäre/r Prüfarzt/in, medizinischer Monitor. Die mITT- und Sicherheitspopulation waren identisch.

<sup>3)</sup> In der Sicherheitspopulation sind alle Personen der Interventionsgruppe, die die Prüfmedikation in beide Augen erhalten haben. Für die Kontrollgruppe wird ein ähnlicher Passus beschrieben wie für die mITT-Population: alle Personen der Kontrollgruppe, die nicht ausgeschieden sind, bevor die folgenden Personen die Behandlungszuteilung wussten: Patient/in, Elternteil, primäre/r Prüfarzt/in, medizinische Aufsicht.

<sup>4)</sup> Ausscheiden aus der Studie erfolgte vor Behandlungszuteilung und somit vor Erhalt des Prüfpräparats.

<sup>5)</sup> Es wurde eine schwere retinale Atrophie/Degeneration diagnostiziert.

<sup>6)</sup> Diese beinhalteten Abweichungen bei Untersuchungen der Retina bzw. zur Sehfähigkeit, nicht jedoch des klinischen Protokolls.

<sup>7)</sup> Abweichungen standen in der Interventionsgruppe in Zusammenhang mit dem Prednisolon/Prednison-Dosierungsregime und in der Kontrollgruppe mit einer über der im Protokoll erlaubten Dosis an Vitamin A.

<sup>8)</sup> Unterschiedliche Definitionen der Beobachtungszeiträume in beiden Behandlungsgruppen: Beobachtungszeitraum Interventionsgruppe: Screening bis Jahr 1B + 30 Tage. Beobachtungszeitraum Kontroll-/Interventionsgruppe: Screening bis Tag vor der ersten Injektion.

<sup>9)</sup> Es wurden für das aktuelle Dossier keine Beobachtungszeiten für beide Behandlungsgruppen eingereicht. Unter Berücksichtigung des vorherigen Verfahrens wird für die Interventionsgruppe eine durchschnittliche Beobachtungszeit für den Zeitraum der ersten Injektion bis 1 Jahr nach der zweiten Injektion (Jahr 1B) von 406,6 Tagen (SD: 20,4) angegeben bzw. lag der Zeitraum zwischen 367,0 und 436,0 Tagen. In der Kontrollgruppe lag der durchschnittliche Zeitraum zwischen Baseline und Visite zu Jahr 1 (Jahr 1C) bei 354,8 Tagen (SD: 20,6) bzw. zwischen 337,0 und 388,0 Tagen [4]. Mediane Beobachtungszeiten wurden nicht vorgelegt.

Abkürzungen: Jahr 1B: Jahr 1 nach Injektion ins zweite Auge (Studie 301: Interventionsgruppe); Jahr 1C: Jahr 1 nach Baseline (Studie 301: Kontrollgruppe); k. A.: keine Angabe; (m)ITT: (modified) Intention-to-Treat; MOP: Manual operating procedures; n. a.: nicht anwendbar; SD: Standardabweichung; SOP: Standard operating procedures

### Studie LTFU

In Tabelle 11 sind die allgemeinen Angaben zur Studie LTFU für Personen, die an der Studie 301 teilgenommen haben, dargestellt. Zum Datenschnitt vom 30. Juni 2020 haben 90 % der Teilnehmenden das Jahr 5B der ehemaligen Interventionsgruppe und 78 % der Teilnehmenden das Jahr 5B der ehemaligen Kontrollgruppe abgeschlossen.

Im Interimsstudienbericht zum aktuellen siebten Datenschnitt werden, seit dem vorherigen Datenschnitt, 6 Protokollabweichungen berichtet. Abweichungen von der Vorgehensweise (SOP/MOP) wurden bei 3 Teilnehmenden, keine Durchführung einer Untersuchung bei 2 Teilnehmenden und Durchführung einer Untersuchung nicht im geforderten Zeitraum bei 1 Teilnehmenden beschrieben. Im Interimsstudienbericht wird die Zuteilung der Personen mit Protokollabweichungen in eine der ursprünglichen Behandlungsgruppen nicht beschrieben.

*Tabelle 11: Allgemeine Angaben; LTFU-Studie; Datenschnitt 30. Juni 2020*

<b>Studie LTFU<sup>1)</sup></b>	<b>Ursprüngliche Intervention: Voretigen Neparvovec N = 20</b>	<b>Ursprüngliche Kontrolle: Beobachtendes Abwarten N = 9</b>
Studie 301 mITT <sup>2)</sup> , n (%)	20 (100,0)	9 (100,0)
Sicherheitspopulation <sup>3)</sup> , n (%)	20 (100,0)	9 (100,0)
Erhalt der Injektion in beide Augen, n (%)	20 (100,0) <sup>4)</sup>	9 (100,0) <sup>5)</sup>
Jahr 1B nach Injektion beendet <sup>3)</sup> , n (%)	20 (100,0)	9 (100,0)
Jahr 2B nach Injektion beendet, n (%)	20 (100,0)	9 (100,0)
Jahr 3B nach Injektion beendet, n (%)	20 (100,0)	8 (88,9)
Jahr 4B nach Injektion beendet, n (%)	20 (100,0)	8 (88,9)
Jahr 5B nach Injektion beendet, n (%)	18 (90,0)	7 (77,8)
Jahr 6B nach Injektion beendet, n (%)	13 (65,0)	1 (11,1)
Jahr 7B nach Injektion beendet, n (%)	2 (10,0)	0 (0,0)
Vorzeitiges Ausscheiden aus der Studie, n (%)	0 (0,0)	0 (0,0)
Jegliche Protokollabweichungen/-verletzungen seit Beginn der Studie 301, n (%)	19 (95,0)	7 (77,8)
Abweichungen von der Vorgehensweise (SOP/MOP)	15 (75,0)	4 (44,4)
Untersuchungen nicht im geforderten Zeitraum durchgeführt	6 (30,0)	3 (33,3)
Abweichungen im Zusammenhang mit der Medikation	5 (25,0)	3 (33,3)
Eine Untersuchung wurde nicht durchgeführt	9 (45,0)	1 (11,1)
Einschlusskriterien nicht erfüllt	1 (5,0)	0 (0,0)
Orientierung und Mobilisierungsuntersuchung außerhalb des Protokolls zu Jahr 2B	6 (30,0)	0 (0,0)
Durchführung Sehschärfe nur mit ETDRS, nicht mit beiden Sehtafeln	0 (0,0)	1 (11,1)
Vektor-Proben nicht bei ordnungsgemäßer Temperatur gelagert	0 (0,0)	1 (11,1)
MLMT-Test mit 14 anstelle 15 Hindernissen durchgeführt	0 (0,0)	1 (11,1)
Beobachtungsdauer von Randomisierung bis zum Datenschnitt (30. Juni 2020) in Tagen		
Mittelwert (SD)	2530 (112)	2600 (133)
Median (min; max)	2532 (2533; 2729)	2613 (2534; 2753)
Beobachtungsdauer von erster Injektion bis zum Datenschnitt (30. Juni 2020) in Tagen		
Mittelwert (SD)	2496 (133)	2204 (123)

<b>Studie LTFU<sup>1)</sup></b>	<b>Ursprüngliche Intervention: Voretigen Neparvovec N = 20</b>	<b>Ursprüngliche Kontrolle: Beobachtendes Abwarten N = 9</b>
Median (min; max)	2489 (2276; 2722)	2197 (1987; 2395)
Beobachtungsdauer von letzter Untersuchung/Visite vor Injektion <sup>6)</sup> und erster Injektion in Tagen Mittelwert (SD) Median (min; max)	35,3 (27,9) 22,5 (7,0; 79,0)	42,3 (28,4) 36 (15,0; 93,0)
Zeit zwischen erster und zweiter Injektion in Tagen Mittelwert (SD) Median (min; max)	9,8 (2,6) 8 (8,0; 15,0)	8,7 (1,4) 8 (8,0; 12,0)

<sup>1)</sup> Als Baseline definiert ist die Untersuchung vor der ersten Injektion. Diese entsprach in der ursprünglichen Interventionsgruppe dem Zeitpunkt der Baseline-Untersuchung zu Beginn der Studie 301. In der ursprünglichen Kontrollgruppe wurde als Baseline vor Injektion für die Wirksamkeitseindpunkte die Visite zu Jahr 1 der Studie 301 definiert.

<sup>2)</sup> Alle Personen, die die Studie 301 abgeschlossen haben, nehmen derzeit auch an der Verlängerungsstudie LTFU teil. 2 Personen wurden zu Beginn der Studie 301 ausgeschlossen und sind nicht Teil der Studienpopulation der LTFU-Studie.

<sup>3)</sup> Die Sicherheitspopulation entspricht der mITT-Population.

<sup>4)</sup> Erhalt der Injektion in beide Augen zu Beginn der Studie 301.

<sup>5)</sup> Erhalt der Injektion in beide Augen bei dem Wechsel zur LTFU-Studie.

<sup>6)</sup> Für den ehemaligen Interventionsarm entspricht dies der Beobachtungsdauer von Randomisierung/Baseline bis zur Injektion ins erste Auge zu Beginn der Studie 301. Für den ehemaligen Kontrollarm entspricht dies der Beobachtungsdauer der letzten Visite zu Jahr 1C bis zur Applikation der Injektion ins erste Auge.

Abkürzungen: ETDRS: Early Treatment of Diabetic Retinopathy Study; Jahr xB: Jahr x nach Erhalt der zweiten Injektion; LTFU: Long-Term Follow-up; mITT: modified Intention-to-Treat; MLMT: Multi-Luminanz-Mobilitätstest; MOP: Manual Operating Procedure; SD: Standardabweichung; SOP: Standard Operating Procedure

## Charakterisierung der Studienpopulation

### Studie 301

In Tabelle 12 sind die Charakteristika der Studienpopulation dargestellt. Der Anteil an Personen im Studienzentrum IOWA in der Interventionsgruppe war mit 20 % niedriger als im Studienzentrum CHOP.

In der medizinischen Vorgeschichte traten in beiden Behandlungsgruppen keine Hornhauttrübung, Glaukoma, Keratokonus oder okuläre Anomalien auf.

*Tabelle 12: Charakterisierung der Studienpopulation in Studie 301*

<b>Studie</b>	<b>Intervention: Voretigen Neparvovec N = 21</b>	<b>Kontrolle: Beobachtendes Abwarten N = 10</b>
<i>Alter zum Zeitpunkt der Randomisierung (Jahre)</i> Mittelwert (SD) Median (min; max)	14,7 (11,8) 11 (4; 44)	15,9 (9,5) 14 (4; 31)
<i>Geschlecht, n (%)</i> Männlich Weiblich	9 (42,9) 12 (57,1)	4 (40,0) 6 (60,0)

<b>Studie</b>	<b>Intervention: Voretigen Neparvovec N = 21</b>	<b>Kontrolle: Beobachtendes Abwarten N = 10</b>
<i>Abstammung (genetisch), n (%)</i>		
Nicht spanisch oder lateinamerikanisch	16 (76,0)	9 (90,0)
Spanisch oder lateinamerikanisch	5 (24,0)	1 (10,0)
<i>Region, n (%)</i>		
EMEA	1 (4,8)	4 (40,0)
Nordamerika	18 (85,7)	6 (60,0)
Lateinamerika	1 (4,8)	0 (0,0)
Asien/Pazifik	1 (4,8)	0 (0,0)
<i>Studienzentren, n (%)</i>		
Children's Hospital of Philadelphia (CHOP), USA	11 (52,4)	8 (80,0)
University of Iowa (IOWA), USA	10 (47,6)	2 (20,0)
<i>Stratifikationsfaktor: Alter, n (%)</i>		
Personen ≥ 10 Jahre	12 (57,1)	6 (60,0)
Personen < 10 Jahre	9 (42,9)	4 (40,0)
<i>Stratifikationsfaktor: Screeningergebnis des MLMT<sup>1)</sup>, n (%)</i>		
Bestehen bei < 125 Lux	12 (57,1)	4 (40,0)
Bestehen bei ≥ 125 Lux	9 (42,9)	6 (60,0)
<i>Verteilung der Personen, bei denen der MLMT bilateral bei der niedrigsten Beleuchtungsstärke noch bestanden wurde zu Baseline, n (%)</i>		
1	0 (0,0)	0 (0,0)
4	4 (19,0)	1 (10,0)
10	5 (23,8)	2 (20,0)
50	7 (33,3)	5 (50,0)
125	3 (14,3)	1 (10,0)
250	0 (0,0)	0 (0,0)
400	0 (0,0)	0 (0,0)
> 400 <sup>2)</sup>	2 (9,5)	1 (10,0)
<i>Medizinische Vorgeschichte: jeglicher Befund<sup>3)</sup>, n (%)</i>		
Allergie	19 (90,5)	8 (80,0)
Hals-Nasen-Ohren	9 (42,9)	5 (50,0)
Gastrointestinaltrakt	10 (47,6)	4 (40,0)
Urogenitales System	3 (14,3)	2 (20,0)
Hämatologisches System	4 (19,0)	3 (30,0)
Integumentsystem	4 (19,0)	0 (0,0)
Muskuloskelettales System	7 (33,3)	3 (30,0)
Muskuloskelettales System	5 (23,8)	3 (30,0)
Neurologisches System	5 (23,8)	4 (40,0)
Andere	7 (33,3)	2 (10,0)

Studie	Intervention: Voretigen Neparvovec N = 21		Kontrolle: Beobachtendes Abwarten N = 10	
	Erstes Auge	Zweites Auge	Erstes Auge	Zweites Auge
<i>Ophthalmologische Vorgeschichte</i>				
<i>Jeglicher Befund, n (%)</i>	21 (100,0)	21 (100,0)	10 (100,0)	10 (100)
Katarakt	1 (4,8)	1 (4,8)	1 (10,0)	1 (10,0)
Linsentrübung	1 (4,8)	1 (4,8)	2 (20,0)	1 (10,0)
Nystagmus	21 (100,0)	21 (100,0)	10 (100,0)	10 (100,0)
Retinale Anomalien	21 (100,0)	21 (100,0)	10 (100,0)	10 (100,0)
Strabismus	8 (38,1)	7 (33,0)	5 (50,0)	4 (20,0)
Andere	3 (14,3)	1 (4,8)	2 (20,0)	2 (20,0)

<sup>1)</sup> Es wurde das Auge mit dem schlechtesten Testergebnis verwendet.

<sup>2)</sup> Personen, die den Test bei 400 Lux nicht bestanden haben.

<sup>3)</sup> Es wurde die medizinische Vorgeschichte ab einer Häufigkeit von  $\geq 20\%$  in einer der beiden Behandlungsgruppen dargestellt.

Abkürzungen: EMEA: Europe, Middle East, Africa; MLMT: Multi-Luminanz-Mobilitätstest; SD: Standardabweichung

### Studie 301: Unterschiede zwischen den Studienzentren

Unterschiede zwischen den Studienzentren betrafen die Dauer zwischen dem Zeitpunkt der Randomisierung und der ersten Injektion. Beim Studienzentrum CHOP betrug die Wartezeit durchschnittlich 23,8 Tage (SD: 21,4), im Studienzentrum von Iowa betrug die Wartezeit 44,7 (30,7) Tage.

In die Interventionsgruppe wurden gleichmäßig Personen in beide Behandlungszentren eingeschlossen (CHOP vs. Iowa: 52 % vs. 48 %), während in die Kontrollgruppe mehr Personen aus dem Studienzentrum CHOP eingeschlossen wurden (CHOP vs. Iowa: 80 % vs. 20 %).

Unterschiede in den Randomisierungsstrata innerhalb der Studienzentren wurden nicht beobachtet. Die Zeit zwischen erster und zweiter Injektion war vergleichbar (CHOP vs. Iowa: 9,8 Tage (2,6) vs. 7,7 Tage (2,2)). Im Studienzentrum IOWA wurden in die Kontrollgruppe Personen aufgenommen, die durchschnittlich 29,5 Jahre (SD: 2,1) alt waren (Median 30 (min; max: 28; 31)), während in Iowa in der Interventionsgruppe jüngere Personen teilnahmen (MW (SD): 17,0 (14,7); Median (min; max): 11 (4; 44)). Weitere Imbalancen konnten sowohl für das Studienzentrum IOWA als auch das Studienzentrum CHOP nicht identifiziert werden.

### Studie LTFU

In Tabelle 13 sind die Charakteristika der Teilnehmenden der Studie 301, die im Anschluss in die LTFU-Studie gewechselt sind, dargestellt. Da es sich um eine Teilpopulation der Studie 301 handelt, sind die Charakteristika vergleichbar.

*Tabelle 13: LTFU-Studie: Charakterisierung der Studienpopulation (MITT)*

Studie	Ursprüngliche Intervention: Voretigen Neparvovec N = 20	Ursprüngliche Kontrolle: Beobachtendes Abwarten N = 9
<i>Alter zum Zeitpunkt der ersten Injektion (Jahre)</i>		
Mittelwert (SD)	14,6 (12,0)	15,2 (8,3)
Median (min; max)	11 (4; 44)	11 (5; 29)



Studie	Ursprüngliche Intervention: Voretigen Neparvovec N = 20	Ursprüngliche Kontrolle: Beobachtendes Abwarten N = 9
<i>Geschlecht, n (%)</i>		
Männlich	8 (40,0)	3 (33,3)
Weiblich	12 (60,0)	6 (66,7)
<i>Abstammung (genetisch), n (%)</i>		
Europäisch	14 (70,0)	6 (66,7)
Asiatisch	2 (10,0)	2 (22,2)
Indigene Völker Nordamerikas	2 (10,0)	1 (11,1)
Schwarz oder Afroamerikanisch	2 (10,0)	0 (0,0)
<i>Herkunftsländer, n (%)</i>		
USA	17 (85,0)	5 (55,6)
Niederlande	1 (5,0)	2 (22,2)
Belgien	0 (0,0)	1 (11,1)
Kanada	1 (5,0)	0 (0,0)
Italien	0 (0,0)	1 (11,1)
Mexiko	1 (5,0)	0 (0,0)
<i>Studienzentren, n (%)</i>		
Children's Hospital of Philadelphia (CHOP), USA	10 (50,0)	8 (88,9)
University of Iowa (IOWA), USA	10 (50,0)	1 (11,1)
<i>Verteilung der Personen, bei denen der MLMT bilateral bei der niedrigsten Beleuchtungsstärke noch bestanden wurde zu Baseline, n (%)</i>		
1	0 (0,0)	0 (0,0)
4	4 (20,0)	3 (33,3)
10	5 (25,0)	2 (22,2)
50	7 (35,0)	2 (22,2)
125	3 (15,0)	1 (11,1)
250	0 (0,0)	1 (11,1)
400	0 (0,0)	0 (0,0)
> 400 <sup>1)</sup>	1 (5,0)	0 (0,0)

<sup>1)</sup> Personen, die den Test bei 400 Lux nicht bestanden haben.

Abkürzungen: mITT: modified Intention-to-Treat; MLMT: Multi-Luminanz-Mobilitätstest; SD: Standardabweichung

## Exposition mit der Studienmedikation

### Studie 301

Das ins erste Auge injizierte Gesamtvolumen betrug durchschnittlich 310,0 µl (SD 44,7) und im Median 300 µl (min; max: 300 µl; 500 µl) und ins zweite Auge durchschnittlich 300,0 µl (0,0) und im Median 300 µl (min; max: 300 µl; 300 µl). Die Anzahl an Injektionen betrug ins erste Auge durchschnittlich 1,4 (SD 0,5) und im Median 1 (min; max: 1; 2) und ins zweite Auge 1,2 (0,5) und im Median 1 (1; 3).

Insgesamt haben 90 % der Personen in der Interventionsgruppe die Prüfmedikation mit dem zulassungsrelevanten Volumen von 300 µl erhalten. Gemäß dem Dossier wurde bei 2 Personen eine Reduktion des Volumens der ersten Injektion (rechtes Auge) auf  $1,0 \times 10^{11}$  Vg (200 µl) bzw.  $1,25 \times 10^{11}$  Vg (250 µl) vorgenommen. Der Grund war, dass bei der Injektion der ersten 100 µl bzw. 50 µl keine subretinale Blase beobachtet werden konnte, sodass eine zweite Injektionsstelle für die verbleibenden 200 µl bzw. 250 µl erforderlich war.

### Studie LTFU

In der ehemaligen Kontrollgruppe der Studie 301 erhielten alle 9 Teilnehmenden das Prüfpräparat in beide Augen. Das durchschnittliche Gesamtvolumen in beide Augen betrug 306 µl (SD: 16,7) bzw. im Median 300 µml (min; max: 300, 350). Die Anzahl an Injektionen betrug durchschnittlich für beide Augen 1,1 (0,2) bzw. im Median 1 (1; 2) und war zwischen den beiden Augen vergleichbar.

### **Begleitmedikation**

#### Studie 301

In Tabelle 14 ist die Begleitmedikation der Studie 301 dargestellt. In der Interventionsgruppe war der Anteil an Begleitmedikation höher als in der Kontrollgruppe bei den folgenden Wirkstoffklassen: Anilide (Intervention vs. Kontrolle: (75 % vs. 11 %), Anticholinergika (50 % vs. 0 %), Kortikosteroide + Antiinfektiva in Kombination (90 % vs. 0 %), Kortikosteroide (30 % vs. 0 %), Propionsäure-Derivate (55 % vs. 11 %). Nur bei anderen Mitteln, die den Lipidstoffwechsel beeinflussen, war der Anteil in der Kontrollgruppe (33 %) höher als in der Interventionsgruppe (0 %).

*Tabelle 14: Studie 301: Begleitmedikation mit  $\geq 30$  % in einer der beiden Behandlungsgruppen (Sicherheitspopulation)*

<b>Studie</b>	<b>Intervention: Voretigen Neparvovec N = 20</b>	<b>Kontrolle: Beobachtendes Abwarten N = 9</b>
<i>Mindestens eine Begleitmedikation, n (%)</i>	20 (100,0)	9 (100,0)
Anilide Paracetamol	17 (85,0) 15 (75,0)	3 (33,0) 1 (11,1)
Anticholinergika Cyclopentolat Hydrochlorid	10 (50,0) 10 (50,0)	0 (0,0) 0 (0,0)
Kortikosteroide + Antiinfektiva in Kombination Dexamethason + Tobramycin <sup>1)</sup>	18 (90,0) 17 (85,0)	0 (0,0) 0 (0,0)
Kortikosteroide, flach Prednisolone acetat	6 (30,0) 6 (30,0)	0 (0,0) 0 (0,0)
Andere Antiinfektiva Moxifloxacin Hydrochlorid	9 (45,0) 9 (45,0)	5 (56,0) 5 (56,0)
Propionsäure-Derivate Ibuprofen	11 (55,0) 9 (45,0)	1 (11,1) 1 (11,1)
Andere Mittel, die den Lipidstoffwechsel beeinflussen Omega-3-Fettsäuren inkl. andere Ester und Säuren	0 (0,0) 0 (0,0)	3 (33,3) 3 (33,3)
Uncodiert	4 (20)	3 (33,3)

<sup>1)</sup> Im Studienbericht: Torbadex

### Studie LTFU

Für die Studie LTFU konnten keine Angaben zur Begleitmedikation identifiziert werden.

## 3.2 Mortalität

### Studie 301 und LTFU

Todesfälle wurden im Rahmen der Sicherheitserhebungen erfasst. Es traten in beiden Studien keine Todesfälle auf.

## 3.3 Morbidität

### Multi-Luminanz-Mobilitätstest (MLMT)

#### Studie 301

Die Ergebnisse für den primären Endpunkt sind in Tabelle 15 dargestellt. Für 2 Personen (je eine Person in der Interventions- bzw. Kontrollgruppe) waren nur Baseline-Werte vorhanden, da diese Personen vorzeitig vor Applikation des Prüfpräparats aus der Studie ausschieden. Die Ergebnisse sind über den Studienverlauf konsistent mit den Ergebnissen zu Jahr 1B/C. Keine Person in der Interventionsgruppe hat beim Test zu Jahr 1 schlechter abgeschnitten als zu Baseline.

Es zeigt sich ein statistisch signifikanter Unterschied zugunsten der Intervention zu Jahr 1. Beim post hoc berechneten SMD nach Hedges' g lag der Wert über der Irrelevanzschwelle von -0,2 bzw. 0,2. Das Vorliegen eines klinisch relevanten Effekts erscheint daher plausibel.

*Tabelle 15: Studie 301: Änderung im MLMT bilateral gemessen, Jahr 1B/C im Vergleich zu Baseline (ITT-Population)*

MLMT <sup>1)</sup>	Intervention: Voretigen Neparvovec (N = 21)	Kontrolle: Beobachtendes Abwarten (N = 10)
<i>Baseline</i>		
n	21	10
Mittelwert (SD)	3,1 (1,7)	2,9 (1,6)
Median (min; max)	k. A.	k. A.
<i>Jahr 1<sup>2)</sup></i>		
n <sup>3)</sup>	21	10
Mittelwert (SD)	5,2 (1,7)	3,6 (1,4)
Änderung zu Baseline: Mittelwert (SD)	1,8 (1,1)	0,2 (1,0)
Änderung zu Baseline: Median (min; max)	2 (0; 4)	0 (-1; 2)
Exakter p-Wert <sup>4)</sup>	0,001 <sup>5)</sup>	
Änderung Intervention - Kontrolle [95%-KI] <sup>6)</sup>	1,6 [0,7; 2,4]	
SMD [95%-KI] <sup>7)</sup>	1,50 [0,66; 2,34]	

<sup>1)</sup> Ein höherer Wert im Mobilitätsscore bedeutet eine Verbesserung im MLMT.

<sup>2)</sup> Primärer Endpunkt

<sup>3)</sup> Fehlende Werte zu Jahr 1 von Personen, die vor Applikation der Prüfmedikation ausgeschieden sind (jeweils eine Person in beiden Behandlungsgruppen), wurden mit 0 imputiert.

<sup>4)</sup> Zur Untersuchung der Änderung zu Jahr 1 im Vergleich zu Baseline zwischen den Behandlungsgruppen wurde ein Wilcoxon-Rangsummen-Test mit exaktem zweiseitigem p-Wert verwendet.

<sup>5)</sup> Im Dossier wird ein p-Wert von 0,002 angegeben.

<sup>6)</sup> Es wurde die Differenz der beobachteten mittleren Veränderungen zu Baseline berechnet. Für die Berechnung des 95%-KI wird seitens des pU ein gemischtes Modell verwendet, welches Terme für die Behandlung und die Studienvisite (bzw. gemäß externem SAP Zeit und Behandlung) einschließt. Konkrete

Angaben zum statistischen Modell, beispielsweise zur Kovarianzstruktur, konnten nicht identifiziert werden. Es erfolgte keine Stratifizierung nach Randomisierungsfaktoren.

<sup>7)</sup> Die SMD wurde post hoc für das Dossier Modul 4 nach Hedges und Olkin berechnet.

Abkürzungen: ITT: Intention-to-Treat; Jahr 1B: Jahr 1 nach Behandlung des zweiten Auges in der Interventionsgruppe; Jahr 1C: Jahr 1 nach Baseline in der Kontrollgruppe; k. A.: keine Angaben; KI: Konfidenzintervall; MLMT: Multi-Luminanz-Mobilitätstest; pU: pharmazeutischer Unternehmer; SAP: Statistischer Analyseplan; SD: Standardabweichung; SMD: standardisierte Mittelwertdifferenz

Der pU gibt im Studienbericht an, dass in der Interventionsgruppe 13 der 21 Personen (61,9 %) zu Jahr 1 den MLMT bei Beleuchtungsstärke 1 Lux bestanden haben. In der Kontrollgruppe hat keine Person den Test bei 1 Lux bestanden. Bis auf eine Person, bei der der Wert des Mobilitätsscores gleich blieb, konnte bei allen Teilnehmenden in der Interventionsgruppe eine positive Änderung im bilateralen Mobilitätsscore zu Jahr 1 beobachtet werden. Bei 8 Personen (31,1 %) konnte eine Änderung im Mobilitätsscore von 1 festgestellt werden, wobei bei 4 dieser Personen (19,0 %) eine Änderung > 1 im Score nicht möglich war. Bei 11 Personen (52,4 %) konnte ein Änderungswert > 1 beobachtet werden.

Bei 4 der 10 in die Studie randomisierten Personen in der Kontrollgruppe konnte eine Verbesserung im Score beobachtet werden. Bei einer dieser Personen trat eine Verschlechterung um mehr als einen Punkt im Mobilitätsscore zu Jahr 1 auf. Bei 2 Personen trat keine Änderung auf und bei 3 Personen trat eine Verschlechterung auf.

### Studie LTFU

In Tabelle 16 sind die Ergebnisse zum MLMT, bilateral gemessen für die LTFU-Studie bis zum Datenschnitt vom 30. Juni 2020, dargestellt. Es wurden nur Ergebnisse mit mind. 70 % der Teilnehmenden (bezogen auf die mITT-Population) dargestellt: Für Jahr 6 waren Ergebnisse für 13 Teilnehmende (65,0 %) der ursprünglichen Interventionsgruppe und nur 1 Teilnehmenden (11,1 %) der ehemaligen Kontrollgruppe verfügbar. Auf eine Darstellung wird daher verzichtet.

Die Ergebnisse zu den jährlichen Visiten sind konsistent mit den Ergebnissen zu Jahr 5. Es zeigte sich ein numerischer Anstieg der Werte im Vergleich zu Baseline in gleicher Größenordnung in beiden ursprünglichen Behandlungsgruppen.

*Tabelle 16: LTFU-Studie: Ergebnisse zum Endpunkt MLMT bilateral gemessen im Vergleich zu Baseline vor Injektion bis zum Datenschnitt vom 30. Juni 2020 (mITT-Population der Studie 301)*

MLMT <sup>1)</sup>	Ursprüngliche Interventionsgruppe (N = 20)		Ursprüngliche Kontrollgruppe (N = 9)	
	Wert <sup>2)</sup>	Änderung (Visite - Baseline)	Wert <sup>2)</sup>	Änderung (Visite - Baseline)
<i>Baseline vor Injektion<sup>3)</sup></i>				
n	20	-	9	-
Mittelwert (SD)	3,3 (1,4)	-	3,6 (1,4) <sup>4)</sup>	-
Median (min; max)	3 (-1; 5)	-	4 (1; 5) <sup>4)</sup>	-
<i>Jahr 1<sup>5)</sup></i>				
n	20	20	9	9
Mittelwert (SD)	5,2 (1,7)	1,9 (1,0)	5,7 (1,0)	2,1 (1,6)
Median (min; max)	6 (-1; 6)	2 (0; 4)	6 (3; 6)	2 (0; 5)

MLMT <sup>1)</sup>	Ursprüngliche Interventionsgruppe (N = 20)		Ursprüngliche Kontrollgruppe (N = 9)	
	Wert <sup>2)</sup>	Änderung (Visite - Baseline)	Wert <sup>2)</sup>	Änderung (Visite - Baseline)
<i>Jahr 3<sup>5)</sup></i>				
n	20	20	8	8
Mittelwert (SD)	5,1 (1,7)	1,8 (1,0)	6,0 (0,0)	2,4 (1,5)
Median (min; max)	6 (-1; 6)	2 (0; 3)	6 (6; 6)	2 (1; 5)
<i>Jahr 5<sup>5)</sup></i>				
n	18	18	7	7
Mittelwert (SD)	4,9 (1,8)	1,6 (1,1)	6,0 (0,0)	2,4 (1,6)
Median (min; max)	6 (-1; 6)	1,5 (1; 3)	6 (6;6)	2 (1; 5)

<sup>1)</sup> Ein höherer Wert im Mobilitätsscore bedeutet eine Verbesserung im MLMT.

<sup>2)</sup> Die Werte zu den Visiten wurden erst im Dossier Modul 4 berichtet.

<sup>3)</sup> Als Baseline definiert ist die Untersuchung vor der ersten Injektion. Diese entsprach in der ursprünglichen Interventionsgruppe dem Zeitpunkt der Baseline-Untersuchung zu Beginn der Studie 301. In der ursprünglichen Kontrollgruppe wurde für die Wirksamkeitsendpunkte die Visite zu Jahr 1C der Studie 301 als Baseline definiert.

<sup>4)</sup> Der Wert zu Studien-Baseline (Baseline zu Beginn der Studie 301) betrug durchschnittlich 3,3 (SD 0,9) bzw. im Median 3 (min; max: 2; 5).

<sup>5)</sup> Jahr nach Behandlung des zweiten Auges.

Abkürzungsverzeichnis: Jahr 1C: Jahr 1 nach Baseline für die Kontrollgruppe der Studie 301; mITT: modified Intention-to-Treat; MLMT: Multi-Luminanz-Mobilitätstest; SD: Standardabweichung.

## Sehschärfe (ETDRS- und HOTV-Sehtafel)

### Studie 301

Personen, bei denen die ETDRS-Sehtafel in der Interventionsgruppe (n = 18) eingesetzt wurde, waren zu Baseline durchschnittlich 16 Jahre (Standardabweichung (SD): 12) alt. Es wurden Personen zwischen 5 und 44 Jahren eingeschlossen. In der Kontrollgruppe waren die Personen zu Baseline (n = 8) durchschnittlich 17 Jahre (SD: 8) alt bzw. wurden Personen zwischen 8 und 29 Jahren in die Studie eingeschlossen. Die HOTV-Sehtafel wurde bei 3 Personen in der Interventionsgruppe eingesetzt, die zwischen 4 und 6 Jahre (MD: 5 (SD: 1)) alt waren. In der Kontrollgruppe wurde die HOTV-Sehtafel bei einem 5-jährigen Kind eingesetzt [4].

Insgesamt wurden bei 4 Personen Off-Chart-Messungen durchgeführt. Bei 2 Personen der Interventionsgruppe wurden diese für den gesamten Studienverlauf vorgenommen. 2 weitere Personen sind nach der Baseline-Visite ausgeschieden.

Im Modul 4 wurden post hoc Responderanalysen für eine Verbesserung und Verschlechterung von mindestens 10 Buchstaben bzw. 15 Buchstaben für beide Sehtafeln zusammen und separat sowohl für die ETDRS- als auch HOTV-Sehtafel eingereicht. Studien zur Übertragbarkeit der MID von der ETDRS- zur HOTV-Sehtafel bzw. zur Untersuchung einer MID von  $\geq 10$  bzw.  $\geq 15$  Buchstaben für die HOTV-Sehtafel wurden nicht eingereicht. Aus diesen Gründen wird nur die Responderanalyse auf Grundlage der ETDRS-Sehtafel dargestellt. Für die Auswertung beider Sehtafeln zusammen werden die a priori definierten Auswertungen zu der mittleren Änderung zu Jahr 1 im Vergleich zu Baseline anhand eines gemischten Modells dargestellt.

Die Ergebnisse zu den a priori definierten Analysen, bei denen beide Sehtafeln (ETDRS, HOTV) zusammen ausgewertet wurden, sind in Tabelle 17 dargestellt. Es zeigten sich zu Jahr 1 keine

statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen, die Ergebnisse waren über den gesamten Studienverlauf bis zu Jahr 1 nach Baseline konsistent.

Sensitivitätsanalysen mit der mITT-Population ergaben vergleichbare Ergebnisse. Neben dem MMRM wurden weitere Sensitivitätsanalysen mit und ohne Imputationen durchgeführt, die ebenfalls vergleichbare Ergebnisse zeigten.

*Tabelle 17: Ergebnisse zum Endpunkt Sehschärfe für beide Sehtafeln (ETDRS, HOTV) zusammen: Jahr 1B/C im Vergleich zu Baseline (ITT-Population)*

Sehschärfe (logMAR) <sup>1)</sup>	Intervention: Voretigen Neparvovec (N = 21)	Kontrolle: Beobachtendes Abwarten (N = 10)
<i>Baseline</i>		
n	21	10
Mittelwert (SE)	1,2 (0,1)	1,3 (0,2)
<i>Jahr 1</i>		
n	20	9
Mittelwert (SE)	1,0 (0,2)	1,3 (0,3)
<u>Änderung zu Baseline:</u>		
n	20	9
Mittelwert (SE)	-0,2 (0,1)	0,01 (0,1)
Mittlere Änderung Intervention - Kontrolle [95%-KI] <sup>2)</sup>	-0,2 [-0,4; 0,1]	
Exakter p-Wert <sup>2)</sup>	0,1703)	

<sup>1)</sup> Es wurde für beide Behandlungsgruppen gemittelt über beide Augen analysiert.

<sup>2)</sup> Die mittlere Änderung und der p-Wert wurde mit einem MMRM mit den Termen Behandlung, Studienvisite und Behandlung\*Studienvisite gerechnet. Es war keine stratifizierte Auswertung anhand der Randomisierungsfaktoren vorgesehen.

<sup>3)</sup> Im Dossier wird ein p-Wert von 0,175 angegeben.

Abkürzungen: ETDRS: Early Treatment of Diabetic Retinopathy Study; ITT: Intention-to-treat-Population; Jahr 1B: ein Jahr nach Behandlung des zweiten Auges in der Interventionsgruppe; Jahr 1C: ein Jahr nach Baseline in der Kontrollgruppe; KI: Konfidenzintervall; logMAR: Logarithmus des minimalen Auflösungswinkels; MMRM: Mixed model for repeated measures; SE: Standardfehler

In Tabelle 18 sind die Ergebnisse der post hoc für Modul 4 durchgeführten Responderanalysen dargestellt. In der Interventionsgruppe war der Anteil an Personen mit einer Verbesserung um  $\geq 10$  Buchstaben bei 29 % der Teilnehmenden und mit einer Verbesserung um  $\geq 15$  Buchstaben bei 19 % der Teilnehmenden, während in der Kontrollgruppe bei keiner Person eine Verbesserung beobachtet wurde. Es wurden keine Effektschätzer und p-Werte vorgelegt.

Responderanalysen für eine Verschlechterung um  $\geq 10$  bzw.  $\geq 15$  Buchstaben nur für die ETDRS-Tafel wurden nicht eingereicht. Jedoch konnte für beide Sehtafeln zusammen jeweils nur bei einer Person in der Interventionsgruppe eine Verschlechterung um  $\geq 10$  bzw.  $\geq 15$  Buchstaben, jedoch bei jeweils keiner Person in der Kontrollgruppe beobachtet werden.

**Tabelle 18: Studie 301: Endpunkt Sehschärfe (ETDRS allein): Ergebnisse der Responderanalysen zur Verbesserung um  $\geq 10$  (logMAR  $\geq 0,2$ ) bzw.  $\geq 15$  Buchstaben (logMAR  $\geq 0,3$ ), Jahr 1B/C im Vergleich zu Baseline (ITT-Population)**

Sehschärfe (logMAR)	Intervention: Voretigen Neparvovec (N = 21)	Kontrolle: Beobachtendes Abwarten (N = 10)	Relatives Risiko [95%-KI] <sup>1)</sup> p-Wert <sup>2)</sup>
Abnahme logMAR $\geq 0,2$ (10 Buchstaben) n (%)	6 (28,6) <sup>3)</sup>	0 (0,0)	k. A. 0,071
Abnahme logMAR $\geq 0,3$ (15 Buchstaben) n (%)	4 (19,0) <sup>4)</sup>	0 (0,0)	k. A. 0,268

<sup>1)</sup> Es wurde die Vierfeldertafel verwendet, das KI wurde mittels exakter Methode berechnet.

<sup>2)</sup> p-Wert wurde nach der exakten Methode nach Fischer mit  $\alpha$  von 0,05 berechnet.

<sup>3)</sup> Prozentualer Anteil selbst ausgerechnet bezogen auf die ITT-Population. Im Dossier Modul 4 wird ein Anteil von 33,3 % Personen mit einer Verbesserung um mind. 10 Buchstaben angegeben.

<sup>4)</sup> Prozentualer Anteil selbst ausgerechnet bezogen auf die ITT-Population. Im Dossier Modul 4 wird ein Anteil von 22,2 % Personen mit einer Verbesserung um mind. 15 Buchstaben angegeben.

Abkürzungen: ETDRS: Early Treatment of Diabetic Retinopathy Study; ITT: Intention-to-Treat; Jahr 1B: Jahr 1 nach Injektion ins zweite Auge (Studie 301: Interventionsgruppe); Jahr 1C: ein Jahr nach Baseline (Studie 301: Kontrollgruppe); k. A.: keine Angabe; KI: Konfidenzintervall; logMAR: Logarithmus des minimalen Auflösungs winkels

### Studie LTFU

In Tabelle 19 sind die Ergebnisse für den Endpunkt Sehschärfe für beide Sehtafeln (ETDRS- und HOTV-Sehtafel) zusammen und ggf. Off-Chart-Analysen zur mittleren Veränderung, gemittelt über beide Augen, zu den Visiten im Vergleich zu Baseline dargestellt. Für die LTFU-Studie konnten keine separaten Responderanalysen für beide Sehtafeln zur Verbesserung bzw. Verschlechterung um mind. 10 bzw. 15 Buchstaben identifiziert werden. Gemäß dem Interimsstudienbericht wurden für 2 Teilnehmende Off-Chart-Ergebnisse zu den Visiten Jahr 4 und 5 verwendet.

Es wurden nur Ergebnisse mit mind. 70 % der Teilnehmenden (bezogen auf die mITT-Population) dargestellt: Für Jahr 6 waren Ergebnisse für 13 Teilnehmende (65,0 %) der ursprünglichen Interventionsgruppe verfügbar. Für Jahr 6 nach Injektion waren Ergebnisse für nur 1 Teilnehmenden (11,1 %) der ehemaligen Kontrollgruppe verfügbar. Auf eine Darstellung wird daher verzichtet.

Die Ergebnisse zu den jährlichen Visiten im Vergleich zu Baseline sind für beide Behandlungsgruppen konsistent mit den Ergebnissen zu Jahr 5. Es zeigte sich zumeist eine numerische Reduktion der Werte in beiden ursprünglichen Behandlungsgruppen.

**Tabelle 19: LTFU-Studie: Endpunkt Sehschärfe (ETDRS- und HOTV-Sehtafel, Holladay Off-Chart), gemittelt über beide Augen, Datenschnitt am 30. Juni 2020 (mITT-Population)**

Sehschärfe (logMAR) <sup>1)2)</sup>	Ursprüngliche Interventionsgruppe (N = 20)		Ursprüngliche Kontrollgruppe (N = 9)	
	Wert	Änderung (Visite - Baseline)	Wert	Änderung (Visite - Baseline)
Baseline vor Injektion <sup>3)</sup> n Mittelwert (SD)	20 1,1 (0,4)	-	9 1,0 (0,3) <sup>4)</sup>	-
Jahr 1 <sup>5)</sup> n Mittelwert (SD)	20 1,0 (0,5)	20 -0,2 (0,3)	9 0,9 (0,3)	9 -0,1 (0,2)

Sehschärfe (logMAR) <sup>1)2)</sup>	Ursprüngliche Interventionsgruppe (N = 20)		Ursprüngliche Kontrollgruppe (N = 9)	
	Wert	Änderung (Visite - Baseline)	Wert	Änderung (Visite - Baseline)
<i>Jahr 3<sup>5)</sup></i>				
n	20	20	8	8
Mittelwert (SD)	1,0 (0,6)	-0,2 (0,4)	0,9 (0,3)	-0,1 (0,2)
<i>Jahr 5<sup>5)</sup></i>				
n	18	18	7	7
Mittelwert (SD)	1,1 (1,8)	0,0 (0,6)	0,8 (0,3)	-0,1 (0,3)

<sup>1)</sup> Es wurde für beide Behandlungsgruppen gemittelt über beide Augen analysiert.

<sup>2)</sup> Es wurde die Sehschärfe für beide Sehtafeln zusammen (ETDRS-, HOTV-Sehtafel) und ggf. Off-Chart-Analysen (nach Holladay) dargestellt.

<sup>3)</sup> Als Baseline definiert ist die Untersuchung vor der ersten Injektion. Diese entsprach in der ursprünglichen Interventionsgruppe dem Zeitpunkt der Baseline-Untersuchung zu Beginn der Studie 301. In der ursprünglichen Kontrollgruppe wurde für die Wirksamkeitsendpunkte die Visite zu Jahr 1 der Studie 301 als Baseline definiert.

<sup>4)</sup> Der Wert zu Studien-Baseline (Baseline zu Beginn der Studie 301) betrug durchschnittlich 1,0 (SD 0,3) bzw. im Median 3 (min; max: 2; 5).

<sup>5)</sup> Jahr x nach Behandlung des zweiten Auges.

Abkürzungen: ETDRS: Early Treatment of Diabetic Retinopathy Study; mITT: modified Intention-to-Treat; logMAR: Logarithmus des minimalen Auflösungs winkels; SD: Standardabweichung

## Lichtempfindlichkeit (Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test (FST))

### Studie 301

In Tabelle 20 sind die Ergebnisse zum Endpunkt FST dargestellt. Die Ergebnisse zum FST, gemessen mit weißem, blauem und rotem Licht, zeigen vergleichbare statistisch signifikante Unterschiede zugunsten der Intervention. Die Ergebnisse sind über den gesamten Studienverlauf bis zu Jahr 1 nach Baseline konsistent. Es zeigt sich ein statistisch signifikanter Unterschied zugunsten der Intervention zu Jahr 1. Beim post hoc berechneten SMD nach Hedges' g lag der Wert über der Irrelevanzschwelle von -0,2 bzw. 0,2. Das Vorliegen eines klinisch relevanten Effekts erscheint daher plausibel.

Sensitivitätsanalysen anhand der mITT-Population zeigten für den FST gemessen mit weißem, blauem und rotem Licht berechnet anhand des gemischten Modells mit und ohne Imputationen vergleichbare Ergebnisse.

Tabelle 20: Studie 301: Ergebnisse zum Endpunkt FST zur Visite Jahr 1B/C im Vergleich zu Baseline (ITT-Population)

FST <sup>1)2)</sup>	Intervention: Voretigen Neparvovec (N = 21)	Kontrolle: Beobachtendes Abwarten (N = 10)
<b>Weißes Licht (log<sub>10</sub> (cd s/m<sup>2</sup>))<sup>3)</sup></b>		
<i>Baseline</i>		
n	20 <sup>5)</sup>	9 <sup>5)</sup>
Mittelwert (SE)	-1,3 (0,1)	-1,7 (0,1)
<i>Jahr 1</i>		
n	20	9
Mittelwert (SE)	-3,4 (0,4)	-1,6 (0,4)



<b>FST<sup>1)2)</sup></b>	<b>Intervention: Voretigen Neparvovec (N = 21)</b>	<b>Kontrolle: Beobachtendes Abwarten (N = 10)</b>
<u>Änderung zu Baseline:</u>		
n	19	9
Mittelwert (SE)	-2,1 (0,3)	0,04 (0,4)
Mittlere Änderung Intervention - Kontrolle [95%-KI] <sup>4)</sup>	-2,1 [-3,2; -1,0]	
p-Wert <sup>4)</sup>	< 0,001	
SMD [95%-KI] <sup>6)</sup>	-1,52 [-2,41; -0,63]	
<b>Blaues Licht (log10 (cd s/m<sup>2</sup>))</b>		
<i>Baseline</i>		
n	20 <sup>5)</sup>	9 <sup>5)</sup>
Mittelwert (SE)	-1,6 (0,1)	-2,0 (0,2)
<i>Jahr 1</i>		
n	17 <sup>7)</sup>	9
Mittelwert (SE)	-3,6 (0,3)	-1,9 (0,4)
<u>Änderung zu Baseline:</u>		
n	17	9
Mittelwert (SE)	-2,0 (0,3)	0,1 (0,5)
Mittlere Änderung Intervention - Kontrolle [95%-KI] <sup>4)</sup>	-2,1 [-3,3; -0,9]	
p-Wert <sup>4)</sup>	0,001	
<b>Rotes Licht (log10 (cd s/m<sup>2</sup>))</b>		
<i>Baseline</i>		
n	20 <sup>5)</sup>	9 <sup>5)</sup>
Mittelwert (SE)	-1,2 (0,1)	-1,7 (0,2)
<i>Jahr 1</i>		
n	17 <sup>7)</sup>	9
Mittelwert (SE)	-2,5 (0,2)	-1,5 (0,3)
<u>Änderung zu Baseline:</u>		
n	17	9
Mittelwert (SE)	-1,3 (0,2)	0,2 (0,2)
Mittlere Änderung Intervention - Kontrolle [95%-KI] <sup>4)</sup>	-1,5 [-2,1; -0,9]	
Exakter p-Wert <sup>4)</sup>	< 0,001	

<sup>1)</sup> Es wurde für beide Behandlungsgruppen gemittelt über beide Augen analysiert.

<sup>2)</sup> Je negativer der Wert, desto besser die Lichtempfindlichkeit.

<sup>3)</sup> Sekundärer Endpunkt.

<sup>4)</sup> Die mittlere Änderung und der p-Wert wurden mit einem MMRM mit den Termen Behandlung, Studienvisite und Behandlung\*Studienvisite gerechnet. Es waren keine Imputationen vorgesehen. Eine Analyse adjustiert nach den Stratifizierungsmerkmalen war nicht vorgesehen.

<sup>5)</sup> Gemäß pU sollte die Analyse anhand der ITT-Population durchgeführt werden. Jeweils eine Person in beiden Studienarmen ist nach der Baseline-Visite ausgeschieden. Bei einer Person konnten zwar Ergebnisse zu Jahr 1 ermittelt werden, jedoch konnte zu den anderen Visiten kein reliables Testergebnis mit weißem Licht erzielt werden.

<sup>6)</sup> SMD nach Hedges' g wurde für den Endpunkt FST mit weißem Licht post hoc für das Modul 4 berechnet.

<sup>7)</sup> Angaben, weshalb Ergebnisse zu Jahr 1 von nur 17 Personen in der Interventionsgruppe berichtet wurden, konnten nicht identifiziert werden.

Abkürzungen: cd s: Candela-Sekunde; FST: Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test; Jahr 1B: Jahr 1 nach Behandlung des zweiten Auges in der Interventionsgruppe; Jahr 1C: Jahr 1 nach Baseline in der Kontrollgruppe; (m)ITT: (modified) Intention-to-Treat; KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; pU: pharmazeutischer Unternehmer; SE: Standardfehler; SMD: Standardisierte Mittelwertdifferenz.

### Studie LTFU

In Tabelle 21 sind die Ergebnisse zum Endpunkt FST, gemessen mit weißem, blauem und rotem Licht, dargestellt. Es wurden nur Ergebnisse mit mind. 70 % der Teilnehmenden (bezogen auf die mITT-Population) dargestellt: Für Jahr 6 waren Ergebnisse für 13 Teilnehmende (65,0 %) der ursprünglichen Interventionsgruppe verfügbar. Für Jahr 6 nach Injektion waren Ergebnisse für nur einen Teilnehmenden (11,1 %) der ehemaligen Kontrollgruppe verfügbar. Auf eine Darstellung wird daher verzichtet.

Die Ergebnisse zu den jährlichen Visiten sind konsistent mit den Ergebnissen zu Jahr 5. Es zeigte sich eine numerische Reduktion der Werte in beiden ursprünglichen Behandlungsgruppen.

*Tabelle 21: LTFU-Studie: Ergebnisse zum Endpunkt FST gemittelt über beide Augen bis zum Datenschnitt am 30. Juni 2020 (mITT-Population der Studie 301)*

FST <sup>1)2)3)4)</sup>	Ursprüngliche Interventionsgruppe (N = 20)		Ursprüngliche Kontrollgruppe (N = 9)	
	Wert	Änderung (Visite - Baseline)	Wert	Änderung (Visite - Baseline)
<b>Weißes Licht (log<sub>10</sub> (cd s/m<sup>2</sup>))</b>				
<i>Baseline vor Injektion<sup>5)</sup></i>		-		-
n	19		9	
Mittelwert (SD)	-1,3 (0,4)		-1,6 (0,5) <sup>5)</sup>	
<i>Jahr 1<sup>6)</sup></i>				
n	20	19	9	9
Mittelwert (SD)	-3,4 (1,5)	-2,1 (1,6)	-4,5 (1,5)	-2,9 (1,5)
<i>Jahr 3<sup>6)</sup></i>				
n	20	19	8	8
Mittelwert (SD)	-3,3 (1,3)	-2,0 (1,4)	-4,5 (1,2)	-2,9 (1,1)
<i>Jahr 5<sup>6)</sup></i>				
n	18	17	7	7
Mittelwert (SD)	-3,2 (1,3)	-2,0 (1,5)	-4,1 (1,3)	-2,6 (1,2)
<b>Blaues Licht (log<sub>10</sub> (cd s/m<sup>2</sup>))</b>				
<i>Baseline vor Injektion<sup>5)</sup></i>				
n	19		9	
Mittelwert (SD)	-1,7 (1,6)		-1,9 (0,4)	
<i>Jahr 1<sup>6)</sup></i>				
n	17	17	9	9
Mittelwert (SD)	-3,7 (1,6)	-2,0 (1,8)	-4,9 (1,5)	-3,0 (1,5)
<i>Jahr 3<sup>6)</sup></i>				
n	19	19	8	8
Mittelwert (SD)	-3,7 (1,5)	-2,1 (1,7)	-5,0 (1,2)	-3,1 (1,1)
<i>Jahr 5<sup>6)</sup></i>				
n	18	17	7	7
Mittelwert (SD)	-3,7 (1,4)	-2,1 (1,6)	-4,6 (1,5)	-2,8 (1,5)
<b>Rotes Licht (log<sub>10</sub> (cd s/m<sup>2</sup>))</b>				
<i>Baseline vor Injektion<sup>5)</sup></i>				
n	19		9	
Mittelwert (SD)	-1,3 (0,4)		-1,5 (0,5)	

FST <sup>1)2)3)4)</sup>	Ursprüngliche Interventionsgruppe (N = 20)		Ursprüngliche Kontrollgruppe (N = 9)	
	Wert	Änderung (Visite - Baseline)	Wert	Änderung (Visite - Baseline)
<i>Jahr 1<sup>6)</sup></i>				
n	17	17	9	9
Mittelwert (SD)	-2,6 (0,9)	-1,3 (0,9)	-3,2 (0,8)	-1,6 (1,0)
<i>Jahr 3<sup>6)</sup></i>				
n	19	19	8	8
Mittelwert (SD)	-2,6 (0,8)	-1,3 (0,8)	-3,3 (0,7)	-1,8 (0,7)
<i>Jahr 5<sup>6)</sup></i>				
n	18	17	7	7
Mittelwert (SD)	-2,5 (0,8)	-1,3 (0,8)	-3,0 (0,8)	-1,6 (0,8)

<sup>1)</sup> Es wurde für beide Behandlungsgruppen, gemittelt über beide Augen, analysiert.

<sup>2)</sup> Je negativer der Wert, desto besser die Lichtempfindlichkeit.

<sup>3)</sup> Es wurden nur die beobachteten Werte berichtet.

<sup>4)</sup> Als Baseline definiert ist die Untersuchung vor der ersten Injektion. Diese entsprach in der ursprünglichen Interventionsgruppe dem Zeitpunkt der Baseline-Untersuchung zu Beginn der Studie 301. In der ursprünglichen Kontrollgruppe wurde für die Wirksamkeitsendpunkte die Visite zu Jahr 1 der Studie 301 als Baseline definiert.

<sup>5)</sup> Der Wert zu Studien-Baseline (Baseline zu Beginn der Studie 301) betrug durchschnittlich -1,7 (SD 0,4).

<sup>6)</sup> Jahr x nach Behandlung des zweiten Auges.

Abkürzungen: cd s: Candela-Sekunde; FST: Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test; mITT: modified Intention-to-Treat; SD: Standardabweichung

## Perimetrie nach Goldmann bzw. Humphrey

### Studie 301

In Tabelle 22 sind die Ergebnisse zur Gesichtsfeldmessung für beide Perimetrie-Tests dargestellt. Es konnten statistisch signifikante Unterschiede zugunsten der Intervention bei der Perimetrie nach Goldmann mit dem Stimulus III4e sowie bei der Perimetrie nach Humphrey für den Bereich Makula gezeigt werden, jedoch nicht für die Perimetrie nach Humphrey für den Bereich Fovea. Aufgrund der geringen Anzahl an Personen (jeweils < 70 %) in der Interventionsgruppe (n = 11) und der Kontrollgruppe (n = 5) wird auf die Darstellung der Ergebnisse für die Perimetrie nach Goldmann für den Stimulus V4e verzichtet.

Die Ergebnisse sind über den gesamten Studienverlauf bis zu Jahr 1 nach Baseline konsistent. Das angewandte statistische Analyseverfahren wurde erst post hoc im Studienbericht bzw. Dossier beschrieben. Es zeigt sich ein statistisch signifikanter Unterschied zugunsten der Intervention zu Jahr 1 für die Perimetrie nach Goldmann und Perimetrie nach Humphrey für die Makula-Region. Beim post hoc berechneten SMD nach Hedges' g lag der Wert über der Irrelevanzschwelle von -0,2 bzw. 0,2. Das Vorliegen eines klinisch relevanten Effekts erscheint daher plausibel.

Tabelle 22: Studie 301: Gesichtsfeldmessung, gemittelt über beide Augen; Jahr 1B/C im Vergleich zu Baseline (ITT-Population)

<b>Perimetrie<sup>1)</sup></b>	<b>Intervention: Voretigen Neparvovec (N = 21)</b>	<b>Kontrolle: Beobachtendes Abwarten (N = 10)</b>
<b>Goldmann III4e (Summenscore)<sup>2)</sup></b>		
<i>Baseline</i>		
n	20 <sup>3)</sup>	10
Mittelwert (SD)	332,9 (413,3)	427,1 (372,0)
<i>Jahr 1</i>		
n	20	9
Mittelwert (SD)	673,9 (423,7)	397,8 (367,3)
<u>Änderung zu Baseline:</u>		
n	19	9
Mittelwert (SD)	302,1 (289,6)	-76,7 (258,7)
Differenz Änderung Intervention - Kontrolle [95%-KI] <sup>4)</sup>	378,8 [145,5; 612,0]	
p-Wert <sup>5)</sup>	0,006	
SMD [95%-KI] <sup>6)</sup>	1,27 [0,41; 2,12]	
<b>Humphrey: Fovea-Sensitivität (dB)</b>		
<i>Baseline</i>		
n	20 <sup>3)</sup>	10
Mittelwert (SD)	22,4 (6,8)	17,6 (8,9)
<i>Jahr 1</i>		
n	20	9
Mittelwert (SD)	25,8 (9,1)	21,1 (8,9)
<u>Änderung zu Baseline:</u>		
n	19	9
Mittelwert (SD)	2,4 (9,7)	2,3 (5,3)
Differenz Änderung Intervention - Kontrolle [95%-KI] <sup>4)</sup>	0,04 [-7,1; 7,2]	
p-Wert <sup>5)</sup>	0,176	
<b>Humphrey: Durchschnittlicher Makula-Grenzwert (dB)</b>		
<i>Baseline</i>		
n	20 <sup>3)</sup>	10
Mittelwert (SD)	16,1 (5,5)	14,4 (8,0)
<i>Jahr 1</i>		
n	20	9
Mittelwert (SD)	24,0 (8,0)	15,8 (7,4)
Median (min; max)	28 (2; 32)	16 (2; 25)
<u>Änderung zu Baseline:</u>		
n	19	9
Mittelwert (SD)	7,7 (6,2)	-0,2 (1,7)
Differenz Änderung Intervention - Kontrolle [95%-KI] <sup>4)</sup>	7,9 [3,5; 12,2]	
p-Wert <sup>5)</sup>	< 0,001	
SMD [95%-KI] <sup>6)</sup>	1,45 [0,61; 2,29]	

<sup>1)</sup> Es wurde für beide Behandlungsgruppen gemittelt über beide Augen analysiert.

<sup>2)</sup> Größe des Stimulus III4e: Größe: 4 mm<sup>2</sup>, Leuchtdichte: 315 cd/m<sup>2</sup>.

<sup>3)</sup> Für eine Person in der Interventionsgruppe konnten zu Baseline keine zuverlässigen Testergebnisse ermittelt werden.

- 4) Es wurde die Differenz der beobachteten mittleren Veränderungen zu Baseline berechnet. Für die Berechnung des 95%-KI wurde ein gemischtes Modell verwendet, welches Terme für die Behandlung und die Studienvisite einschließt. Es wurde keine Stratifizierung nach den Randomisierungsfaktoren vorgenommen.
- 5) Der zweiseitige p-Wert wurde post hoc mittels Wilcoxon-Rangsummen-Test berechnet. Es wurden keine Imputationen für fehlende Werte durchgeführt.
- 6) SMD und Hedges' g wurden post hoc für das Dossier berechnet.

Abkürzungen: cd: Candela; dB: Dezibel; ITT: Intention-to-Treat; Jahr 1B: Jahr 1 nach Injektion ins zweite Auge (Studie 301: Interventionsgruppe); Jahr 1C: ein Jahr nach Baseline (Studie 301: Kontrollgruppe); KI: Konfidenzintervall; SD: Standardabweichung; SMD: Standardisierte Mittelwertdifferenz.

### Studie LTFU

Die Ergebnisse zur Perimetrie nach Goldmann sind in Tabelle 23 dargestellt. Für die Goldmann-Perimetrie unter Verwendung des Stimulus V4e waren ab Jahr 1 nach Injektion in der ursprünglichen Behandlungsgruppe Daten von < 70 % der Personen und in der ursprünglichen Kontrollgruppe von ≤ 11 % verfügbar. Aus diesem Grund werden die Ergebnisse nicht dargestellt.

Für die Goldmann-Perimetrie mit Stimulus III4e zeigte sich zu allen Zeitpunkten ein numerischer Anstieg der Werte im Vergleich zu Baseline in beiden ursprünglichen Behandlungsgruppen. Zwischen Jahr 1 und Jahr 5 fluktuieren (ehemalige Kontrollgruppe) oder verringern sich die Werte (ursprüngliche Interventionsgruppe).

*Tabelle 23: LTFU-Studie: Perimetrie nach Goldmann, gemittelt über beide Augen, Datenschnitt am 30. Juni 2020 (mITT-Population der Studie 301)*

Perimetrie nach Goldmann <sup>1)2)</sup>	Ursprüngliche Interventionsgruppe (N = 20)		Ursprüngliche Kontrollgruppe (N = 9)	
	Wert	Änderung (Visite - Baseline)	Wert	Änderung (Visite - Baseline)
<b>Perimetrie nach Goldmann Stimulus III4e<sup>3)</sup> (Summenscore)</b>				
<i>Baseline vor Injektion<sup>4)</sup></i>				
n	19 <sup>5)</sup>		9	
Mittelwert (SD)	350,4 (416,9)		397,8 (367,3) <sup>6)</sup>	
<i>Jahr 1<sup>7)</sup></i>				
n	20	19	9	9
Mittelwert (SD)	673,9 (423,7)	302,1 (289,6)	592,1 (296,6)	194,3 (244,7)
<i>Jahr 3<sup>7)</sup></i>				
n	19	18	8	8
Mittelwert (SD)	625,9 (413,3)	282,2 (256,5)	579,1 (247,8)	157,9 (325,3)
<i>Jahr 5<sup>7)</sup></i>				
n	16	15	7	7
Mittelwert (SD)	533,4 (428,5)	166,6 (208,7)	506,9 (219,8)	188,9 (222,3)

1) Es wurde für beide Behandlungsgruppen gemittelt über beide Augen analysiert.

2) Es wurde die Sehschärfe für beide Sehtafeln zusammen (ETDRS-, HOTV-Sehtafel) und ggf. Off-Chart-Analysen dargestellt.

3) Größe des Stimulus III4e: Größe: 4 mm<sup>2</sup>, Leuchtdichte: 315 cd/m<sup>2</sup>.

4) Als Baseline definiert ist die Untersuchung vor der ersten Injektion. Diese entsprach in der ursprünglichen Interventionsgruppe dem Zeitpunkt der Baseline-Untersuchung zu Beginn der Studie 301. In der ursprünglichen Kontrollgruppe wurde für die Wirksamkeitseindpunkte die Visite zu Jahr 1 der Studie 301 als Baseline definiert.

5) Für 1 Person in der Interventionsgruppe konnten zu Baseline keine zuverlässigen Testergebnisse ermittelt werden.

6) Der Wert zu Studien-Baseline (Baseline zu Beginn der Studie 301) betrug durchschnittlich 474,5 Grad (SD 361,0).

7) Jahr x nach Behandlung des zweiten Auges.

Abkürzungen: cd: Candela; ETDRS: Early Treatment of Diabetic Retinopathy Study; mITT: modified Intention-to-Treat; SD: Standardabweichung.

Die Ergebnisse zur Perimetrie nach Humphrey sind in Tabelle 24 dargestellt. Für die Perimetrie nach Humphrey im Bereich der Fovea zeigte sich zu allen Zeitpunkten ein numerischer Anstieg der Werte im Vergleich zu Baseline in beiden ursprünglichen Behandlungsgruppen. Zwischen Jahr 1 und Jahr 5 fluktuieren (ehemalige Interventionsgruppe) oder erhöhen sich die Werte bis Jahr 3 und bleiben bis Jahr 5 in einer vergleichbaren Größenordnung (ursprüngliche Kontrollgruppe).

Für die Perimetrie nach Humphrey im Bereich der Macula zeigte sich ebenfalls zu allen Zeitpunkten ein numerischer Anstieg der Werte im Vergleich zu Baseline in beiden ursprünglichen Behandlungsgruppen. Zwischen Jahr 1 und Jahr 5 kam es zu einem Anstieg (ursprüngliche Kontrollgruppe) oder einer Reduktion (ursprüngliche Interventionsgruppe) der Werte.

*Tabelle 24: LTFU-Studie: Perimetrie nach Humphrey: gemittelt über beide Augen, Datenschnitt am 30. Juni 2020 (mITT-Population der Studie 301)*

Perimetrie nach Humphrey <sup>1)</sup>	Ursprüngliche Interventionsgruppe (N = 20)		Ursprüngliche Kontrollgruppe (N = 9)	
	Wert	Änderung (Visite - Baseline)	Wert	Änderung (Visite - Baseline)
<b>Perimetrie nach Humphrey: Bereich Fovea (db)</b>				
<i>Baseline vor Injektion<sup>2)</sup></i>				
n	19		9	
Mittelwert (SD)	23,3 (5,5)		21,5 (8,9) <sup>4)</sup>	
<i>Jahr 1<sup>3)</sup></i>				
n	20	19	9	9
Mittelwert (SD)	25,8 (9,1)	2,4 (9,7)	24,7 (9,4)	3,2 (11,5)
<i>Jahr 3<sup>3)</sup></i>				
n	20	19	8	8
Mittelwert (SD)	26,6 (8,1)	3,0 (8,7)	26,8 (4,1)	4,7 (77,0)
<i>Jahr 5<sup>3)</sup></i>				
n	17	16	7	7
Mittelwert (SD)	25,7 (8,1)	1,2 (8,1)	25,6 (9,4)	4,7 (6,9)
<b>Perimetrie nach Humphrey: Bereich Macula (db)</b>				
<i>Baseline vor Injektion<sup>2)</sup></i>				
n	19		9	
Mittelwert (SD)	16,6 (5,3)		15,8 (7,4) <sup>5)</sup>	
<i>Jahr 1<sup>3)</sup></i>				
n	20	19	9	9
Mittelwert (SD)	24,0 (8,0)	7,7 (6,2)	21,0 (11,9)	5,2 (9,9)
<i>Jahr 3<sup>3)</sup></i>				
n	20	19	8	8
Mittelwert (SD)	22,9 (6,9)	6,5 (5,8)	23,0 (8,8)	6,8 (6,4)
<i>Jahr 5<sup>3)</sup></i>				
n	18	17	7	7
Mittelwert (SD)	21,7 (7,1)	4,8 (6,7)	23,3 (5,5)	8,2 (5,2)

<sup>1)</sup> Es wurde für beide Behandlungsgruppen gemittelt über beide Augen analysiert.

<sup>2)</sup> Als Baseline definiert ist die Untersuchung vor der ersten Injektion. Diese entsprach in der ursprünglichen Interventionsgruppe dem Zeitpunkt der Baseline-Untersuchung zu Beginn der Studie 301. In der ursprünglichen Kontrollgruppe wurde für die Wirksamkeitsendpunkte die Visite zu Jahr 1 der Studie 301 als Baseline definiert.

<sup>3)</sup> Jahr x nach Behandlung des zweiten Auges.

<sup>4)</sup> Der Wert zu Studienbaseline (Baseline zu Beginn der Studie 301) betrug durchschnittlich 19,2 (7,8).

<sup>5)</sup> Der Wert zu Studienbaseline (Baseline zu Beginn der Studie 301) betrug durchschnittlich 15,8 (7,4).

Abkürzungen: db: dezibel; mITT: modified Intention-to-Treat; SD: Standardabweichung.

### 3.4 Lebensqualität

Es wurden keine bewertungsrelevanten Daten eingereicht.

### 3.5 Sicherheit

#### Studie 301

Es werden Sicherheitsdaten für die Studie 301 bis Jahr 1B für die Interventionsgruppe (1 Jahr nach Erhalt der ersten Injektion), 1C für die Kontrollgruppe (1 Jahr nach Baseline) dargestellt. In den vom pU eingereichten Unterlagen werden in der Interventionsgruppe – abweichend von der Kontrollgruppe und dem üblichen Vorgehen einer zeitgleichen Erfassung – UE nicht ab Baseline, sondern erst ab der ersten Injektion berichtet. Hierdurch werden in der Interventionsgruppe UE, die zwischen Baseline und Injektion ins erste Auge aufgetreten sind, nicht erfasst. Die Beobachtungsdauer zwischen Baseline und Injektion ins erste Auge beträgt in der Interventionsgruppe im Durchschnitt 34,3 Tage (SD 27,9) bzw. im Median 22 Tage (min; max: 6; 78).

Im Modul 4 werden post hoc die Ergebnisse für UE Grad  $\geq 3$  neben der studienindividuellen Schweregradeinteilung aus dem Studienprotokoll dargestellt. Es bleibt unklar, nach welchem Klassifizierungssystem die Einteilung erfolgte und inwieweit diese mit der beschriebenen Schweregradeinteilung im Studienprotokoll zusammenhängt. Aus diesem Grund werden die UE Grad  $\geq 3$  nach SOC und PT im Anhang der Nutzenbewertung dargestellt.

Die zusammenfassende Darstellung der UE anhand der Sicherheitspopulation ist in Tabelle 25 dargestellt. Es wurden keine a priori definierten UE von besonderem Interesse angegeben.

*Tabelle 25: Studie 301: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE zu Jahr 1 nach Baseline (Sicherheitspopulation)*

Personen mit mindestens einem ...	Intervention: Voretigen Neparvovec <sup>1)2)</sup> (N = 20) n (%)	Kontrolle: Beobachtendes Abwarten <sup>1)2)</sup> (N = 9) n (%)	RR [95%-KI] <sup>3)</sup> ; p-Wert <sup>4)</sup>
UE	20 (100)	9 (100)	-
Schwere UE <sup>5)</sup>	3 (15,0)	0 (0,0)	k. A. <sup>6)</sup> /0,532
SUE	2 (10,0)	0 (0,0)	k. A. <sup>6)</sup> /1,000
UE, die zum Studienabbruch führten	0 (0,0)	0 (0,0)	n. b.

<sup>1)</sup> Unterschiedliche Erhebungszeiträume zwischen beiden Behandlungsgruppen: In der Interventionsgruppe wurden UE zwischen Tag 0A und Jahr 1B und in der Kontrollgruppe zwischen Baseline und Jahr 1C erhoben. Die durchschnittliche Beobachtungsdauer zwischen Baseline und Injektion ins erste Auge beträgt 34,3 Tage (SD 27,9) bzw. im Median 22 Tage (min; max: 6; 78) in der Interventionsgruppe. Aufgrund des späteren Erfassungsbeginns in der Interventionsgruppe im Vergleich zur Kontrollgruppe ist eine einheitliche und vollständige Erhebung in diesem Zeitraum für die Interventionsgruppe nicht gegeben.

<sup>2)</sup> Im aktuellen Dossier konnten keine konkreten Angaben zu den Beobachtungszeiten identifiziert werden. Unter Berücksichtigung des vorherigen Verfahrens wird für die Interventionsgruppe die Beobachtungszeit zwischen Tag 0A bis Jahr 1B mit durchschnittlich 406,6 Tagen (SD: 20,4) angegeben. In der Kontrollgruppe lag der durchschnittliche Zeitraum zwischen Baseline und Jahr 1C bei 354,8 Tagen (SD: 20,6) [4]. Der Unterschied zwischen beiden Gruppen betrug somit ca. 50 Tage und wird als hinreichend ähnlich zur Verwendung der nicht-Zeit-adjustierten Effektschätzer angesehen. Angaben zu den medianen Beobachtungszeiten konnten nicht identifiziert werden.

<sup>3)</sup> Vom pU post hoc berechneter Effektschätzer mittels Vier-Felder-Tafel. Das KI wurde mit einer exakten Methode (Tail-Methode) berechnet.

<sup>4)</sup> P-Wert berechnet post hoc anhand des exakten Tests nach Fischer.

<sup>5)</sup> Es wurde gemäß Studienprotokoll eine studienindividuelle Schweregradeinteilung durch den Prüfarzt bzw. die Prüfarztin vorgenommen.

<sup>6)</sup> Es wurden seitens des pU keine Effektschätzer berechnet.

Abkürzungen: Jahr 1B: Jahr 1 nach Injektion ins zweite Auge (Studie 301: Interventionsgruppe); Jahr 1C: 1 Jahr nach Baseline (Studie 301: Kontrollgruppe); k. A.: keine Angaben, KI: Konfidenzintervall; n. b.: nicht bestimmbar; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SD: Standardabweichung; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis; Tag 0A: Tag der Injektion ins erste Auge

## Studie LTFU

Im Modul 4 werden UE für den Zeitraum Injektion ins erste Auge bis Jahr 1B und Injektion ins erste Auge bis Jahr 5 (= Zeitpunkt der letzten Injektion + 1.825 Tage) angegeben. Jedoch werden keine (medianen) Beobachtungszeiten gegeben. Für die ursprüngliche Interventionsgruppe betrug die mittlere Beobachtungszeit 406,6 Tage (SD: 20,4) zwischen dem Tag der ersten Injektion und Jahr 1 nach der zweiten Injektion.

UE Grad  $\geq 3$  wurden nur im Dossier Modul 4 post hoc berichtet. Aufgrund der beschriebenen Unklarheiten zum Klassifizierungssystem, wie in Studie 301, werden die Ergebnisse auf SOC- und PT-Ebene im Anhang dargestellt (siehe Ausführungen zur Schweregradeinteilung in Studie 301 weiter oben).

*Tabelle 26: Studie LTFU: Ergebnisse der Nebenwirkungen<sup>1)</sup> – Zusammenfassung der UE bis zum Datenschnitt vom 30. Juni 2020 (Sicherheitspopulation)*

Personen mit mindestens einem ...	Ursprüngliche Intervention <sup>2)</sup> N = 20 n (%)	Ursprüngliche Kontrolle <sup>2)</sup> N = 9 n (%)
<b>Erhalt Injektion ins erste Auge bis Jahr 1B<sup>3)</sup></b>		
UE	20 (100,0)	9 (100,0)
Schwere UE <sup>4)</sup>	3 (15,0)	0 (0,0)
SUE	2 (10,0)	1 (11,1)
UE, das zum Abbruch der Studienmedikation führte	0 (0,0)	0 (0,0)
UE von besonderem Interesse <sup>5)</sup>	k. A.	k. A.
<b>Erhalt Injektion ins erste Auge bis Jahr 5<sup>6)7)</sup></b>		
UE	20 (100,0)	9 (100,0)
Schwere UE <sup>4)</sup>	4 (20,0)	0 (0,0)
SUE	4 (20,0)	1 (11,1)
UE, das zum Abbruch der Studienmedikation führte	0 (0,0)	0 (0,0)
UE von besonderem Interesse <sup>5)</sup>	k. A. <sup>8)</sup>	k. A. <sup>8)</sup>

<sup>1)</sup> Gemäß SAP (22. Juni 2016) war in der LTFU-Studie eine vollständige Erhebung der UE nur bis Jahr 1B, jedoch nicht im Rahmen einer Langzeiterhebung für den gesamten Beobachtungszeitraum vorgesehen, ab Jahr 1B sollten im CRF nur bestimmte UE, nämlich SUE, UE, die möglicherweise oder wahrscheinlich in Verbindung mit der Administration des Prüfpräparats stehen, und neue oder sich verschlechternde UE in einer der 4 Kategorien (onkologische, hämatologische, neurologische Ereignisse und/oder Autoimmunerkrankungen) erfasst werden. Ab Studienprotokoll Version 2 (15. Juni 2018) wird nur angegeben, dass die Dokumentation von UE sich auf bestimmte UE fokussiert. Es bleibt unklar, inwieweit bzw. ab wann keine vollständige Erhebung der UE in der LTFU-Studie stattfand.

<sup>2)</sup> Im Modul 4 werden UE für den Zeitraum Injektion ins erste Auge bis Jahr 1B und Injektion ins erste Auge bis Jahr 5 (= Zeitpunkt der letzten Injektion + 1.825 Tage) angegeben. Jedoch konnten keine Angaben zu (medianen) Beobachtungszeiten identifiziert werden. Für die ursprüngliche Interventionsgruppe betrug die mittlere Beobachtungszeit 406,6 Tage (SD: 20,4) zwischen Tag erster Injektion und Jahr 1 nach zweiter Injektion [4].

<sup>3)</sup> Für die ursprüngliche Interventionsgruppe entspricht der Zeitraum „Erhalt der Injektion ins erste Auge bis Jahr 1B“ dem Beobachtungszeitraum der Studie 301. In diesem Zeitraum war eine vollständige Erhebung der UE vorgesehen. Für die ursprüngliche Kontrollgruppe werden UE ab Beginn der Studie LTFU (ab Injektion ins erste Auge) erfasst.



- 4) Studienindividuelle Einteilung der Schweregrade: Wenn möglich, sollte gemäß dem Studienprotokoll eine Schweregradeinteilung (Grad 1–4) in Orientierung der WHO-Toxizitätsskala erfolgen. Vom pU wurden ergänzend Schweregrade für ophthalmologische UE in Orientierung an diese Einteilung vorgenommen. UE Grad  $\geq 3$  wurden nur im Dossier Modul 4 post hoc berichtet und werden, wie in Studie 301, im Anhang dargestellt.
- 5) Neu oder sich verschlechternde UE der 4 genannten Kategorien wurden als UE von besonderem Interesse post hoc im Interimsstudienbericht 2020 definiert.
- 6) In Modul 4 post hoc berechnet: Jahr 5 entspricht dem Zeitpunkt der letzten Injektion + 1.825 Tagen.
- 7) Für die ursprüngliche Interventionsgruppe geht in die Darstellung „Erhalt Injektion ins erste Auge bis Jahr 5“ der Beobachtungszeitraum aus Studie 301 (ab Injektion ins erste Auge) ein. Für die ursprüngliche Kontrollgruppe werden UE ab Beginn der Studie LTFU (ab Injektion ins erste Auge) erfasst.
- 8) Gemäß Interimsstudienbericht 2020 wurden 2 UE (Parästhesie bei 1 Person in der ursprünglichen Kontrollgruppe; Krämpfe bei 1 Person der ursprünglichen Interventionsgruppe), die in den vorherigen Interimsstudienberichten nicht als AESI klassifiziert waren, als solche im neuen Studienbericht gekennzeichnet.

Abkürzungen: AESI: Adverse events of special interest; k. A.: keine Angabe; KI: Konfidenzintervall; Jahr 1B: Jahr 1 nach Injektion ins zweite Auge; LTFU: Long-term Follow-up; SD: Standardabweichung; SAP: Statistical Analysis Plan; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

## Unerwünschte Ereignisse

### Studie 301

In Tabelle 27 sind die UE nach SOC und PT der Studie 301 dargestellt. Für einzelne UE unterscheidet sich die Anzahl an Personen mit UE auf SOC- und PT-Ebene zwischen den Angaben im Dossier Modul 4 und den Angaben im Studienbericht.

Es konnte nur für folgende UE ein statistisch signifikanter Nachteil zuungunsten der Intervention beobachtet werden: SOC „Störungen des Blutes und lymphatischen Systems“, PT „Leukozytose“ (jeweils Intervention vs. Kontrolle: 45,0 % vs. 0,0 %: RR: k. A., p-Wert: 0,027).

Tabelle 27: Studie 301: UE mit Inzidenz  $\geq 15$  % der Sicherheitspopulation innerhalb eines Jahres nach Baseline (Sicherheitspopulation)

MedDRA-Systemorganklasse Preferred Term	Intervention: Voretigen Neparvovec (N = 20) <sup>1)2)</sup> n (%)	Kontrolle: Beobachtendes Abwarten (N = 9) <sup>1)2)</sup> n (%)	RR [95%-KI] <sup>3)</sup> p-Wert <sup>4)</sup>
<b>Störungen des Blutes und lymphatischen Systems</b>	9 (45,0)	0 (0,0)	k. A. <sup>4)</sup> 0,027
Leukozytose	9 (45,0)	0 (0,0)	k. A. <sup>4)</sup> 0,027
<b>Erkrankungen des Auges</b>	10 (50,0)	1 (11,1)	4,50 [0,95; 124,40] 0,096
Katarakt	3 (15,0)	0 (0,0)	k. A. <sup>4)</sup> 0,532
<b>Erkrankungen des Gastrointestinaltrakts<sup>5)</sup></b>	12 (60,0)	3 (33,0)	1,80 [0,76; 9,21] 0,245
Übelkeit <sup>6)</sup>	6 (30,0)	1 (11,1)	k. A. <sup>7)</sup>
Erbrechen <sup>8)</sup>	8 (40,0)	2 (22,2)	k. A. <sup>7)</sup>
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort<sup>9)</sup></b>	10 (50,0)	1 (11,1)	4,50 [0,95; 124,40] 0,096 <sup>7)</sup>
Fieber <sup>10)</sup>	7 (35,0)	1 (11,1)	k. A. <sup>7)</sup>

MedDRA-Systemorganklasse Preferred Term	Intervention: Voretigen Neparvovec (N = 20) <sup>1)2)</sup> n (%)	Kontrolle: Beobachtendes Abwarten (N = 9) <sup>1)2)</sup> n (%)	RR [95%-KI] <sup>3)</sup> p-Wert <sup>4)</sup>
<b>Infektionen und parasitäre Erkrankungen<sup>11)12)</sup></b>	11 (55,0)	4 (44,4)	1,24 [0,57; 6,56] 0,700 <sup>7)</sup>
Nasopharyngitis <sup>13)</sup>	7 (35,0)	2 (22,2)	k. A.
Infektion der oberen Atemwege	2 (10,0)	3 (33,3)	0,30 [0,03; 1,65] 0,287
<b>Verletzung; Vergiftung und durch Eingriffe bedingte Komplikationen</b>	5 (25,0)	2 (22,2)	1,13 [0,26; 7,34] 1,000
<b>Untersuchungen</b>	7 (35,0)	1 (11,0)	3,15 [0,60; 82,84] 0,372
Erhöhter Augeninnendruck	4 (20,0)	0 (0,0)	k. A. 0,280
<b>Erkrankungen des Nervensystems</b>	10 (50,0)	3 (33,3)	1,50 [0,60; 6,61] 0,453
Kopfschmerzen <sup>14)</sup>	7 (35,0)	2 (22,2)	k. A. <sup>7)</sup>
<b>Erkrankungen der Nieren und Harnwege</b>	3 (15,0)	1 (11,1)	1,35 [0,15; 34,78] 1,000
Hämaturie	3 (15,0)	1 (11,1)	1,35 [0,15; 34,78] 1,000
<b>Erkrankungen der Geschlechtsorgane und der Brustdrüse</b>	3 (15,0)	0 (0,0)	k. A. 0,532
<b>Erkrankungen der Atemwege, des Brustraums und Mediastinums<sup>15)</sup></b>	10 (50,0)	5 (55,6)	0,90 [0,43; 2,49] 1,000 <sup>7)</sup>
Husten <sup>16)</sup>	6 (30,0)	1 (11,0)	k. A. <sup>7)</sup>
Oropharyngealer Schmerz <sup>17)</sup>	6 (30,0)	4 (44,0)	k. A. <sup>7)</sup>

<sup>1)</sup> Im aktuellen Dossier konnten keine konkreten Angaben zu den Beobachtungszeiten identifiziert werden. Unter Berücksichtigung des vorherigen Verfahrens werden für die Interventionsgruppe zwischen erster Injektion bis ein Jahr nach der zweiten Injektion durchschnittlich 406,6 Tage (SD: 20,4) angegeben. In der Kontrollgruppe lag der durchschnittliche Zeitraum zwischen Baseline und Jahr 1C bei 354,8 Tagen (SD: 20,6). Der Unterschied zwischen beiden Gruppen betrug somit ca. 50 Tage und wird als hinreichend ähnlich angesehen. Die nicht-Zeit-adjustierten Effektschätzer werden als geeignet angesehen.

<sup>2)</sup> Unterschiedlicher Beginn des Erhebungszeitraums zwischen beiden Gruppen: Die Erhebung begann in der Interventionsgruppe ab der ersten Injektion und in der Kontrollgruppe ab der Baseline-Untersuchung. Der Unterschied beträgt im Mittel 34,3 Tage (SD 27,9) bzw. im Median 22 Tage (min; max: 6; 78).

<sup>3)</sup> Post hoc berechneter Effektschätzer mittels Vier-Felder-Tafel. Das KI wurde mittels exakter Methode berechnet sowie der p-Wert post hoc anhand des exakten Tests nach Fischer berechnet.

<sup>4)</sup> Keine Berechnung der Effektschätzer seitens des pU.

<sup>5)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für die SOC „Erkrankungen des Gastrointestinaltrakts“. Die in der Tabelle dargestellten Effektschätzer sind der vorherigen Nutzenbewertung entnommen [7]. Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 13 Personen (65,0 %) und in der Kontrollgruppe bei 3 Personen (33,3 %) auf. Das relative Risiko betrug 1,95 [95%-KI: 0,84; 12,91] und der p-Wert 0,226.

<sup>6)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für den PT „Übelkeit“: Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 7 Personen (35,0 %) und in der Kontrollgruppe bei 1 Person (11,1 %) auf. Das relative Risiko betrug 3,15 [95%-KI: 0,60; 82,84] und der p-Wert 0,371.

<sup>7)</sup> Bei diskrepanten Angaben zwischen dem aktuellen Dossier Modul 4 und Studienbericht wurden die Ergebnisse aus dem vorherigen Verfahren berichtet.

- <sup>8)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für den PT „Erbrechen“: Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 9 Personen (45,0 %) und in der Kontrollgruppe bei 2 Personen (22,2 %) auf. Das relative Risiko betrug 2,03 [95%-KI: 0,64; 15,92] und der p-Wert 0,412.
- <sup>9)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für die SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“. Die in der Tabelle dargestellten Effektschätzer sind der vorherigen Nutzenbewertung entnommen [7]. Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 12 Personen (60,0 %) und in der Kontrollgruppe bei 1 Person (11,1 %) auf. Das relative Risiko betrug 5,40 [95%-KI: 1,05; 154,72] und der p-Wert 0,020.
- <sup>10)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für den PT „Fieber“: Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 9 Personen (45,0 %) und in der Kontrollgruppe bei 1 Person (11,1 %) auf. Das relative Risiko betrug 4,05 [95%-KI: 0,80; 110,04] und der p-Wert 0,107.
- <sup>11)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für die SOC „Infektionen und parasitäre Erkrankungen“. Die in der Tabelle dargestellten Effektschätzer sind der vorherigen Nutzenbewertung entnommen [7]. Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 12 Personen (60,0 %) und in der Kontrollgruppe bei 5 Personen (55,6 %) auf. Das relative Risiko betrug 1,08 [95%-KI: 0,56; 3,47] und der p-Wert 1,000.
- <sup>12)</sup> Gemäß Studienbericht traten nur bei 1 Person (5,0 %) Infektionen der unteren Atemwege auf. Im Dossier Modul 4 werden jedoch 2 Personen (10,0 %) angegeben.
- <sup>13)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für den PT „Nasopharyngitis“: Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 7 Personen (35,0 %) und in der Kontrollgruppe bei 3 Personen (33,3 %) auf. Das relative Risiko betrug 1,05 [95%-KI: 0,34; 6,56] und der p-Wert 1,000.
- <sup>14)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für den PT „Kopfschmerz“: Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 8 Personen (40,0 %) und in der Kontrollgruppe bei 2 Personen (22,2 %) auf. Das relative Risiko betrug 1,80 [95%-KI: 0,54; 14,86] und der p-Wert 0,431.
- <sup>15)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für die SOC „Erkrankungen der Atemwege, des Brustraums und Mediastinums“. Die in der Tabelle dargestellten Effektschätzer sind der vorherigen Nutzenbewertung entnommen [7]. Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 13 Personen (65,0 %) und in der Kontrollgruppe bei 6 Personen (66,7 %) auf. Das relative Risiko betrug 0,98 [95%-KI: 0,54; 2,45] und der p-Wert 1,000.
- <sup>16)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für den PT „Husten“: Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 9 Personen (45,0 %) und in der Kontrollgruppe bei 2 Personen (22,2 %) auf. Das relative Risiko betrug 2,03 [95%-KI: 0,64; 15,92] und der p-Wert 0,412.
- <sup>17)</sup> Die Angaben unterscheiden sich zwischen Dossier Modul 4 und Studienbericht für den PT „Oropharyngealer Schmerz“: Gemäß Modul 4 trat das UE in der ursprünglichen Interventionsgruppe bei 7 Personen (35,0 %) und in der Kontrollgruppe bei 4 Personen (44,4 %) auf. Das relative Risiko betrug 0,79 [95%-KI: 0,30; 2,94] und der p-Wert 0,694.

Abkürzungen: k. A.: keine Angaben; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; PT: Preferred Term; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SD: Standardabweichung; SOC: Systemorganklasse; UE: Unerwünschte/s Ereignis/se.

### Studie LTFU

In Tabelle 28 sind die UE mit Inzidenz  $\geq 10$  % der ehemaligen Interventions- und Kontrollgruppe bis Jahr 1B bzw. Jahr 5 dargestellt. Die Anzahl an Personen mit UE ist auf SOC- und PT-Ebene für die Beobachtungszeiträume zu Jahr 1B und bis Jahr 5 sehr ähnlich, nach Jahr 1B wurden lediglich für einzelne SOC und PT zusätzliche Ereignisse berichtet.

Es finden sich für einzelne UE nach SOC und PT diskrepante Angaben zwischen Interimsstudienbericht und Dossier Modul 4 für den Zeitraum bis Jahr 1B. Für den Zeitraum bis Jahr 5 konnten nur Angaben im Dossier Modul 4 identifiziert werden.

Tabelle 28: Studie LTFU: UE<sup>1)</sup> mit Inzidenz  $\geq 15\%$  bis Jahr 1B bzw. Jahr 5 (Sicherheitspopulation)

MedDRA-Systemorganklasse Preferred Term	Bis Jahr 1B <sup>2)</sup>		Bis Jahr 5 <sup>2)3)</sup>	
	Ursprüngliche Intervention (N = 20) n (%)	Ursprüngliche Kontrolle (N = 9) n (%)	Ursprüngliche Intervention (N = 20) n (%)	Ursprüngliche Kontrolle (N = 9) n (%)
<b>Störungen des Blutes und lymphatischen Systems</b>	9 (45,0)	2 (22,2)	9 (45,0)	2 (22,2)
Leukozytose	9 (45,0)	2 (22,2)	9 (45,0)	2 (22,2)
<b>Erkrankungen des Auges</b>	10 (50,0)	6 (66,7)	11 (55,0)	6 (66,7)
Katarakt	3(15,0) <sup>4)</sup>	0 (0,0)	5 (25,0)	1 (11,1)
Netzhautablagerungen	0 (0,0)	3 (33,3)	0 (0,0)	3 (33,3)
<b>Erkrankungen des Gastrointestinaltrakts</b>	12 (60,0)	5 (55,6)	12 (60,0)	5 (55,6)
Übelkeit	6 (30,0)	4 (44,4)	6 (30,0)	4 (44,4)
Erbrechen	8 (40,0)	2 (22,2)	8 (40,0)	2 (22,2)
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	10 (50,0)	2 (22,2)	10 (50,0)	2 (22,2)
Fieber	7 (35,0)	2 (22,2)	7 (35,0)	2 (22,2)
<b>Infektionen und parasitäre Erkrankungen</b>	11 (55,0)	3 (33,3)	11 (55,0)	3 (33,3)
Nasopharyngitis	7 (35,0)	1 (11,1)	7 (35,0)	1 (11,1)
<b>Verletzung; Vergiftung und durch Eingriffe bedingte Komplikationen</b>	5 (25,0)	2 (22,2)	5 (25,0)	2 (22,2)
<b>Untersuchungen</b>	7 (35,0)	2 (22,2)	7 (35,0)	2 (22,2)
Erhöhter Augeninnendruck	4 (20,0)	1 (11,1)	4 (20,0)	1 (11,1)
<b>Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen</b>	1 (5,0)	2 (22,2)	1 (5,0)	2 (22,2)
<b>Erkrankungen des Nervensystems</b>	10 (50,0)	6 (66,7)	10 (50,0)	6 (66,7)
Kopfschmerzen	7 (35,0)	6 (66,7)	7 (35,0)	6 (66,7)
<b>Psychiatrische Erkrankungen</b>	2 (10,0)	2 (22,2)	2 (10,0)	2 (22,2)
Angst	0 (0,0)	2 (22,2)	0 (0,0)	2 (22,2)
<b>Erkrankungen der Nieren und Harnwege</b>	3 (15,0)	0 (0,0)	3 (15,0)	0 (0,0)
Hämaturie	3 (15,0)	0 (0,0)	3 (15,0)	0 (0,0)
<b>Erkrankungen der Geschlechtsorgane und der Brustdrüse</b>	3 (15,0)	0 (0,0)	3 (15,0)	0 (0,0)
<b>Erkrankungen der Atemwege, des Brustraums und Mediastinums</b>	10 (50,0) <sup>5)</sup>	4 (44,4)	10 (50,0)	4 (44,4)
Husten	6 (30,0) <sup>6)</sup>	2 (22,2)	6 (30,0)	2 (22,2)
Nasenschleimhautschwellung	2 (10,0)	2 (22,2)	2 (10,0)	2 (22,2)
Oropharyngealer Schmerz	6 (30,0)	1 (11,1)	6 (30,0)	1 (11,1)
<b>Erkrankungen der Haut und des Unterhautzellgewebes</b>	2 (10,0) <sup>7)</sup>	4 (44,4)	1 (5,0)	4 (44,4)

- 1) Gemäß SAP (22. Juni 2016) war in der LTFU-Studie eine vollständige Erhebung der UE nur bis Jahr 1B, jedoch nicht im Rahmen einer Langzeiterhebung für den gesamten Beobachtungszeitraum vorgesehen, ab Jahr 1B sollten im CRF nur bestimmte UE, nämlich SUE, UE, die möglicherweise oder wahrscheinlich in Verbindung mit der Administration des Prüfpräparats stehen, und neue oder sich verschlechternde UE in einer der 4 Kategorien (onkologische, hämatologische, neurologische Ereignisse und/oder Autoimmunerkrankungen) erfasst werden. Ab Studienprotokoll Version 2 (15. Juni 2018) wird nur angegeben, dass die Dokumentation von UE sich auf bestimmte UE fokussiert. Es bleibt unklar, inwieweit bzw. ab wann keine vollständige Erhebung der UE in der LTFU-Studie stattfand.
- 2) Im Modul 4 werden UE für den Zeitraum Injektion ins erste Auge bis Jahr 1B und Injektion ins zweite Auge bis Jahr 5 (= Zeitpunkt der letzten Injektion + 1.825 Tage) angegeben. Jedoch konnten keine Angaben zu (medianen) Beobachtungszeiten für diesen Untersuchungszeitraum identifiziert werden. Für die ursprüngliche Interventionsgruppe betrug die mittlere Beobachtungszeit 406,6 Tage (SD: 20,4) zwischen dem Tag der ersten Injektion und Jahr 1 nach der zweiten Injektion [4].
- 3) Für die ursprüngliche Interventionsgruppe geht in die Darstellung „Erhalt Injektion ins erste Auge bis Jahr 5“ der Beobachtungszeitraum aus der Studie 301 (ab Injektion ins erste Auge) ein. Für die ursprüngliche Kontrollgruppe werden UE ab Beginn der Studie LTFU (ab Injektion ins erste Auge) erfasst.
- 4) Gemäß Dossier Modul 4 wurde nur bei 1 Person (5,0 %) das UE nach PT „Katarakt“ beobachtet.
- 5) Gemäß Dossier Modul 4 wurde bei 9 Personen (45,0 %) das UE SOC „Erkrankungen der Atemwege, des Brustraums und Mediastinums“ beobachtet.
- 6) Gemäß Dossier Modul 4 wurde bei 5 Personen (25,0 %) das UE nach PT „Husten“ beobachtet.
- 7) Gemäß Dossier Modul 4 wurde nur bei 1 Person (5,0 %) das UE SOC „Erkrankungen der Haut und des Unterhautzellgewebes“ beobachtet.

Abkürzungen: CRF: Case report form; Jahr 1B: Jahr 1 nach Injektion ins zweite Auge; MedDRA: Medical Dictionary for Regulatory Activities; PT: Preferred Term; SAP: Statistischer Analyseplan; SD: Standardabweichung; SOC: Systemorganklasse; (S)UE: (Schwerwiegende) unerwünschte/s Ereignis/se.

## Schwerwiegende unerwünschte Ereignisse

### Studie 301

In Tabelle 29 sind die Ergebnisse zu den SUE der Studie 301 dargestellt. Es traten insgesamt nur wenige SUE und nur in der Interventionsgruppe auf.

Tabelle 29: Studie 301: Alle SUE innerhalb eines Jahres nach Baseline (Sicherheitspopulation)

MedDRA-Systemorganklasse Preferred Term	Intervention: Voretigen Neparvovec <sup>1)2)</sup> (N = 20) n (%)	Kontrolle: Beobachtendes Abwarten <sup>1)2)</sup> (N = 9) n (%)	RR [95%-KI] <sup>3)</sup> p-Wert <sup>3)</sup>
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	2 (10,0)	0 (0,0)	k. A., 1,000
Reaktionen auf das Arzneimittel	2 (10,0)	0 (0,0)	k. A., 1,000
<b>Erkrankungen des Nervensystems</b>	1 (5,0)	0 (0,0)	k. A., 1,000
Krampfanfall	1 (5,0)	0 (0,0)	k. A., 1,000

- 1) Im aktuellen Dossier konnten keine konkreten Angaben zu den Beobachtungszeiten identifiziert werden. Unter Berücksichtigung des vorherigen Verfahrens werden für die Interventionsgruppe zwischen Injektion ins erste Auge bis 1 Jahr nach Injektion ins zweite Auge durchschnittlich 406,6 Tage (SD: 20,4) angegeben. In der Kontrollgruppe lag der durchschnittliche Zeitraum zwischen Baseline und Jahr 1C bei 354,8 Tagen (SD: 20,6 [4]). Der Unterschied zwischen beiden Gruppen betrug somit ca. 50 Tage und wird als hinreichend ähnlich angesehen. Die nicht-zeit-adjustierten Effektschätzer werden als geeignet angesehen.
- 2) Unterschiedlicher Beginn des Erhebungszeitraums zwischen beiden Gruppen: Die Erhebung begann in der Interventionsgruppe ab der ersten Injektion und in der Kontrollgruppe ab der Baseline-Untersuchung. Der Unterschied beträgt im Mittel 34,3 Tage (SD 27,9) bzw. im Median 22 Tage (min; max: 6; 78).
- 3) Post hoc berechneter Effektschätzer mittels Vier-Felder-Tafel. Das KI wurde mittels exakter Methode berechnet. P-Wert berechnet post hoc anhand des exakten Tests nach Fischer.

Abkürzungen: k. A.: keine Angaben; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; RR: Relatives Risiko; SD: Standardabweichung; SUE: Schwerwiegende/s unerwünschte/s Ereignis/se.

### Studie LTFU

In Tabelle 30 sind die SUE der Studie LTFU bis Jahr 1B und Jahr 5 dargestellt. In beiden Behandlungsgruppen traten jeweils nur bei einzelnen Personen SUE auf.

Tabelle 30: Studie LTFU: alle SUE<sup>1)</sup> bis Jahr 1B bzw. Jahr 5 (Sicherheitspopulation)

MedDRA-Systemorganklasse Preferred Term	Bis Jahr 1B <sup>1)</sup>		Bis Jahr 5 <sup>1)</sup>	
	Ursprüngliche Intervention (N = 20) <sup>2)</sup> n (%)	Ursprüngliche Kontrolle (N = 9) <sup>2)</sup> n (%)	Ursprüngliche Intervention (N = 20) <sup>2)</sup> n (%)	Ursprüngliche Kontrolle (N = 9) <sup>2)</sup> n (%)
<b>Augenerkrankungen</b>	0 (0,0)	1 (11,1)	1 (5,0)	1 (11,1)
Netzhautablösung	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
Retinale Erkrankung	0 (0,0)	1 (11,1)	0 (0,0)	1 (11,1)
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
Unerwünschte Arzneimittelwirkung	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
<b>Infektionen und parasitäre Erkrankungen</b>	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
Pneumonie	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
<b>Erkrankungen des Nervensystems</b>	1 (5,0)	0 (0,0)	1 (5,0)	0 (0,0)
Krampfanfall	1 (5,0)	0 (0,0)	1 (5,0)	0 (0,0)
<b>Erkrankungen der Geschlechts- organe und der Brustdrüse</b>	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
Menorrhagie	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)

<sup>1)</sup> Für die ursprüngliche Interventionsgruppe geht in die Darstellung „Erhalt Injektion ins erste Auge bis Jahr 5“ der Beobachtungszeitraum aus der Studie 301 (ab Injektion ins erste Auge) ein. Für die ursprüngliche Kontrollgruppe werden UE ab Beginn der Studie LTFU (ab Injektion ins erste Auge) erfasst.

<sup>2)</sup> Im Modul 4 werden UE für den Zeitraum Injektion ins erste Auge bis Jahr 1B und Injektion ins erste Auge bis Jahr 5 (= Zeitpunkt der letzten Injektion + 1.825 Tage) angegeben. Jedoch konnten keine Angaben zu (medianen) Beobachtungszeiten für diesen Untersuchungszeitraum identifiziert werden. Für die ursprüngliche Interventionsgruppe betrug die mittlere Beobachtungszeit 406,6 Tage (SD: 20,4) zwischen Tag erster Injektion und Jahr 1 nach zweiter Injektion [4].

Abkürzungen: KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; SD: Standardabweichung; (S)UE: (Schwerwiegendes) Unerwünschte/s Ereignis/se.

### Schwere unerwünschte Ereignisse

In Tabelle 31 sind schwere UE gemäß Studienbericht mit studienindividueller Schweregrad-einteilung dargestellt. Milde und moderate UE wurden nicht dargestellt.

Tabelle 31: Studie 301: Alle schweren UE<sup>1)</sup> innerhalb eines Jahres nach Baseline (Sicherheitspopulation)

MedDRA-Systemorganklasse Preferred Term	Intervention: Voretigen Neparovvec <sup>2)3)</sup> (N = 20) n (%)	Kontrolle: Beobachtendes Abwarten <sup>2)3)</sup> (N = 9) n (%)	RR [95%-KI] <sup>4)</sup> p-Wert <sup>4)</sup>
<b>Erkrankungen des Herzens</b>	1 (5,0)	0 (0,0)	k. A., 1,000
Tachykardie	1 (5,0)	0 (0,0)	k. A., 1,000

MedDRA-Systemorganklasse Preferred Term	Intervention: Voretigen Neparovvec <sup>2)3)</sup> (N = 20) n (%)	Kontrolle: Beobachtendes Abwarten <sup>2)3)</sup> (N = 9) n (%)	RR [95%-KI] <sup>4)</sup> p-Wert <sup>4)</sup>
<b>Erkrankungen des Gastrointestinaltrakts</b>	2 (10,0)	0 (0,0)	k. A., 1,000
Übelkeit	2 (10,0)	0 (0,0)	k. A., 1,000
Erbrechen	2 (10,0)	0 (0,0)	k. A., 1,000
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	2 (10,0)	0 (0,0)	k. A., 1,000
Reaktionen auf das Arzneimittel	2 (10,0)	0 (0,0)	k. A., 1,000
Brustkorbschmerzen	1 (5,0)	0 (0,0)	k. A., 1,000
<b>Erkrankungen des Nervensystems</b>	1 (5,0)	0 (0,0)	k. A., 1,000
Krampfanfall	1 (5,0)	0 (0,0)	k. A., 1,000

<sup>1)</sup> Es erfolgte gemäß Studienprotokoll eine studienindividuelle Einteilung der UE in Schweregrade.

<sup>2)</sup> Im aktuellen Dossier konnten keine konkreten Angaben zu den Beobachtungszeiten identifiziert werden. Unter Berücksichtigung des vorherigen Verfahrens werden für die Interventionsgruppe zwischen erster Injektion bis 1 Jahr nach der zweiten Injektion durchschnittlich 406,6 Tage (SD: 20,4) angegeben. In der Kontrollgruppe lag der durchschnittliche Zeitraum zwischen Baseline und Jahr 1C bei 354,8 Tagen (SD: 20,6) [4]. Der Unterschied zwischen beiden Gruppen betrug somit ca. 50 Tage und wird als hinreichend ähnlich angesehen. Die nicht-Zeit-adjustierten Effektschätzer werden als geeignet angesehen.

<sup>3)</sup> Unterschiedlicher Beginn des Erhebungszeitraums zwischen beiden Gruppen: Die Erhebung begann in der Interventionsgruppe ab der ersten Injektion und in der Kontrollgruppe ab der Baseline-Untersuchung. Der Unterschied beträgt im Mittel 34,3 Tage (SD 27,9) bzw. im Median 22 Tage (min; max: 6; 78).

<sup>4)</sup> Berechnung des relativen Risikos anhand einer Vier-Felder-Tafel. Das KI wurde mit einer exakten Methode berechnet. Berechnung des p-Wertes mittels exakten Test nach Fischer.

Abkürzungen: k. A.: keine Angaben; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SD: Standardabweichung; UE: Unerwünschte/s Ereignis/se.

### Studie LTFU

In Tabelle 32 sind alle schweren UE der Studie LTFU bis Jahr 1B und Jahr 5 dargestellt. Es finden sich diskrepante Angaben zu den UE zwischen Interimsstudienbericht zu Jahr 1B und den Angaben aus dem Dossier Modul 4 für diesen Zeitraum. Es werden im Interimsstudienbericht folgende UE berichtet, die nicht im Dossier Modul 4 genannt sind: SOC „Kardiologische Erkrankungen“, PT „Tachykardie“, PT „Schmerzen in der Brust“.

In den Studienunterlagen konnten keine Ergebnisse zu den UE bis Jahr 5 identifiziert werden. Diese werden nur im Dossier Modul 4 dargestellt.

Schwere UE traten überwiegend bei einzelnen Personen in der ursprünglichen Interventionsgruppe auf.

Tabelle 32: Studie LTFU: alle schweren UE<sup>1)2)</sup> bis Jahr 1B bzw. Jahr 5 (Sicherheitspopulation)

MedDRA-Systemorganklasse Preferred Term	Bis Jahr 1B <sup>3)</sup>		Bis Jahr 5 <sup>4)</sup>	
	Ursprüngliche Intervention (N = 20) n (%)	Ursprüngliche Kontrolle (N = 9) n (%)	Ursprüngliche Intervention (N = 20) n (%)	Ursprüngliche Kontrolle (N = 9) n (%)
<b>Kardiologische Erkrankungen<sup>5)</sup></b>	1 (5,0)	1 (5,0)	k. A.	k. A.
Tachykardie <sup>5)</sup>	1 (5,0)	1 (5,0)	k. A.	k. A.

MedDRA-Systemorganklasse Preferred Term	Bis Jahr 1B <sup>3)</sup>		Bis Jahr 5 <sup>4)</sup>	
	Ursprüngliche Intervention (N = 20) n (%)	Ursprüngliche Kontrolle (N = 9) n (%)	Ursprüngliche Intervention (N = 20) n (%)	Ursprüngliche Kontrolle (N = 9) n (%)
<b>Erkrankungen des Gastrointestinaltrakts</b>	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
Übelkeit	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
Erbrechen	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
Unerwünschte Arzneimittelwirkung	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
Schmerzen in der Brust <sup>5)</sup>	1 (5,0)	0 (0,0)	k. A.	k. A.
<b>Infektionen und parasitäre Erkrankungen</b>	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
Pneumonie	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
<b>Erkrankungen des Nervensystems</b>	1 (5,5)	0 (0,0)	1 (5,5)	0 (0,0)
Krampfanfall	1 (5,5)	0 (0,0)	1 (5,5)	0 (0,0)
<b>Erkrankungen der Geschlechts- organe und der Brustdrüse</b>	0 (0,0)	0 (0,0)	1 (5,5)	0 (0,0)
Menorrhagie	0 (0,0)	0 (0,0)	1 (5,5)	0 (0,0)

<sup>1)</sup> Es erfolgte gemäß Studienprotokoll eine studienindividuelle Einteilung der UE in Schweregrade.

<sup>2)</sup> Gemäß SAP (22. Juni 2016) war in der LTFU-Studie eine vollständige Erhebung der UE nur bis Jahr 1B, jedoch nicht im Rahmen einer Langzeiterhebung für den gesamten Beobachtungszeitraum vorgesehen, ab Jahr 1B sollten im CRF nur bestimmte UE, nämlich SUE, UE, die möglicherweise oder wahrscheinlich in Verbindung mit der Administration des Prüfpräparats stehen, und neue oder sich verschlechternde UE in einer der 4 Kategorien (onkologische, hämatologische, neurologische Ereignisse und/oder Autoimmunerkrankungen) erfasst werden. Ab Studienprotokoll Version 2 (15. Juni 2018) wird nur angegeben, dass die Dokumentation von UE sich auf bestimmte UE fokussiert. Es bleibt unklar, inwieweit bzw. ab wann keine vollständige Erhebung der UE in der LTFU-Studie stattfand.

<sup>3)</sup> Im Modul 4 werden UE für den Zeitraum Injektion ins erste Auge bis Jahr 1B und Injektion ins erste Auge bis Jahr 5 (= Zeitpunkt der letzten Injektion + 1.825 Tage) angegeben. Jedoch konnten keine Angaben zu (medianen) Beobachtungszeiten für diesen Untersuchungszeitraum identifiziert werden. Für die ursprüngliche Interventionsgruppe betrug die mittlere Beobachtungszeit 406,6 Tage (SD: 20,4) zwischen dem Tag der ersten Injektion und Jahr 1 nach der zweiten Injektion [4].

<sup>4)</sup> Für die ursprüngliche Interventionsgruppe geht in die Darstellung „Erhalt Injektion ins erste Auge bis Jahr 5“ der Beobachtungszeitraum aus der Studie 301 (ab Injektion ins erste Auge) ein. Für die ursprüngliche Kontrollgruppe werden UE ab Beginn der Studie LTFU (ab Injektion ins erste Auge) erfasst.

<sup>5)</sup> Angaben zu den UE konnten nur im Interimsstudienbericht für den Zeitraum Injektion ins erste Auge bis Jahr 1B, nicht jedoch im Dossier Modul 4 identifiziert werden. Angaben zu Jahr 5 wurden im Dossier Modul 4 nicht vorgenommen.

Abkürzungen: CRF: Case report form; k. A.: keine Angaben; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; SAP: Statistischer Analyseplan; SD: Standardabweichung; (S)UE: (Schwerwiegende) unerwünschte/s Ereignis/se.

### Unerwünschte Ereignisse von besonderem Interesse

Für die Studie 301 waren keine AESI definiert.

Gemäß Interimsstudienbericht 2020 wurden für die LTFU-Studie 2 UE (Parästhesie bei 1 Person in der ursprünglichen Kontrollgruppe; Krämpfe bei 1 Person der ursprünglichen Interventionsgruppe), die in den vorherigen Interimsstudienberichten nicht als AESI klassifiziert waren, als solche im neuen Studienbericht gekennzeichnet.



## 4 Diskussion der Methodik und Ergebnisse

### 4.1 Zulassungsstatus und Zulassungspopulation von Voretigen Neparvovec

Voretigen Neparvovec ist zugelassen zur Behandlung von erwachsenen und pädiatrischen Patientinnen und Patienten mit Sehverlust aufgrund einer erblichen Netzhautdystrophie, die auf nachgewiesenen biallelischen *RPE65*-Mutationen beruht, und die über ausreichend lebensfähige Netzhautzellen verfügen. Die zulassungsrelevante Dosis beträgt  $1,5 \times 10^{11}$  Vg. Es wird sequentiell ein Volumen von je 0,3 ml ins Auge injiziert. Die individuelle Verabreichung pro Auge erfolgt an unterschiedlichen Tagen innerhalb eines kurzen Zeitintervalls, das jedoch mindestens 6 Tage betragen muss [11].

An der Studie 301 konnten Personen ab einem Alter von mindestens 3 Jahren teilnehmen. Es wurden Personen zwischen 4 und 44 Jahren in die Studie eingeschlossen; sie waren in beiden Behandlungsgruppen durchschnittlich 15 Jahre alt. Es wurden überwiegend weiße Personen (ca. 70 %) eingeschlossen. Die Teilnehmenden stammten zum Großteil aus dem nordamerikanischen Raum (86 % in der Interventionsgruppe und 60 % in der Kontrollgruppe), insgesamt nahmen ca. 16 % Personen aus dem europäischen Raum an der Studie teil. Eine Übertragbarkeit auf den deutschen Versorgungsraum für Personen zwischen 4 und 44 Jahren erscheint plausibel. Unklar bleibt die Übertragbarkeit bei Patientinnen/Patienten, die jünger bzw. älter sind. Es haben bis auf jeweils 1 Person in beiden Behandlungsgruppen alle Teilnehmenden die Studie 301 abgeschlossen und nahmen an der LTFU-Studie teil. Aus diesem Grund sind die Kriterien der Studie 301 auf die LTFU-Studie anwendbar.

Als Ein-/Ausschlusskriterien der Studie 301 wurden detaillierte ophthalmologische Kriterien vorgegeben, um gemäß dem pU sowohl eine ausreichend lebensfähige Netzhautfunktion sicherzustellen, aber auch um eine bestimmte Krankheitsschwere untersuchen zu können. Da an der LTFU-Studie Personen teilnahmen, die die Studie 301 abgeschlossen haben, galten diese Kriterien somit ebenfalls für die LTFU-Studie. Gemäß Modul 4 sollten Personen mit einer vollständig atrophierten Netzhaut, die von einer Therapie mit der Prüfmedikation nicht profitieren, ausgeschlossen werden, dabei aber Personen mit nur wenigen lebensfähigen Netzhautzellen trotzdem eine Therapie ermöglicht werden. Als Begründung für die Auswahl der Kriterien wird auf klinische Expertise verwiesen. Zusätzlich sollte der MLMT durchführbar sein. Dabei sollte sichergestellt werden, dass Personen mit einer annähernd normal funktionierenden Netzhaut (Bestehen bei 1 Lux), aber auch Personen, bei denen keine ausreichende Netzhautfunktion vorhanden ist (Nicht-Bestehen bei 400 Lux) an der Studie nicht teilnehmen können.

### 4.2 Design und Methodik der Studie

#### Studie 301

Bei der Studie 301 handelt es sich um eine randomisierte, multizentrische, offene Studie, bei der das Prüfpräparat im Vergleich zu beobachtendem Abwarten verglichen wurde. Die Randomisierung erfolgte im Verhältnis 2:1 in Form einer Block-Randomisierung (Blockgröße von 3), stratifiziert nach Alter der Personen ( $\geq 10$  vs.  $< 10$  Jahre) und Ergebnislevel des Mobilitätstests zum Screening mit dem am schlechtesten im Test abschneidenden Auge (Bestehen bei  $\geq 125$  vs.  $< 125$  Lux). Der primäre Endpunkt war definiert als die Veränderung im MLMT bilateral gemessen nach Jahr 1B/C zwischen den Behandlungsgruppen. Bis auf den primären Endpunkt (MLMT) fand keine Verblindung der Bewertung der Testergebnisse statt.

Es wurden 21 Personen in die Interventions- und 10 Personen in die Kontrollgruppe eingeschlossen. Die Studie wurde von 20 Personen in der Interventionsgruppe und von

9 Personen zu Jahr 1 in der Kontrollgruppe beendet. 2 Personen (jeweils eine in beiden Behandlungsarmen) schieden bereits vor Bekanntgabe der Behandlungszuteilung aus. Der Großteil der Personen (ca. 90 %) hat die zulassungsrelevante Dosierung bzw. das Volumen des Prüfpräparats in der Interventionsgruppe erhalten.

Die Randomisierung erfolgte stratifiziert nach 2 Faktoren, während die statistischen Analysen gemäß dem pU aufgrund der geringen Fallzahl (N = 31) und damit verbundener möglicher Konvergenzproblemen, nicht stratifiziert durchgeführt wurden.

Die Studiendauer zwischen Screening und Studienende betrug in der Interventionsgruppe (Screening bis Jahr 1B (= Jahr 1 nach Injektion ins zweite Auge)) im Median 475 Tage (min; max: 412; 593) und in der Kontrollgruppe (Screening bis Tag der ersten Injektion) im Median 420 Tage (378; 568). Angaben zur (medianen) Beobachtungsdauer ab Randomisierung/Baseline bis Jahr 1B/C konnten im aktuellen Dossier nicht identifiziert werden. Im vorherigen Verfahren wurde für die Interventionsgruppe eine durchschnittliche Beobachtungszeit für den Zeitraum der ersten Injektion bis ein Jahr nach der zweiten Injektion (Jahr 1B) von 406,6 Tagen (SD: 20,4) angegeben. In der Kontrollgruppe lag der durchschnittliche Zeitraum zwischen Baseline und Visite zu Jahr 1 (Jahr 1C) bei 354,8 Tagen (SD: 20,6) [4].

Aufgrund des offenen Studiendesigns wird das Verzerrungspotential als hoch eingeschätzt. Es finden sich zudem Hinweise auf Unterschiede in der Studiendurchführung zwischen den Zentren, die in der Gesamtschau zu einer Verzerrung der Ergebnisse führen können. Generell wurden Protokollabweichungen häufiger in der Interventionsgruppe (81 %) im Vergleich zur Kontrollgruppe (50 %) beobachtet, die größtenteils auf Abweichungen in der Vorgehensweise (SOP/MOP) zurückzuführen waren (Unterschied zwischen den Gruppen: 47 %). Es wurden mehr Abweichungen von den SOP/MOP für das Studienzentrum in Iowa als für das Studienzentrum CHOP berichtet. Zudem zeigte sich eine ungleiche Verteilung der Kontrollpopulation (N = 10) auf die Studienzentren (CHOP: 80 % vs. Iowa: 20 %), während diese in der Interventionsgruppe (N = 21) ähnlich war (CHOP: 52 % vs. Iowa: 48 %). Die ungleiche Verteilung der Teilnehmenden der Kontrollgruppe zwischen den Zentren kann möglicherweise auf die Blockrandomisierung zurückzuführen sein.

Des Weiteren konnten Imbalancen in der Begleitmedikation mit einem > 40 Prozentpunkte höheren Anteil in der Interventionsgruppe beobachtet werden: Schmerzmittel, Anticholinergika, Kortikosteroide + Antiinfektiva in Kombination. Da dies auf die Applikation des Prüfpräparats in der Interventionsgruppe im Gegensatz zum beobachtenden Abwarten in der Kontrollgruppe rückführbar ist, ergeben sich dadurch keine Hinweise auf ein erhöhtes Verzerrungspotential.

### Studie LTFU

Es handelt sich hierbei um eine multizentrische, nichtrandomisierte Beobachtungsstudie. Es soll die Langzeitsicherheit und -Wirksamkeit des Prüfpräparats bei Patientinnen und Patienten untersucht werden, die eine Injektion bereits in vorherigen Studien erhalten haben. Dadurch können Teilnehmende der Studien 101 bzw. 102 und 301 an dieser Studie teilnehmen. Für die Nutzenbewertung werden nur Ergebnisse der Teilnehmenden, die vorab an der Studie 301 teilgenommen haben, dargestellt. Personen der ursprünglichen Interventionsgruppe der Studie 301 haben das Prüfpräparat bereits zu Beginn der Studie 301 erhalten. Nach Abschluss der Studie 301 konnten Personen der Kontrollgruppe, sofern sie die Ein- und Ausschlusskriterien der Studie 301 zu diesem Zeitpunkt erfüllten, das Prüfpräparat erhalten und an der LTFU-Studie teilnehmen. Die Vorgaben zu den Injektionen in beide Augen entsprachen denen der ursprünglichen Interventionsgruppe.

Alle Teilnehmenden, die die Studie 301 erfolgreich abgeschlossen haben, nehmen an der LTFU-Studie teil: 20 Personen der ursprünglichen Interventionsgruppe und 9 Personen der ursprünglichen Kontrollgruppe. Alle Teilnehmenden der ursprünglichen Kontrollgruppe erhielten die Intervention. Die Studiendauer soll 15 Jahre betragen. Die Studie ist derzeit nicht abgeschlossen. Es sind jährliche Datenschnitte geplant. Für das aktuelle Dossier wurden Ergebnisse für den Datenschnitt vom 30. Juni 2020 vorgelegt. Zu diesem Zeitpunkt waren für beide Behandlungsgruppen Daten zu Jahr 5 von > 70 % verfügbar. Es werden daher für die Nutzenbewertung die Ergebnisse zu Jahr 5 für beide Behandlungsgruppen dargestellt.

Aufgrund des einarmigen Studiendesigns wird das Verzerrungspotential als hoch angesehen. Die Aussagesicherheit der Ergebnisse ist aufgrund der deskriptiven, nicht vergleichenden Analysen und geringer Fallzahl eingeschränkt.

Gemäß dem SAP Version 1.0 (2. Juni 2016) waren nur deskriptive Auswertungen vorgesehen. Es wurden im Laufe der Studie 3 Protokollamendments erstellt, die letzte Version 3.0 ist vom 30. Mai 2019. Für die Version 2.0 vom 15. Juni 2018 konnte keine Übersicht zu den Änderungen in den Studienunterlagen identifiziert werden. Ersichtlich waren jedoch im Vergleich zur ersten Protokollversion Änderungen in der Erfassung/Dokumentation der Sicherheitsendpunkte (siehe Abschnitt zur Sicherheit).

Es wurden Beobachtungszeiten für den Zeitraum von Randomisierung/Baseline bzw. Injektion ins erste Auge bis zum Datenschnitt vom 30. Juni 2020 vorgelegt. Die mediane Beobachtungszeit zwischen Randomisierung und Datenschnitt war in der ursprünglichen Interventionsgruppe im Median um ca. 80 Tage kürzer als in der ursprünglichen Kontrollgruppe, während der Unterschied zwischen dem Zeitpunkt der Injektion ins erste Auge bis Datenschnitt im Median um ca. 365 Tage länger war, was dem Studienzeitraum der Studie 301 entspricht. Es wurden keine Angaben zur (medianen) Beobachtungszeit zwischen Injektion ins erste Auge bis Jahr 5 für beide Behandlungsgruppen vorgelegt.

Gemäß Studienprotokoll wird für die Wirksamkeitsendpunkte ein Bewertungsparadigma (im Studienprotokoll als „evaluation paradigm“ bezeichnet), bezogen auf die Auswahl der Tests und deren Durchführung, individuell an die Patientinnen und Patienten definiert. Die Bewertung ist gemäß diesem Paradigma von den individuellen Zielen, dem Grad der Sehfunktion, dem Grad des Verständnisses der Durchführungen, der Ausprägung des Nystagmus und der Kooperationsfähigkeit und Geduld der Patientinnen und Patienten bei längeren Testsitzungen abhängig. Es ist unklar, inwieweit dies die standardisierte und einheitliche Erhebung gemäß SOP in der Studie und somit die generierten Ergebnisse beeinflusst hat.

Für beide Studien werden bilaterale Testszenarien bzw. Ergebnisse für beide Augen gemittelt, gegenüber anderer Testszenarien mit einzelnen Augen als übergeordnet bewertet, da diese der Alltagssituation von betroffenen Personen am nächsten ist.

### **4.3 Mortalität**

Todesfälle wurden als UE während der Studie im Rahmen der Sicherheit erfasst. Es traten während der Erfassung von UE in beiden Studien keine Todesfälle auf.

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

Es traten in beiden Studien (301 und LTFU) keine Todesfälle auf.

## 4.4 Morbidität

### Multi-Luminanz-Mobilitätstest

Der Multi-Luminanz-Mobilitätstest (MLMT) dient dazu, Veränderungen im funktionalen Sehen zu messen, insbesondere die Fähigkeit, sich unter verschiedenen Lichtverhältnissen in einem Hindernisparcours zu orientieren und selbstständig bewegen zu können. Maßgeblich für die Mobilität war die Beleuchtungsstärke, bei der die Patientin oder der Patient den Mobilitätstest in der entsprechenden Testsituation noch bestehen konnte. Dieser Lux-Wert wurde dann für jede Testsituation in einen Mobilitätstest-Score zwischen 6 (1 Lux) und -1 (400 Lux) umgewandelt, wobei ein höherer Mobilitätsscore einer besseren Mobilität entspricht. Die standardisierten Beleuchtungsstärken waren Alltagssituationen angeglichen.

#### Studie 301

Die Analyse der Veränderung im Mobilitätsscore, bilateral gemessen, zu Jahr 1B/C im Vergleich zu Baseline zwischen den Behandlungsgruppen war als primärer Endpunkt definiert und zeigte anhand der ITT-Population einen statistisch signifikanten Unterschied zugunsten der Interventionsgruppe ( $p < 0,001$ ). Dieses Ergebnis wird durch Sensitivitätsanalysen bestätigt. Beim post hoc berechneten SMD nach Hedges' g lag der Wert über der Irrelevanzschwelle von -0,2 bzw. 0,2. Das Vorliegen eines klinisch relevanten Effekts erscheint daher plausibel. Keine Person in der Interventionsgruppe hat beim Test zu Jahr 1 schlechter abgeschnitten als zu Baseline.

Eine ähnliche Verteilung der Teilnehmenden auf die Beleuchtungsstärken und damit verbundenen Deckeneffekte konnte in der Validierungsstudie und in der Studie 301 zu Baseline bzw. zu Jahr 1 beobachtet werden. Im Studienbericht wurde angemerkt, dass bei 40 % der Patientinnen/Patienten der Kontrollgruppe ein positives Ergebnis beobachtet wurde, was u. a. durch Lerneffekte, eine verbesserte Fähigkeit, sich zu bewegen, Variabilität in der Beurteilung erklärt wurde. Es ist nicht auszuschließen, dass diese Faktoren auch zu einer Verbesserung in der Interventionsgruppe beigetragen haben.

Die Testdurchführung wurde anhand der SOP umfassend beschrieben und lässt auf ein standardisiertes Vorgehen schließen, das die Teilnahme von Kindern berücksichtigt. Auch ist die verblindete Bewertung bei offenem Studiendesign zu begrüßen. Zudem wurden Maßnahmen zur Qualitätssicherung während der Studie bei der verblindeten Bewertung der Testdurchläufe und eine Randomisierung der Hindernisparcours zur Reduktion von Lerneffekten angewendet. In der Validierungsstudie konnten die Reliabilität, Validität und Änderungssensitivität nachgewiesen werden, wenn auch Unsicherheiten bei einigen Faktoren bestehen (u. a. Entwicklung des Parcours, Deckeneffekte, unzureichende Untersuchung bei einzelnen Beleuchtungsstärken/Mobilitätsscores).

Aufgrund des offenen Studiendesigns wird das Verzerrungspotential als hoch angesehen. So kann die Kenntnis der Behandlungszuordnung die Durchführung des Tests durch den/die Patienten/Patientin beeinflusst haben. Der Anteil an Personen mit prozedurbedingten Abweichungen war gemäß Dossier in der Interventionsgruppe größer als in der Kontrollgruppe. Insgesamt kam es zu 5 Prozedurabweichungen in der Interventionsgruppe (3 zu Baseline und 2 zu Jahr 1) und zu 1 in der Kontrollgruppe. Konkrete Angaben zur Anzahl an Personen, bei denen die genannten Abweichungen auftraten, konnten nicht identifiziert werden. Es bleibt unklar, inwieweit der höhere Anteil an Abweichungen in der Interventionsgruppe zu einer Verzerrung des Ergebnisses führt.

### Studie LTFU

Die deskriptiven, nicht vergleichenden Ergebnisse zu den jährlichen Visiten im Vergleich zu Baseline für beide Behandlungsgruppen sind konsistent mit den Ergebnissen bis Jahr 5. Es zeigt sich ein numerischer Anstieg in gleicher Größenordnung der Werte in beiden ursprünglichen Behandlungsgruppen.

Aufgrund des einarmigen, offenen Studiendesigns und geringer Fallzahl wird das Verzerrungspotential als hoch angesehen. Es bleibt unklar, inwieweit eine Anpassung der Testdurchführung an die individuelle Situation der Teilnehmenden vorgenommen wurde und dies die Standardisierung und Vergleichbarkeit der Erhebung des Endpunkts während der Studie LTFU und zwischen beiden Studien beeinflusste.

### **Sehschärfe (ETDRS- und HOTV-Sehtafel)**

Die Sehschärfe wird als unmittelbar patientenrelevant angesehen. Die ETDRS-Sehtafel ist ein etabliertes Instrument und wird als valide und geeignet zur Erhebung und Untersuchung der Sehschärfe angesehen. Der HOTV-Test ist ein verbreitetes Instrument zur Messung der Sehschärfe bei Kindern. Es bleibt unklar, inwieweit beide Tafeln austauschbar sind und somit eine gemeinsame Auswertung als geeignet angesehen werden kann.

### Studie 301

Der Endpunkt wurde in Abhängigkeit der kognitiven Fähigkeiten des Kindes entweder mit der ETDRS-Sehtafel oder mit der HOTV-Sehtafel erfasst. Für das aktuelle Dossier konnten keine Angaben zur Anzahl an Personen, bei denen die HOTV-Tafel eingesetzt wurde, identifiziert werden. Gemäß dem vorherigen Verfahren wurde die ETDRS-Sehtafel bei 18 Personen der Interventionsgruppe sowie 8 Personen in der Kontrollgruppe angewandt. Die HOTV-Sehtafel wurde bei 3 Personen der Interventionsgruppe und 1 Person der Kontrollgruppe eingesetzt [4]. Insgesamt wurden bei 4 Personen Off-Chart-Messungen durchgeführt. Bei 2 Personen der Interventionsgruppe wurden diese für den gesamten Studienverlauf vorgenommen. 2 weitere Personen sind nach der Baseline-Visite ausgeschieden.

Trotz der oben beschriebenen Unklarheiten werden die Ergebnisse der a priori definierten Analyse anhand eines MMRM mit den Termen Behandlung, Studienvisite und Behandlung\*Studienvisite für beide Sehtafeln zusammen in der Nutzenbewertung dargestellt. Es konnte zwischen den Behandlungsgruppen zu Jahr 1 kein statistisch signifikanter Unterschied zwischen den beiden Behandlungsgruppen gezeigt werden (mittlere Änderung Intervention - Kontrolle: -0,2 [95%-KI: -0,4; 0,1]; p-Wert: 0,170). Dieses Ergebnis war über den gesamten Studienverlauf bis zu Jahr 1 nach Baseline konsistent. Sensitivitätsanalysen anhand der mITT-Population mit oder ohne Imputationen ergaben vergleichbare Ergebnisse.

Im Modul 4 wurden post hoc Responderanalysen für eine Verbesserung und Verschlechterung von  $\geq 10$  bzw.  $\geq 15$  Buchstaben für beide Sehtafeln zusammen und separat für eine Verbesserung um  $\geq 10$  bzw. 15 Buchstaben sowohl für die ETDRS- als auch HOTV-Sehtafel eingereicht. Studien zur Übertragbarkeit der MID von ETDRS- zur HOTV-Sehtafel bzw. zur Untersuchung einer MID von  $\geq 10$  bzw.  $\geq 15$  Buchstaben für die HOTV-Sehtafel wurden nicht vorgelegt. Aus diesen Gründen werden nur die Responderanalysen für die ETDRS-Sehtafel berücksichtigt. Es wurden vom pU keine Effektschätzer oder p-Werte vorgelegt. Es konnte keine Verbesserung um  $\geq 10$  bzw.  $\geq 15$  Buchstaben in der Kontrollgruppe bis Jahr 1 beobachtet werden, dafür in der Interventionsgruppe bei 29 % der Teilnehmenden um  $\geq 10$  Buchstaben bzw. bei 19 % der Teilnehmenden um  $\geq 15$  Buchstaben.

Das Verzerrungspotential wird aufgrund des offenen Studiendesigns und der fehlenden Verblindung bei der Auswertung als hoch angesehen. Unklar bleibt, inwieweit beide Sehtafeln austauschbar sind. Angaben hierzu wurden seitens des pU nicht gemacht.

### Studie LTFU

Es wurden deskriptive Ergebnisse für beide Sehtafeln (ETDRS- und HOTV-Sehtafel) zusammen und ggf. Off-Chart-Analysen zur mittleren Veränderung, gemittelt über beide Augen, zu den Visiten im Vergleich zu Baseline dargestellt. Für die LTFU-Studie konnte keine separate Responderanalyse zur Verbesserung bzw. Verschlechterung um  $\geq 10$  bzw.  $\geq 15$  Buchstaben für die ETDRS-Tafel identifiziert werden.

Die deskriptiven, nicht vergleichenden Ergebnisse zu den jährlichen Visiten im Vergleich zu Baseline sind für beide Behandlungsgruppen konsistent mit den Ergebnissen zu Jahr 5. Es zeigte sich zumeist eine numerische Reduktion der Werte in beiden ursprünglichen Behandlungsgruppen. Gemäß dem Interimsstudienbericht wurden für 2 Teilnehmende Off-Chart-Ergebnisse zu der Visite Jahr 5 verwendet.

Aufgrund des einarmigen, offenen Studiendesigns wird das Verzerrungspotential als hoch angesehen. Es bleibt unklar, inwieweit eine Anpassung der Testdurchführung an die individuelle Situation der Teilnehmenden vorgenommen wurde und dies die Standardisierung und Vergleichbarkeit der Erhebung des Endpunkts während der Studie und zwischen beiden Studien beeinflusste. Version 3 der SOP wurde am 13.01.2014 erstellt, nachdem die erste Person am 15.11.2012 in die Studie eingeschlossen wurde. Änderungen in den SOP-Versionen können nicht nachvollzogen werden, da hierzu keine Informationen vorgelegt wurden.

### **Lichtempfindlichkeit (FST)**

Der Vollfeld-Lichtempfindlichkeits-Schwellenwert-Test (Full-light Sensitivity Test, FST) hat das Ziel die subjektive Lichtempfindlichkeit des gesamten Gesichtsfeldes, bei der die Testperson noch sehen kann, zu erfassen. Die Messung der Lichtempfindlichkeit wird als patientenrelevant angesehen.

Untersuchungen zur Reliabilität und Konstruktvalidität liegen für Erwachsene, nicht jedoch für Kinder vor. Informationen zur Änderungssensitivität konnten nicht identifiziert werden. Eine abschließende Beurteilung unter Berücksichtigung der pädiatrischen Patientenpopulation ist daher derzeit nicht möglich.

### Studie 301

Beim FST zeigen sich statistisch signifikante Unterschiede zugunsten der Intervention für das weiße (sekundärer Endpunkt), blaue und rote Licht (zusätzliche Endpunkte). Beim post hoc berechneten SMD nach Hedges'  $g$ , berechnet für das weiße Licht, lag der Wert über der Irrelevanzschwelle von  $-0,2$  bzw.  $0,2$ . Das Vorliegen eines klinisch relevanten Effekts erscheint daher plausibel.

Das Verzerrungspotential wird aufgrund des offenen Studiendesigns als hoch angesehen. Kritisch anzumerken ist die unverblindete Durchführung und Bewertung der Testergebnisse durch die Untersucherin / den Untersucher bei offenem Studiendesign. Zudem konnte der Test wiederholt werden, bis akzeptable Reliabilitätswerte erreicht wurden. Bei der Entscheidung zur wiederholten Messung zur Erreichung eines zuverlässigeren Testergebnisses kann neben den in der SOP angegebenen Kriterien zur Reliabilitätsbeurteilung auch das subjektive Ermessen des Untersuchers / der Untersucherin einen Einfluss ausüben. Es konnten keine Angaben identifiziert werden, weshalb für die Messung mit blauem und rotem

Licht Angaben für 3 Personen in der Interventionsgruppe fehlen, obwohl gemäß SAP eine Auswertung anhand der ITT-Population vorgesehen war. Version 3 der SOP wurde am 13. Januar 2014 erstellt, mehr als ein Jahr nachdem die erste Person am 15. November 2012 in die Studie eingeschlossen wurde. Änderungen in den Versionen der SOP können nicht nachvollzogen werden, da hierzu keine Daten eingereicht wurden.

### Studie LTFU

Die Ergebnisse zu den jährlichen Visiten sind konsistent mit den Ergebnissen zu Jahr 5 für beide Behandlungsgruppen. Es zeigte sich eine numerische Reduktion der Werte in beiden ursprünglichen Behandlungsgruppen.

Aufgrund des einarmigen, offenen Studiendesigns wird das Verzerrungspotential als hoch angesehen. Es bleibt unklar, inwieweit eine Anpassung der Testdurchführung an die individuelle Situation der Teilnehmenden vorgenommen wurde und dies die Standardisierung und Vergleichbarkeit der Erhebung des Endpunkts während der Studie LTFU und zwischen beiden Studien beeinflusste.

### **Perimetrie nach Goldmann bzw. Humphrey**

Zur Messung des Gesichtsfeldes wurden in beiden Studien sowohl statische (nach Humphrey) als auch kinetische Untersuchungsmethoden (nach Goldmann) eingesetzt. Bei beiden Perimetrieverfahren handelt es sich um in der klinischen Praxis weit verbreitete Methoden zur Messung des Gesichtsfeldes. Die Goldmann-Perimetrie erfasst das gesamte Gesichtsfeld, während die Humphrey-Perimetrie sich auf bestimmte Regionen im Gesichtsfeld fokussiert. Eine Ausdehnung des Gesichtsfelddefektes bzw. der Gesichtsfeldeinschränkungen werden als patientenrelevant erachtet.

Da die Prüfmedikation in eine bestimmte Region (Makula) des Auges appliziert wird, wurden bei der Humphrey-Perimetrie Unterschiede in der Funktion in dieser Region vor und nach Applikation der Prüfmedikation untersucht. Da eine Applikation in der Fovea-Region vermieden werden sollte, wurde ebenfalls eine Untersuchung in dieser Region vorgenommen.

### Studie 301

Bei der Perimetrie nach Goldmann mit dem Stimulus III4e, der regelhaft während der Studie eingesetzt werden sollte, zeigte sich bei der statistischen Analyse ein statistisch signifikanter Unterschied zugunsten der Intervention (mittlere Änderung Intervention - Kontrolle: 378,7 (95%-KI: [145,5; 612,0]; p-Wert: 0,006). Beim post hoc berechneten SMD nach Hedges' g lag der Wert über der Irrelevanzschwelle von -0,2 bzw. 0,2. Das Vorliegen eines klinisch relevanten Effekts erscheint daher plausibel. Der Stimulus V4e sollte gemäß SOP zur Durchführung der Perimetrie nach Goldmann nur eingesetzt werden, sofern der Test mit dem kleineren Stimulus III4e nicht möglich war. Da der Stimulus V4e bei < 70 % der Teilnehmenden zu Jahr 1 in beiden Behandlungsgruppen eingesetzt wurde, wird auf die Darstellung der Ergebnisse verzichtet.

Bei der Perimetrie nach Humphrey konnte im Bereich der Makula ein statistisch signifikanter Unterschied zugunsten der Intervention gezeigt werden. Beim post hoc berechneten SMD nach Hedges' g lag der Wert über der Irrelevanzschwelle von -0,2 bzw. 0,2. Das Vorliegen eines klinisch relevanten Effekts erscheint daher plausibel. Im Bereich der Fovea konnte kein statistisch signifikanter Unterschied gezeigt werden.

Aufgrund des offenen Studiendesigns wird das Verzerrungspotential als hoch angesehen. Es wurde ein statistisches Analyseverfahren, analog zum MLMT, verwendet, das erst post hoc im Studienbericht/Modul 4 beschrieben wurde.

### Studie LTFU

Für die Goldmann-Perimetrie mit Stimulus III4e zeigte sich zu allen Zeitpunkten ein numerischer Anstieg der Werte im Vergleich zu Baseline in beiden ursprünglichen Behandlungsgruppen. Zwischen Jahr 1 und Jahr 5 fluktuieren (ursprüngliche Kontrollgruppe) oder verringern sich die Werte (ursprüngliche Interventionsgruppe).

Für die Perimetrie nach Humphrey im Bereich der Fovea und der Macula zeigten sich zu allen Zeitpunkten höhere Werte im Vergleich zu Baseline in beiden ursprünglichen Behandlungsgruppen. Für den Bereich der Macula nehmen zwischen Jahr 1 und Jahr 5 die Werte kontinuierlich ab (ursprüngliche Interventionsgruppe) oder zu (ursprüngliche Kontrollgruppe) und für den Bereich der Fovea fluktuieren die Werte (ursprüngliche Interventionsgruppe) oder nehmen kontinuierlich zu bzw. bleiben ab Jahr 3 in einer vergleichbaren Größenordnung (ursprüngliche Kontrollgruppe).

Aufgrund des einarmigen, offenen Studiendesigns wird das Verzerrungspotential als hoch angesehen. Es bleibt unklar, inwieweit eine Anpassung der Testdurchführung an die individuelle Situation der Teilnehmenden vorgenommen wurde und dies die Standardisierung und Vergleichbarkeit der Erhebung des Endpunkts während der Studie LTFU und zwischen beiden Studien beeinflusste.

### **Zusammenfassende Einschätzung zur Morbidität**

Für die Studie 301 zeigten sich statistisch signifikante Unterschiede zugunsten der Intervention zu Jahr 1B/C im Vergleich zu Baseline für den MLMT, die FST, bei der Perimetrie nach Goldmann für den Stimulus III4e und bei der Perimetrie nach Humphrey für den Bereich der Macula. Die Konfidenzintervalle der post hoc berechneten SMD nach Hedges' g liegen außerhalb des Bereichs von -0,2 bis 0,2, sodass das Vorliegen eines klinisch relevanten Effekts plausibel erscheint. Für die Perimetrie im Bereich der Fovea zeigten sich keine statistisch signifikanten Unterschiede zwischen beiden Behandlungsgruppen sowie für den Endpunkt Sehschärfe für die kontinuierliche, gemeinsame Auswertung der ETDRS- und HOTV-Sehtafeln. Die Responderanalysen für eine Verbesserung der Sehschärfe um  $\geq 10$  bzw.  $\geq 15$  Buchstaben gemessen mit der ETDRS-Tafel zeigten eine numerische Verbesserung in der Interventionsgruppe, es wurden keine Effektschätzer und p-Werte vorgelegt.

Für alle Endpunkte wird das Verzerrungspotential u. a. aufgrund des offenen Studiendesigns als hoch angesehen.

Für die LTFU-Studie konnte für die Endpunkte MLMT, FST, Sehschärfe und Perimetrie bei den deskriptiven, nicht vergleichenden Auswertungen bis Jahr 5 in beiden ursprünglichen Behandlungsgruppen eine Verbesserung der Werte im Vergleich zu Baseline beobachtet werden.

Aufgrund des einarmigen, offenen Studiendesigns wird das Verzerrungspotential als hoch angesehen. Es bleibt unklar, inwieweit eine Anpassung der Testdurchführung an die individuelle Situation der Teilnehmenden vorgenommen wurde.



## 4.5 Lebensqualität

Es wurden keine geeigneten Daten für die Kategorie Lebensqualität eingereicht.

## 4.6 Sicherheit

### Studie 301

Es werden Sicherheitsdaten für die Studie 301 bis Jahr 1B für die Interventionsgruppe (1 Jahr nach Erhalt der ersten Injektion) / 1C für die Kontrollgruppe (1 Jahr nach Baseline) dargestellt.

Es wurden für das aktuelle Dossier keine Beobachtungszeiten für beide Behandlungsgruppen eingereicht. Im vorherigen Verfahren wurde für die Interventionsgruppe eine durchschnittliche Beobachtungszeit für den Zeitraum der ersten Injektion ins erste Auge (Tag 0A) bis Jahr 1B von 406,6 Tagen (SD: 20,4) angegeben. In der Kontrollgruppe lag der durchschnittliche Zeitraum zwischen Baseline und Visite zu Jahr 1C bei 354,8 Tagen (SD: 20,6) [4]. Der Unterschied zwischen beiden Gruppen betrug somit ca. 50 Tage. In der vorliegenden Konstellation wird davon ausgegangen, dass die vom pU im Dossier Modul 4 post hoc berechneten Effektschätzer für die Sicherheitsendpunkte anwendbar sind. Es wurden vom pU für diese Ereignisse, die ausschließlich im Interventionsarm auftraten, keine Effektschätzer vorgelegt.

Es werden in der Interventionsgruppe – abweichend von der Kontrollgruppe und dem üblichen Vorgehen einer zeitgleichen Erfassung – UE nicht ab Baseline, sondern erst ab der ersten Injektion berichtet. Hierdurch werden UE in der Interventionsgruppe zwischen Baseline und Injektion ins erste Auge nicht berichtet. Die mediane Beobachtungsdauer zwischen Baseline und Injektion ins erste Auge beträgt in der Interventionsgruppe 22 Tage (min; max: 6; 78). Im Dossier Modul 4 wird angegeben, dass UE ab Screening bis Studienende erfasst wurden, wenn auch in den Studienunterlagen eine Erhebung erst ab Baseline geplant war. Die Auswertung der UE erfolgte für die Interventionsgruppe allerdings erst ab der Injektion ins erste Auge.

Im Studienprotokoll wird a priori eine studienindividuelle Schweregradeinteilung für UE beschrieben. Eine Begründung für die Auswahl des Klassifizierungssystems konnte nicht identifiziert werden. Neben dieser studienindividuellen Schweregradeinteilung konnte, sofern anwendbar, auch die WHO-Toxizitätstabelle eingesetzt werden. Im Modul 4 werden post hoc die Ergebnisse für UE Grad  $\geq 3$  neben der a priori beschriebenen studienindividuellen Schweregradeinteilung dargestellt. Es bleibt unklar, nach welchem Klassifizierungssystem die Einteilung UE Grad  $\geq 3$  erfolgte und inwieweit diese mit der beschriebenen Schweregradeinteilung im Studienprotokoll übereinstimmen. Aus diesem Grund werden die UE Grad  $\geq 3$  nach SOC und PT im Anhang der Nutzenbewertung dargestellt.

Es konnten bei den Gesamtraten keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen beobachtet werden. UE, die zum Studienabbruch führten, sind nicht aufgetreten. Ein statistisch signifikanter Unterschied zuungunsten der Intervention konnte bei dem UE SOC „Störungen des Blutes und lymphatischen Systems“ und PT „Leukozytose“ (jeweils p-Wert: 0,027) beobachtet werden. Es wurden vom pU für Ereignisse, die ausschließlich im Interventionsarm auftraten, keine Effektschätzer vorgelegt. Insgesamt war numerisch der Anteil an UE im Interventionsarm höher als im Kontrollarm. SUE und schwere UE wurden bei einzelnen Personen nur im Interventionsarm, jedoch nicht im Kontrollarm berichtet. UE von besonderem Interesse wurden in der Studie nicht definiert und erhoben.

Bei UE auf SOC- und PT-Ebene unterscheiden sich die Angaben der Anzahl an Personen mit UE zwischen finalem Studienbericht und Dossier Modul 4. Für diese Diskrepanzen wurden keine Begründungen identifiziert.

Das Verzerrungspotential wird aufgrund des offenen Studiendesigns als hoch angesehen.

### Studie LTFU

Gemäß SAP V1 (22. Juni 2016) sollten bis Jahr 1B alle UE vollständig überwacht werden. Im Anschluss sollte der Fokus der Dokumentation im CRF nur auf bestimmten UE (onkologische Ereignisse, hämatologische Ereignisse, neurologische Ereignisse und/oder Autoimmunerkrankungen) liegen. Alle weiteren UE wurden in den Patientenakten, jedoch nicht im CRF dokumentiert. Ab Studienprotokoll V2 (15. Juni 2018) ist der Passus zur vollständigen Überwachung der UE bis Jahr 1B nicht mehr enthalten. Es bleibt somit unklar, für welchen Zeitraum ggf. eine vollständige Darstellung der UE im Studienbericht bzw. im Dossier Modul 4 vorgenommen wurde oder ob nur das Berichten bestimmter UE vorgesehen war.

Für die ursprüngliche Interventionsgruppe geht in die Darstellung „Erste Injektion bis Jahr 1B“ bzw. „Erste Injektion bis Jahr 5“ der Beobachtungszeitraum aus der Studie 301 (ab Injektion ins erste Auge) ein. Sofern bereits ab Studienbeginn der Studie LTFU nur bestimmte UE dokumentiert wurden, erfolgte, anders als in Studie 301, keine einheitliche Erhebung ab Applikation ins erste Auge.

Im Studienprotokoll wird eine Schweregradeinteilung nur in mild, moderat und schwer beschrieben, während im Dossier Modul 4 die Schweregradeinteilung zusätzlich um lebensbedrohlich und Tod, ähnlich der Studie 301, ergänzt wurde. UE auf SOC- und PT-Ebene Grad  $\geq 3$  werden ergänzend im Anhang dargestellt, da wie in Studie 301 unklar ist, nach welchem Klassifikationssystem die Schweregradeinteilung erfolgte.

In der Nutzenbewertung werden die im Dossier Modul 4 dargestellten Analysen zum Zeitpunkt der Injektion ins erste Auge bis Jahr 1B bzw. Jahr 5 nach Injektion ins zweite Auge dargestellt, da hier für beide Interventionsgruppen „vergleichbare“ Beobachtungszeiträume zugrunde gelegt werden.

Es konnten keine Angaben zu den (medianen) Beobachtungszeiten bis Jahr 5 identifiziert werden. Für die ursprüngliche Interventionsgruppe betrug die mittlere Beobachtungszeit 406,6 Tage (SD: 20,4) zwischen dem Zeitpunkt der ersten Injektion und Jahr 1 nach zweiter Injektion.

Bei den Gesamtraten traten numerisch in der ursprünglichen Interventionsgruppe sowohl für den Zeitraum Injektion ins erste Auge bis Jahr 1B als auch bis Jahr 5 mehr schwere UE und SUE auf als in der ursprünglichen Kontrollgruppe. Es wurden keine UE, die zum Studienabbruch führten, beobachtet.

Vereinzelt finden sich bei UE und schweren UE auf SOC- und PT-Ebene unterschiedliche Angaben zur Anzahl und Häufigkeit von Teilnehmenden mit UE zwischen Interimsstudienbericht und Dossier für die Auswertung „Injektion ins erste Auge bis Jahr 1B“. Für diese Diskrepanzen wurden keine Begründungen identifiziert. Ergebnisse zu Jahr 5 nach Injektion konnten nur im Dossier Modul 4 identifiziert werden.

Gemäß Interimsstudienbericht 2020 wurden für die LTFU-Studie 2 UE (Parästhesie und Krämpfe), die in den vorherigen Interimsstudienberichten nicht als UE von besonderem Interesse klassifiziert waren, als solche im neuen Studienbericht gekennzeichnet.

Das Verzerrungspotential wird aufgrund des offenen einarmigen Studiendesigns als hoch angesehen.

### **Zusammenfassende Einschätzung zur Sicherheit**

In der Studie 301 zeigte sich für SUE und schwere UE bei hohem Verzerrungspotential kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen. Nur im Interventionsarm der Studie 301 wurden vereinzelt SUE und schwere UE beobachtet. Ein statistisch signifikanter Unterschied zuungunsten der Intervention zeigte sich im PT „Leukozytose“ und dadurch in der SOC „Störungen des Blutes und lymphatischen Systems“.

In der Studie LTFU bleibt unklar, für welchen Zeitraum eine vollständige Darstellung der UE im Studienbericht bzw. im Dossier Modul 4 vorgenommen wurde bzw. ab wann nur bestimmte UE berichtet wurden. Auf Basis der vorhandenen Angaben und bei hohem Verzerrungspotential sind nach Jahr 1 nur noch bei einzelnen Personen neue UE aufgetreten.

Insgesamt ist die Aussagekraft der Sicherheitsauswertungen durch die geringe Fallzahl limitiert.

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

Die Angaben der Fachinformation sind zu berücksichtigen.

## 6 Zusammenfassung der Nutzenbewertung

Voretigen Neparvovec ist zugelassen für die Behandlung von erwachsenen und pädiatrischen Patientinnen und Patienten mit Sehverlust aufgrund einer erblichen Netzhautdystrophie, die auf nachgewiesenen biallelischen RPE65-Mutationen beruht, und die über ausreichend lebensfähige Netzhautzellen verfügen. Die Nutzenbewertung von Voretigen Neparvovec basiert auf der zulassungsbegründenden Studie 301 und der Nachbeobachtungsstudie LTFU. Es handelt sich bei der Studie 301 um eine randomisierte, kontrollierte, offene Phase-III-Studie im Parallelgruppenvergleich, bei der Voretigen Neparvovec als gentherapeutische Intervention in Form einer subretinalen Injektion mit beobachtendem Abwarten in der Kontrollgruppe verglichen wurde. Ein Wechsel von Personen der Kontrollgruppe in den Interventionsarm war nach mindestens einem Jahr möglich. Die Beobachtungsdauer nach Injektion der Prüfmedikation ins zweite Auge betrug in beiden Gruppen 1–1,5 Jahre, wobei in der Kontrollgruppe die Möglichkeit der Behandlung mit dem Prüfpräparat erst nach Visite zu Jahr 1 bestand. Die Patientinnen und Patienten konnten im Anschluss in der LTFU-Studie bis zu 15 Jahre weiterbeobachtet werden.

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

Weitergehende, über die Angaben in der nachfolgenden zusammenfassenden Tabelle hinausgehende Erläuterungen zur Methodik und den Ergebnissen der Endpunkte finden sich in den jeweiligen Kapiteln 2.3 bis 2.5 und 3.2 bis 3.5 der Nutzenbewertung.

*Tabelle 33: Zusammenfassende Darstellung der Ergebnisse der Studie 301 (ITT-Population)*

Studie 301 Darstellung der Ergebnisse	Voretigen Neparvovec N = 21			Beobachtendes Abwarten N = 10			Voretigen Neparvovec vs. Beobachten- des Abwarten	Effekt
	N <sup>1)</sup>	Personen mit Ereignis, n (%)	N <sup>1)</sup>	Personen mit Ereignis, n (%)	RR [95%-KI] <sup>2)</sup> ; p-Wert <sup>2)</sup>			
<b>Mortalität</b>								
Todesfälle <sup>3)</sup>	20	0 (0,0)	9	0 (0,0)	-	-		
<b>Morbidität</b>								
	N Base- line/ N Jahr 1	Baseline MW (SD)	Ände- rung zu Jahr 1B MW (SD)	N Base- line/ N Jahr 1	Baseline MW (SD)	Ände- rung zu Jahr 1C MW (SD)	Mittlere Differenz [95%-KI] <sup>4)5)</sup> ; p-Wert <sup>5)6)</sup>	
MLMT <sup>7)8)</sup>	21/21	3,1 (1,7)	1,8 (1,1)	10/10	2,9 (1,6)	0,2 (1,0)	1,6 [0,7; 2,4] 0,001	$\uparrow$ <sup>9)</sup>
Goldmann III4e <sup>10)</sup>	20/19	332,9 (413,3)	302,1 (289,6)	10/9	427,1 (372,0)	-76,7 (258,7)	378,7 [145,5; 612,0] <sup>11)</sup> 0,006 <sup>11)</sup>	$\uparrow$ <sup>12)</sup>

Studie 301 Darstellung der Ergebnisse	Voretigen Neparvovec N = 21			Beobachtendes Abwarten N = 10			Voretigen Neparvovec vs. Beobachten- des Abwarten	Effekt
	N	Baseline MW (SE)	Ände- rung zu Jahr 1B MW (SE)	N	Baseline MW (SE)	Ände- rung zu Jahr 1C MW (SE)		
Humphrey: Fovea- Sensitivität <sup>13)</sup>	20/19	22,4 (6,8)	2,4 (9,7)	10/9	17,6 (8,9)	2,3 (5,3)	0,04 [-7,1; 7,2] <sup>11)</sup> 0,18 <sup>11)</sup>	↔
Humphrey: Makula- Grenzwert <sup>13)</sup>	20/19	16,1 (5,5)	7,7 (6,2)	10/9	14,4 (8,0)	-0,2 (1,7)	7, [3,5; 12,2] <sup>11)</sup> < 0,001 <sup>11)</sup>	↑ <sup>14)</sup>
	<i>N Base- line/ N Jahr 1</i>	<i>Baseline MW (SE)</i>	<i>Ände- rung zu Jahr 1B MW (SE)</i>	<i>N Base- line/ N Jahr 1</i>	<i>Baseline MW (SE)</i>	<i>Ände- rung zu Jahr 1C MW (SE)</i>	<i>Mittlere Änderung [95%-KI]<sup>4)15)</sup>; p-Wert<sup>4)15)</sup></i>	
Lichtempfind- lichkeit <sup>16)</sup> Weißes Licht	20/19	-1,3 (0,1)	-2,1 (0,3)	9/9	1,7 (0,1)	0,04 (0,4)	-2,1 [-3,2; -1,0] < 0,001	↑ <sup>17)</sup>
Lichtempfind- lichkeit <sup>16)</sup> Blaues Licht	20/17	-1,6 (0,1)	-2,0 (0,3)	9/9	-2,0 (0,2)	0,1 (0,5)	-2,1 [-3,3; -0,9] 0,001	↑
Lichtempfind- lichkeit <sup>16)</sup> Rotes Licht	20/17	-1,2 (0,1)	-1,3 (0,2)	9/9	-1,7 (0,2)	0,2 (0,2)	-1,5 [-2,1; -0,9] < 0,001	↑
Sehschärfe (ETDRS, HOTV) <sup>18)</sup>	21/20	1,2 (0,1)	-0,2 (0,1)	10/9	1,3 (0,2)	0,01 (0,1)	-0,2 [-0,4; 0,1] 0,170 <sup>19)</sup>	↔
	<i>N</i>	<i>Personen mit Ereignis, n (%)</i>		<i>N</i>	<i>Personen mit Ereignis, n (%)</i>		<i>RR [95%-KI]<sup>2)</sup>; p-Wert<sup>2)</sup></i>	
Sehschärfe (ETDRS) ≥ 10 Buchstaben	21	6 (28,6) <sup>20)</sup>		10	0 (0,0)		k. A. 0,071	n. b.
Sehschärfe (ETDRS) ≥ 15 Buchstaben	21	4 (19,0) <sup>21)</sup>		10	0 (0,0)		k. A. 0,268	n. b.
<b>Lebensqualität</b>								
	<i>N</i>	<i>Personen mit Ereignis, n (%)</i>		<i>N</i>	<i>Personen mit Ereignis, n (%)</i>		<i>RR [95%-KI]<sup>2)</sup>; p-Wert<sup>2)</sup></i>	
_22)								n. b.
<b>Sicherheit<sup>1)23)</sup></b>								
	<i>N</i>	<i>Personen mit Ereignis, n (%)</i>		<i>N</i>	<i>Personen mit Ereignis, n (%)</i>		<i>RR [95%-KI]<sup>2)</sup>; p-Wert<sup>2)</sup></i>	
Schwere UE <sup>24)</sup>	20	3 (15,0)		9	0 (0,0)		k. A.; 0,532	↔
SUE	20	2 (10,0)		9	0 (0,0)		k. A.; 1,000	↔

Studie 301 Darstellung der Ergebnisse	Voretigen Neparvovec N = 21		Beobachtendes Abwarten N = 10		Voretigen Neparvovec vs. Beobachten- des Abwarten	Effekt
UE, die zum Studienab- bruch führten	20	0 (0,0)	9	0 (0,0)	n. b.	n. b.

- <sup>1)</sup> Sicherheitspopulation
- <sup>2)</sup> Für die Berechnung wurde das relative Risiko mittels Vier-Felder-Tafel berechnet, das KI wurde mittels exakter Methode berechnet. Der p-Wert wurde nach der exakten Methode nach Fischer berechnet.
- <sup>3)</sup> Todesfälle wurden über die Sicherheit erhoben.
- <sup>4)</sup> Es wurde die Differenz der beobachteten mittleren Veränderungen zu Baseline berechnet. Für die Berechnung des 95%-KI wird seitens des pU ein gemischtes Modell verwendet, welches Terme für die Behandlung und die Studienvisite einschließt.
- <sup>5)</sup> Es war keine stratifizierte Analyse gemäß Randomisierungsfaktoren aufgrund der geringen Fallzahl vorgesehen.
- <sup>6)</sup> Wilcoxon-Rangsummen-Test mit exaktem zweiseitigem p-Wert.
- <sup>7)</sup> Ein höherer Wert im Mobilitätsscore bedeutet eine Verbesserung im MLMT.
- <sup>8)</sup> Primärer Endpunkt
- <sup>9)</sup> Es wurde post hoc für das Dossier Modul 4 eine SMD nach Hedges und Olkin berechnet. Der Wert betrug 1,50 [95%-KI: 0,66; 2,34]. Auf Basis von Hedges' g (einem möglichen Hilfskriterium zur Einschätzung klinischer Relevanz) erscheint das Vorliegen eines klinisch relevanten Effekts plausibel.
- <sup>10)</sup> Summenscore
- <sup>11)</sup> Effektschätzer und p-Wert wurden post hoc berechnet. Imputationen waren gemäß SAP nicht vorgesehen.
- <sup>12)</sup> Es wurde post hoc für das Dossier Modul 4 eine SMD nach Hedges und Olkin berechnet. Der Wert betrug 1,27 [95%-KI: 0,41; 2,12]. Auf Basis von Hedges' g (einem möglichen Hilfskriterium zur Einschätzung klinischer Relevanz) erscheint das Vorliegen eines klinisch relevanten Effekts plausibel.
- <sup>13)</sup> Angaben in dB
- <sup>14)</sup> Es wurde post hoc für das Dossier Modul 4 eine SMD nach Hedges und Olkin berechnet. Der Wert betrug 1,45 [95%-KI: 0,61; 2,29]. Auf Basis von Hedges' g (einem möglichen Hilfskriterium zur Einschätzung klinischer Relevanz) erscheint das Vorliegen eines klinisch relevanten Effekts plausibel.
- <sup>15)</sup> Die mittlere Änderung und der p-Wert wurden mit einem MMRM mit den Termen Behandlung, Studienvisite und Behandlung\*Studienvisite gerechnet. Imputationen wurden für die primäre Analyse nicht vorgenommen.
- <sup>16)</sup> Gemessen mit dem Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test (FST) ( $\log_{10}$  (cd s/m<sup>2</sup>)). Je negativer der Wert, desto besser die Lichtempfindlichkeit.
- <sup>17)</sup> Es wurde post hoc für das Dossier Modul 4 eine SMD nach Hedges und Olkin berechnet. Der Wert betrug -1,52 [95%-KI: -2,41; -0,63]. Auf Basis von Hedges' g (einem möglichen Hilfskriterium zur Einschätzung klinischer Relevanz) erscheint das Vorliegen eines klinisch relevanten Effekts plausibel.
- <sup>18)</sup> Es werden a priori die ETDRS- und HOTV-Sehtafeln zusammen ausgewertet. Angaben in logMAR.
- <sup>19)</sup> Im Dossier wird ein p-Wert von 0,175 angegeben.
- <sup>20)</sup> Prozentualer Anteil selbst ausgerechnet bezogen auf die ITT-Population. Im Dossier Modul 4 wird ein Anteil von 33,3 % Personen mit einer Verbesserung um  $\geq 10$  Buchstaben angegeben.
- <sup>21)</sup> Prozentualer Anteil selbst ausgerechnet bezogen auf die ITT-Population. Im Dossier Modul 4 wird ein Anteil von 22,2 % Personen mit einer Verbesserung um  $\geq 15$  Buchstaben angegeben.
- <sup>22)</sup> Es wurden keine geeigneten bewertungsrelevanten Daten eingereicht.
- <sup>23)</sup> Die UE von besonderem Interesse wurden im Ergebnisteil der Sicherheit dargestellt.
- <sup>24)</sup> Es wurde gemäß Studienprotokoll eine studienindividuelle Schweregradeinteilung durch den Prüfarzt bzw. die Prüfarztin vorgenommen.

Abkürzungen: cds: Candela-Sekunde; ETDRS: Early Treatment of Diabetic Retinopathy Study; FST: Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test; ITT: Intention to Treat; k. A.: keine Angaben; KI: Konfidenzintervall; logMAR: Logarithmus des minimalen Auflösungswinkels; MLMT: Multi-Luminanz-Mobilitätstest; MW: Mittelwert; n. b.: nicht bestimmbar; RR: Relatives Risiko; SAP: Statistischer Analyseplan; SD: Standardabweichung; SE: Standardfehler; SMD: Standardisierte Mittelwertdifferenz; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis

**Tabelle 34: Zusammenfassende Darstellung der Ergebnisse der LTFU-Studie für die ehemalige Studie 301 (MITT-Population) bis Datenschnitt vom 30. Juni 2020**

Studie LTFU (301) Darstellung der Ergebnisse	Ursprüngliche Interventionsgruppe N = 20			Ursprüngliche Kontrollgruppe N = 9		
	N <sup>1)</sup>	Personen mit Ereignis, n (%)		N <sup>1)</sup>	Personen mit Ereignis, n (%)	
<b>Mortalität</b>						
Todesfälle <sup>2)</sup>	20	0 (0,0)		9	0 (0,0)	
<b>Morbidität</b>						
	N <sup>1)</sup>	MW (SD)	Differenz MW (SD)	N <sup>1)</sup>	MW (SD)	Differenz MW (SD)
MLMT <sup>3)</sup>						
Baseline	20	3,3 (1,4)		9	3,6 (1,4)	
Jahr 5 <sup>4)</sup>	18	4,9 (1,8)	1,6 (1,1)	7	6,0 (0,0)	2,4 (1,6)
Perimetrie nach Goldmann (III4e) <sup>5)</sup>						
Baseline	19	350,4 (416,9)		9	397,8 (367,3)	
Jahr 5	16	533,4 (428,5)	166,6 (208,7)	7	506,9 (219,8)	188,9 (222,3)
Perimetrie nach Humphrey (Fovea) <sup>6)</sup>						
Baseline	19	23,3 (5,5)		9	21,5 (8,9)	
Jahr 5	17	25,7 (5,3)	1,2 (8,1)	7	25,6 (9,4)	4,7 (6,9)
Perimetrie nach Humphrey (Macula) <sup>6)</sup>						
Baseline	19	16,6 (5,3)		9	15,8 (7,4)	
Jahr 5	18	21,7 (7,1)	4,8 (6,7)	7	23,3 (5,5)	8,2 (5,2)
Lichtempfindlichkeit, Weißes Licht <sup>7)</sup>						
Baseline	19	-1,3 (0,4)		9	-1,6 (0,5)	
Jahr 5	18	-3,2 (1,3)	-2,0 (1,5)	7	-4,1 (1,3)	-2,6 (1,2)
Lichtempfindlichkeit, Blaues Licht <sup>7)</sup>						
Baseline	19	-1,7 (1,6)		9	-1,9 (0,4)	
Jahr 5	18	-3,7 (1,4)	-2,1 (1,6)	7	-4,6 (1,5)	-2,8 (1,5)
Lichtempfindlichkeit, Rotes Licht <sup>7)</sup>						
Baseline	19	-1,3 (0,4)		9	-1,5 (0,5)	
Jahr 5	18	-2,5 (0,8)	-1,3 (0,8)	7	-3,0 (0,8)	-1,6 (0,8)
Sehschärfe (ETDRS, HOTV)						
Baseline	20	1,1 (0,4)		9	1,0 (0,3)	
Jahr 5	18	1,1 (1,8)	0,0 (0,6)	7	0,8 (0,3)	-0,1 (0,3)



Studie LTFU (301) Darstellung der Ergebnisse	Ursprüngliche Interventionsgruppe N = 20		Ursprüngliche Kontrollgruppe N = 9	
<b>Sicherheit<sup>8)9)</sup></b>				
	<i>N<sup>1)</sup></i>	<i>Personen mit Ereignis, n (%)</i>	<i>N<sup>1)</sup></i>	<i>Personen mit Ereignis, n (%)</i>
Erhalt Injektion ins erste Auge bis Jahr 1B <sup>10)</sup>				
Schwere UE <sup>11)</sup>	20	3 (15,0)	9	0 (0,0)
SUE	20	2 (10,0)	9	1 (11,1)
UE, das zum Abbruch der Studie führte	20	0 (0,0)	9	0 (0,0)
Erhalt Injektion ins erste Auge bis Jahr 5 <sup>12)</sup>				
Schwere UE <sup>11)</sup>	20	4 (20,0)	9	0 (0,0)
SUE	20	4 (20,0)	9	1 (11,1)
UE, das zum Abbruch der Studie führte	20	0 (0,0)	9	0 (0,0)

<sup>1)</sup> Entsprach der ehemaligen mITT-Population der Studie 301, bei der 20 Teilnehmende der ursprünglichen Interventionsgruppe und 9 Teilnehmende der ursprünglichen Kontrollgruppe in die LTFU-Studie eingeschlossen wurden. Die mITT-Population entsprach der Sicherheitspopulation.

<sup>2)</sup> Todesfälle wurden über die Sicherheit erfasst.

<sup>3)</sup> Ein höherer Wert im Mobilitätsscore bedeutet eine Verbesserung im MLMT.

<sup>4)</sup> Die Werte zu den Visiten wurden ausschließlich im Dossier Modul 4 berichtet.

<sup>5)</sup> Summenscore

<sup>6)</sup> Angaben in db

<sup>7)</sup> Gemessen mit dem Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test (FST) [ $\log_{10}(\text{cd s/m}^2)$ ]

<sup>8)</sup> Gemäß SAP (22. Juni 2016) war in der LTFU-Studie eine vollständige Erhebung der UE nur bis Jahr 1B, jedoch nicht im Rahmen einer Langzeiterhebung für den gesamten Beobachtungszeitraum vorgesehen, ab Jahr 1B sollten im CRF nur bestimmte UE, nämlich SUE, UE, die möglicherweise oder wahrscheinlich in Verbindung mit der Administration des Prüfpräparats stehen, und neue oder sich verschlechternde UE in einer der 4 Kategorien (onkologische, hämatologische, neurologische Ereignisse und/oder Autoimmunerkrankungen) erfasst werden. Ab Studienprotokoll Version 2 (15. Juni 2018) wird angegeben, dass die Dokumentation von UE sich auf bestimmte UE fokussiert. Es bleibt unklar, inwieweit bzw. ab wann keine vollständige Erhebung der UE in der LTFU-Studie stattfand.

<sup>9)</sup> Die UE von besonderem Interesse wurden im Ergebnisteil der Sicherheit beschrieben.

<sup>10)</sup> Für die ursprüngliche Interventionsgruppe entspricht der Zeitraum „Erhalt der Injektion ins erste Auge bis Jahr 1B“ dem Beobachtungszeitraum der Studie 301. Für die ursprüngliche Kontrollgruppe werden UE ab Beginn der LTFU-Studie (ab Injektion ins erste Auge) erfasst.

<sup>11)</sup> Studienindividuelle Einteilung der Schweregrade: Wenn möglich, sollte gemäß dem Studienprotokoll eine Schweregradeinteilung (Grad 1–4) in Orientierung der WHO-Toxizitätsskala erfolgen.

<sup>12)</sup> In Modul 4 post hoc berechnet: Jahr 5 entspricht dem Zeitpunkt der letzten Injektion + 1.825 Tage.

Abkürzungen: CRF: Case report form; ETDRS: Early Treatment of Diabetic Retinopathy Study; KI: Konfidenzintervall; LTFU: Long-Term Follow-up; mITT: Modified Intention-to-Treat; MLMT: Multi-Luminanz-Mobilitätstest; MW: Mittelwert; SAP: Statistical Analysis Plan; SD: Standardabweichung; SE: Standardfehler; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis

## Referenzen

1. **Chung DC, McCague S, Yu ZF, Thill S, DiStefano-Pappas J, Bennett J, et al.** Novel mobility test to assess functional vision in patients with inherited retinal dystrophies. *Clinical & experimental ophthalmology* 2018;46(3):247-259.
2. **European Medicines Agency (EMA).** Luxturna (Voretigene neparvovec): European public assessment report EMEA/H/C/004451/0000 [online]. 20.09.2018. Amsterdam (NED): EMA. [Zugriff: 13.04.2022]. URL: [https://www.ema.europa.eu/en/documents/assessment-report/luxturna-epar-public-assessment-report\\_en.pdf](https://www.ema.europa.eu/en/documents/assessment-report/luxturna-epar-public-assessment-report_en.pdf).
3. **European Medicines Agency (EMA).** Luxturna (Voretigene neparvovec): Summary of product characteristics (EPAR Product information) [online]. Amsterdam (NED): EMA; 2021. [Zugriff: 13.04.2022]. URL: [https://www.ema.europa.eu/en/documents/product-information/luxturna-epar-product-information\\_en.pdf](https://www.ema.europa.eu/en/documents/product-information/luxturna-epar-product-information_en.pdf).
4. **Gemeinsamer Bundesausschuss (G-BA).** Amendment Teil 1 und Teil 2 zur Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V; Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO: Wirkstoff Voretigen Neparvovec [online]. Berlin (GER): G-BA; 2019. [Zugriff: 13.04.2022]. URL: [https://www.g-ba.de/downloads/92-975-3163/2019-10-17\\_AM-RL-XII\\_Voretigen-Neparvovec\\_D-436\\_Amendment.pdf](https://www.g-ba.de/downloads/92-975-3163/2019-10-17_AM-RL-XII_Voretigen-Neparvovec_D-436_Amendment.pdf).
5. **Gemeinsamer Bundesausschuss (G-BA).** Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Voretigen Neparvovec (Erbliche Netzhautdystrophie), vom 20. Mai 2021 [online]. Berlin (GER): G-BA; 2021. [Zugriff: 13.04.2022]. URL: [https://www.g-ba.de/downloads/39-261-4834/2021-05-20\\_AM-RL-XII\\_Voretigen-Neparvovec\\_D-436\\_BAnz.pdf](https://www.g-ba.de/downloads/39-261-4834/2021-05-20_AM-RL-XII_Voretigen-Neparvovec_D-436_BAnz.pdf).
6. **Gemeinsamer Bundesausschuss (G-BA).** Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Voretigen Neparvovec, vom 17. Oktober 2019 [online]. Berlin (GER): G-BA; 2019. [Zugriff: 13.04.2022]. URL: [https://www.g-ba.de/downloads/39-261-3984/2019-10-17\\_AM-RL-XII\\_Voretigen-Neparvovec\\_D-436\\_aktualisiert-2020-12-01.pdf](https://www.g-ba.de/downloads/39-261-3984/2019-10-17_AM-RL-XII_Voretigen-Neparvovec_D-436_aktualisiert-2020-12-01.pdf).
7. **Gemeinsamer Bundesausschuss (G-BA).** Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V; Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO: Wirkstoff Voretigen Neparvovec [online]. Berlin (GER): G-BA; 2019. [Zugriff: 13.04.2022]. URL: [https://www.g-ba.de/downloads/92-975-3039/2019-04-15\\_Nutzenbewertung-G-BA-Voretigen-Neparvovec-D-436.pdf](https://www.g-ba.de/downloads/92-975-3039/2019-04-15_Nutzenbewertung-G-BA-Voretigen-Neparvovec-D-436.pdf).
8. **Gemeinsamer Bundesausschuss (G-BA).** Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Voretigen Neparvovec (Erbliche Netzhautdystrophie), vom 20. Mai 2021 [online]. Berlin (GER): G-BA; 2021. [Zugriff: 13.04.2022]. URL: [https://www.g-ba.de/downloads/40-268-7540/2021-05-20\\_AM-RL-XII\\_Voretigen-Neparvovec\\_D-436\\_TrG.pdf](https://www.g-ba.de/downloads/40-268-7540/2021-05-20_AM-RL-XII_Voretigen-Neparvovec_D-436_TrG.pdf).

9. **Gemeinsamer Bundesausschuss (G-BA).** Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V - Voretigen Neparvovec, vom 17. Oktober 2019 [online]. Berlin (GER): G-BA; 2019. [Zugriff: 13.04.2022]. URL: [https://www.g-ba.de/downloads/40-268-6053/2019-10-17\\_AM-RL-XII\\_Voretigen-Neparvovec\\_D-436\\_TrG\\_akt.pdf](https://www.g-ba.de/downloads/40-268-6053/2019-10-17_AM-RL-XII_Voretigen-Neparvovec_D-436_TrG_akt.pdf).
10. **Novartis Pharma.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4 A; Voretigen Neparvovec (Luxturna) Behandlung von erwachsenen und pädiatrischen Patienten mit Sehverlust aufgrund einer erblichen Netzhautdystrophie, die auf nachgewiesenen biallelischen RPE65-Mutationen beruht, und die über ausreichend lebensfähige Netzhautzellen verfügen: Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 30.03.2022.
11. **Novartis Pharma.** Luxturna  $5 \times 10^{12}$  Vektorgenome/ml Konzentrat und Lösungsmittel zur Herstellung einer Injektionslösung [Fachinformation] [online]. 11.2021. [Zugriff: 04.04.2022]. URL: <https://www.fachinfo.de/>.
12. **Novartis Pharma.** A post-authorization, multicenter, multinational, longitudinal observational safety registry study for patients treated with Voretigene Neparvovec (CLTW888A12401): 2nd study interim report [unveröffentlicht]. 2022.
13. **Novartis Pharma.** A post-authorization, multicenter, multinational, longitudinal observational safety registry study for patients treated with Voretigene Neparvovec (CLTW888A12401): statistical analysis plan, version 2.0 [unveröffentlicht]. 2020.
14. **Novartis Pharma.** A post-authorization, multicenter, multinational, longitudinal observational safety registry study for patients treated with Voretigene Neparvovec (CLTW888A12401): study protocol amendment Switzerland, version v01-CH.01 [unveröffentlicht]. 2020.
15. **Patel DE, Cumberland PM, Walters BC, Russell-Eggitt I, Rahi JS, Optic study group.** Study of Optimal Perimetric Testing in Children (OPTIC): feasibility, reliability and repeatability of perimetry in children. PLoS One 2015;10(6):e0130895.
16. **Spark Therapeutics.** AAV2-hRPE65v2-301 - study procedures: clinical study report [unveröffentlicht]. 2016.
17. **Spark Therapeutics.** AAV2-hRPE65v2-301: protocol and protocol amendments [unveröffentlicht]. 2016.
18. **Spark Therapeutics.** A long-term follow-up study in subjects who received an adenovirus-associated viral vector serotype 2 containing the human RPE65 gene (AAV2-hRPE65v2) administered via subretinal injection (AAV2-hRPE65v2-LTFU-01): study protocol and amendments 1 and 2 [unveröffentlicht]. 2018.
19. **Spark Therapeutics.** A long-term follow-up study in subjects who received an adenovirus-associated viral vector serotype 2 containing the human RPE65 gene (AAV2-hRPE65v2) administered via subretinal injection (AAV2-hRPE65v2-LTFU-01): study protocol version 2 [unveröffentlicht]. 2016.
20. **Spark Therapeutics.** Mobility Test Validation Study: single-center observational study comparing the performance of normal-sighted and visually-impaired subjects on standardized obstacle courses [unveröffentlicht]. 2015.

21. **Spark Therapeutics.** A safety and efficacy study in subjects with leber congenital amaurosis (LCA) using adeno-associated viral vector to deliver the gene for human RPE65 to the retinal pigment epithelium (RPE) - AAV2-hRPE65v2-301: clinical study report addendum 2020 [unveröffentlicht]. 2021.
22. **Spark Therapeutics.** A safety and efficacy study in subjects with leber congenital amaurosis (LCA) using adeno-associated viral vector to deliver the gene for human RPE65 to the retinal pigment epithelium (RPE) (AAV2-hRPE65v2-301): clinical study report, addendum, amendment [unveröffentlicht]. 2017.
23. **Spark Therapeutics.** A Safety and Efficacy Study in Subjects with Leber Congenital Amaurosis (LCA) Using Adeno-Associated Viral Vector to Deliver the Gene for Human RPE65 to the Retinal Pigment Epithelium (RPE) [AAV2-hRPE65v2-301]; statistical analysis plan; Version 4.0 [unveröffentlicht]. 19.08.2015.
24. **Spark Therapeutics.** A safety and efficacy study in subjects with leber congenital amaurosis (LCA) using adeno-associated viral vector to deliver the gene for human RPE65 to the retinal pigment epithelium (RPE): clinical study report [unveröffentlicht]. 2017.
25. **Spark Therapeutics.** A safety and efficacy study in subjects with leber congenital amaurosis (LCA) using adeno-associated viral vector to deliver the gene for human RPE65 to the retinal pigment epithelium (RPE): statistical analysis plan [unveröffentlicht]. 2015.
26. **Spark Therapeutics.** SOP LCA 001 Performing Visual Fields using the Goldmann Perimeter [unveröffentlicht]. 2014.
27. **Spark Therapeutics.** SOP LCA 007 Performing a Visual Acuity Test [unveröffentlicht]. 2014.
28. **Spark Therapeutics.** SOP LCA 015 Performing Static Threshold Visual Fields using the Humphrey Visual Field Analyzer [unveröffentlicht]. 2014.
29. **Spark Therapeutics.** SOP LCA 016 Performing Full-field Light Sensitivity Threshold Testing [unveröffentlicht]. 2014.
30. **Submacular Surgery Trials Research Group.** Evaluation of minimum clinically meaningful changes in scores on the National Eye Institute Visual Function Questionnaire (NEI-VFQ) SST Report Number 19. *Ophthalmic Epidemiology* 2007;14(4):205-215.

## Anhang

### A 1 Studie 301 und LTFU

#### Ergebnisse zur Sicherheit

##### Unerwünschte Ereignisse Grad $\geq 3$

##### Studie 301

UE Grad  $\geq 3$  traten bei 11 Personen (55,0 %) im Interventionsarm und bei 2 Personen (22,2 %) im Kontrollarm auf (RR: 2,48 [95%-KI: 0,80; 24,37]; p-Wert: 0,130). Es bleibt unklar, nach welchem Klassifizierungssystem die Schweregradeinteilung erfolgte.

In Tabelle 35 sind die im Dossier dargestellten UE  $\geq$  Grad 3 auf SOC- und PT-Ebene dargestellt.

Tabelle 35: Studie 301: alle UE Grad  $\geq 3^1$  innerhalb eines Jahres nach Baseline (Sicherheitspopulation)

MedDRA-Systemorganklasse Preferred Term	Intervention: Voretigen Neparvovec <sup>2)</sup> (N = 20) n (%)	Kontrolle: Beobachtendes Abwarten <sup>2)</sup> (N = 9) n (%)	RR [95%-KI] <sup>3)</sup> p-Wert <sup>4)</sup>
<b>Erkrankungen des Blutes und des Lymphsystems</b>	9 (45,0)	0 (0,0)	k. A., 0,027
Leukozytose <sup>5)</sup>	9 (45,0)	0 (0,0)	k. A., 0,027
<b>Erkrankungen des Gastrointestinaltrakts</b>	2 (10,0)	0 (0,0)	k. A., 1,000
Übelkeit	2 (10,0)	0 (0,0)	k. A., 1,000
Erbrechen	2 (10,0)	0 (0,0)	k. A., 1,000
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	1 (5,0)	0 (0,0)	k. A., 1,000
Reaktionen auf das Arzneimittel	1 (5,0)	0 (0,0)	k. A., 1,000
<b>Untersuchungen</b>	0 (0,0)	1 (11,1)	n. a., 0,310
Gewicht erniedrigt	0 (0,0)	1 (11,1)	n. a., 0,310
<b>Stoffwechsel- und Ernährungsstörungen</b>	1 (5,0)	0 (0,0)	k. A., 1,000
Hyperkaliämie	1 (5,0)	0 (0,0)	k. A., 1,000
<b>Erkrankungen des Nervensystems</b>	1 (5,0)	0 (0,0)	k. A., 1,000
Krampfanfall	1 (5,0)	0 (0,0)	k. A., 1,000
<b>Gefäßerkrankungen</b>	0 (0,0)	1 (11,1)	n. a., 0,310
Hypertonie	0 (0,0)	1 (11,1)	n. a., 0,310

<sup>1)</sup> Eine Darstellung der Ergebnisse zu UE Grad  $\geq 3$  wird post hoc nur im Dossier vorgenommen. Es bleibt unklar, nach welchem Klassifizierungssystem die Einteilung erfolgte und inwieweit diese mit der a priori beschriebenen Schweregradeinteilung im Protokoll zusammenhängt.

<sup>2)</sup> Unterschiedlicher Beginn des Erhebungszeitraum zwischen beiden Gruppen: Die Erhebung begann in der Interventionsgruppe ab der ersten Injektion und in der Kontrollgruppe ab der Baseline-Untersuchung. Weitere Angaben zur gesamten Beobachtungs-/Studiendauer konnten nicht identifiziert werden.

<sup>3)</sup> Keine Berechnung der Effektschätzer seitens des pU.

<sup>4)</sup> P-Wert berechnet post hoc anhand des exakten Tests nach Fischer.

<sup>5)</sup> Erfassung des UE Leukozytose nur ab einem bestimmten Schweregrad im Interventionsarm, während der Gabe von Glukokortikoiden (18–30 Tage, je nach Zeitpunkt der zweiten Injektion) nur als UE bei Grad 3 oder höher.

Abkürzungen: k. A.: keine Angaben; KI: Konfidenzintervall; MedDRA: Medical Dictionary for Regulatory Activities; n. a.: nicht anwendbar; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; UE: Unerwünschte/s Ereignis/se.

*Studie LTFU*

Es traten UE Grad  $\geq 3$  bis Jahr 1B bei 11 (55,0 %) Personen der ursprünglichen Interventionsgruppe bzw. bei 2 Personen (22,2 %) der ursprünglichen Kontrollgruppe und bis Jahr 5 bei 12 Personen (60,0 %) der ursprünglichen Interventions- bzw. bei 2 Personen (22,2 %) der ursprünglichen Kontrollgruppe auf. Es bleibt unklar, nach welchem Klassifizierungssystem die Schweregradeinteilung erfolgte.

In Tabelle 36 werden die in Modul 4 post hoc berichteten UE Grad  $\geq 3$  für den Zeitraum Tag der Injektion ins erste Auge bis Jahr 1B bzw. Jahr 5B berichtet. Angaben im Studienbericht konnten nicht identifiziert werden.

UE Grad  $\geq 3$  traten bis auf 2 Teilnehmende in der ursprünglichen Kontrollgruppe in der ursprünglichen Interventionsgruppe auf. Das häufigste UE Grad  $\geq 3$  war in beiden Behandlungsgruppen SOC „Erkrankungen des Blutes und des Lymphsystems“ und „Leukozytose“ (jeweils in der ursprünglichen Interventionsgruppe vs. ursprünglichen Kontrollgruppe: 45,0 % vs. 22,2 %).

*Tabelle 36: LTFU-Studie: alle UE Grad  $\geq 3$ <sup>1)</sup> bis Jahr 1B bzw. Jahr 5 (Sicherheitspopulation)*

MedDRA-Systemorganklasse Preferred Term	Bis Jahr 1B <sup>2)3)</sup>		Bis Jahr 5 <sup>2)3)</sup>	
	Ursprüngliche Intervention (N = 20) <sup>4)</sup> n (%)	Ursprüngliche Kontrolle (N = 9) <sup>4)</sup> n (%)	Ursprüngliche Intervention (N = 20) <sup>4)</sup> n (%)	Ursprüngliche Kontrolle (N = 9) <sup>4)</sup> n (%)
<b>Erkrankungen des Blutes und des Lymphsystems</b>	9 (45,0)	2 (22,2)	9 (45,0)	2 (22,2)
Leukozytose <sup>5)</sup>	9 (45,0)	2 (22,2)	9 (45,0)	2 (22,2)
<b>Kongenitale, familiäre und genetische Erkrankungen</b>	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
Kongenitale Pneumonie	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
<b>Erkrankungen des Gastrointestinaltrakts</b>	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
Übelkeit	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
Erbrechen	2 (10,0)	0 (0,0)	2 (10,0)	0 (0,0)
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	1 (5,0)	0 (0,0)	1 (5,0)	0 (0,0)
Unerwünschte Arzneimittelwirkung	1 (5,0)	0 (0,0)	1 (5,0)	0 (0,0)
<b>Infektionen und parasitäre Erkrankungen</b>	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
Pneumonie	0 (0,0)	0 (0,0)	1 (5,0)	0 (0,0)
<b>Stoffwechsel- und Ernährungsstörungen</b>	1 (5,0)	0 (0,0)	1 (5,0)	0 (0,0)
Hyperkaliämie	1 (5,0)	0 (0,0)	1 (5,0)	0 (0,0)

MedDRA-Systemorganklasse Preferred Term	Bis Jahr 1B <sup>2)3)</sup>		Bis Jahr 5 <sup>2)3)</sup>	
	Ursprüngliche Intervention (N = 204) <sup>4)</sup> n (%)	Ursprüngliche Kontrolle (N = 9) <sup>4)</sup> n (%)	Ursprüngliche Intervention (N = 20) <sup>4)</sup> n (%)	Ursprüngliche Kontrolle (N = 9) <sup>4)</sup> n (%)
<b>Erkrankungen des Nervensystems</b>	1 (5,5)	0 (0,0)	1 (5,5)	0 (0,0)
Krampfanfall	1 (5,5)	0 (0,0)	1 (5,5)	0 (0,0)
<b>Erkrankungen der Geschlechts- organe und der Brustdrüse</b>	0 (0,0)	0 (0,0)	1 (5,5)	0 (0,0)
Menorrhagie	0 (0,0)	0 (0,0)	1 (5,5)	0 (0,0)
<b>Erkrankungen der Atemwege, des Brustraums und Mediastinums</b>	0 (0,0)	0 (0,0)	0 (0,0)	0 (0,0)
Atemnotsyndrom bei Neugeborenem	0 (0,0)	0 (0,0)	1 (5,5)	0 (0,0)

<sup>1)</sup> Die Darstellung erfolgt post hoc im Dossier Modul 4. Angaben im Studienbericht konnten nicht identifiziert werden. Es bleibt unklar, welches Klassifizierungssystem für den Schweregrad der UE verwendet wurde. Im Studienprotokoll erfolgte eine studienindividuelle Einteilung der UE in Schweregrade. Eine Schweregradeinteilung anhand der Toxizitätstabelle (Grad 1–4) sollte, sofern möglich, erfolgen.

<sup>2)</sup> Gemäß SAP war in der LTFU-Studie eine vollständige Erhebung der UE nur bis Jahr 1, jedoch nicht im Rahmen einer Langzeiterhebung für den gesamten Beobachtungszeitraum vorgesehen, ab Jahr 1 sollten im CRF nur bestimmte UE (onkologische Ereignisse, hämatologische Ereignisse, neurologische Ereignisse und/oder Autoimmunerkrankungen) dokumentiert, jedoch alle UE in den Patientenunterlagen dokumentiert werden.

<sup>3)</sup> Im Modul 4 werden UE für den Zeitraum Injektion ins erste Auge bis Jahr 1B und Injektion ins erste Auge bis Jahr 5 (= Zeitpunkt der letzten Injektion + 1.825 Tage) angegeben. Jedoch konnten keine Angaben zu (medianen) Beobachtungszeiten für diesen Untersuchungszeitraum identifiziert werden. Für die ursprüngliche Interventionsgruppe betrug die mittlere Beobachtungszeit 406,6 Tage (SD: 20,4) zwischen Tag erster Injektion und Jahr 1 nach zweiter Injektion [4].

<sup>4)</sup> Angaben zu den UE konnten nur im Interimsstudienbericht für den Zeitraum Injektion ins erste Auge bis Jahr 1B, nicht jedoch zu Jahr 5 identifiziert werden. Diese wurden aus dem Dossier Modul 4 entnommen.

<sup>5)</sup> Für die Studie 301 wird beschrieben, dass die Erfassung des UE Leukozytose nur ab einem bestimmten Schweregrad im Interventionsarm, während der Gabe von Glukokortikoiden (18–30 Tage, je nach Zeitpunkt der zweiten Injektion) nur als UE bei Schweregrad 3 oder höher erfolgen soll.

Abkürzungen: CRF: Case report form; MedDRA: Medical Dictionary for Regulatory Activities; SAP: Statistischer Analyseplan; SD: Standardabweichung; UE: Unerwünschte/s Ereignis/se.

## A 2 Beschreibung der ergänzend dargestellten Registerstudie 401

Die Registerstudie 401 war Bestandteil der Befristungsaufgaben zur Neubewertung der Langzeiteffekte nach Fristablauf, insbesondere zur Sicherheit [6,9]. Es wurden insgesamt 106 Teilnehmende, von denen 103 das Prüfpräparat erhalten haben, in die Studie aufgenommen. Die mediane Beobachtungszeit seit Injektion ins erste Auge bis zum Datenschnitt vom 31.08.2021 beträgt jedoch nur 0,65 Jahre (min; max: 0; 2,3). Da gemäß Befristungsgrundlagen eine Untersuchung von Langzeiteffekten von Relevanz ist und die Registerstudie derzeit die Anforderung aufgrund der kurzen Beobachtungszeit bis zum aktuellen Datenschnitt nicht erfüllt, wird die Studie nur ergänzend im Anhang dargestellt.

Tabelle 37: Charakterisierung der Registerstudie 401

Charakteristikum	Beschreibung
<b>Design und Studienablauf</b>	<p>Multizentrische, internationale, longitudinale, beobachtende nicht-interventionelle „Post-Authorisation-Safety Study“ (PASS) Registerstudie nach europäischer Zulassungserteilung von Voretigen Neparvovec. Die Studie wurde durch die EMA beauftragt.</p> <p>Das primäre Ziel der Studie ist die langfristige Erhebung von Sicherheitsdaten, insbesondere unerwünschter Ereignisse von besonderem Interesse, jeglicher anderer UE und schwerwiegender unerwünschter Ereignisse.</p> <p>Es soll möglichst allen Personen, die eine Behandlung mit Voretigen Neparvovec erhalten haben, die Teilnahme an der Studie ermöglicht werden. Teilnehmende der Registerstudie werden bis zu 5 Jahre nach der Behandlung mit Voretigen Neparvovec beobachtet.</p> <p>In die Studie eingeschlossene Patienten und Patientinnen werden gemäß den lokalen Zulassungsinformationen und der lokalen medizinischen Routinepraxis in Bezug auf Häufigkeit der Visiten und der durchgeführten Untersuchungen behandelt bzw. untersucht, und nur diese Daten werden im Rahmen der Studie erhoben. Es existieren kein vorgegebenes Behandlungsprotokoll, keine Vorgaben zu diagnostischen oder therapeutischen Prozeduren oder ein Visitenplan. Das zuständige ärztliche Personal bestimmte die Anzahl und den Zeitpunkt der Visiten. Im Studienprotokoll wird jedoch ein Zeitplan für eine empfohlene Mindestanzahl an Datenerhebungen vorgeschlagen. Die minimale Erfassung beinhaltet Baseline-/Vorbehandlungsdaten (u. a. klinische Diagnose in Zusammenhang zur erblichen retinalen Dystrophie, Begleitmedikation und -prozeduren, ophthalmologische Untersuchungen), Postbehandlungsdaten (Zeitpunkt und Informationen zur Injektion, Einsatz Steroide, Sicherheitsdaten) und jährliche Verlaufskontrollen zur Sicherheit und zu ophthalmologischen Untersuchungen (Jahr 1–5).</p> <p>Sofern Teilnehmende, die vor Einschluss die Intervention erhalten haben, in die Registerstudie eingeschlossen werden, sollen die im Protokoll vorgeschlagenen Mindestinformationen zu Baseline / vor der Behandlung, Informationen zur Applikation der Intervention und zu möglichen Begleitprozeduren/-medikation sowie Daten zur Sicherheit, ophthalmologische Untersuchungen möglichst vollständig retrospektiv erhoben werden. Für Teilnehmende, die die Intervention erst nach Einschluss in die Studie erhalten haben, können Informationen zu ophthalmologischen Untersuchungen bis zu 3 Monate vor Behandlung als Baseline-Werte erfasst werden.</p> <p>Es werden nur qualifizierte Studienzentren mit Studienpersonal, das ein Schulungsprogramm zur Anwendung der Intervention und zum Erkrankungsbild der erblichen retinalen Dystrophie erhalten hat, als primäre Behandlungszentren benannt. In diesen soll die Datenerhebung schwerpunktmäßig erfolgen.</p>



Charakteristikum	Beschreibung
<b>Population</b>	<p><b>Wesentliche Einschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Geplante oder bereits erhaltene Injektion von Voretigen Neparvec in mindestens ein Auge</li> </ul> <p><b>Wesentliche Ausschlusskriterien</b></p> <ul style="list-style-type: none"> <li>• Derzeitige oder vorherige Teilnahme an einer interventionellen klinischen Studie mit Voretigen Neparvec</li> </ul>
<b>Intervention und Zahl der Patientinnen und Patienten</b>	<p>N = 106 eingeschlossen (Datenschnitt: 31. August 2021)</p> <p>Die Applikation der Intervention soll nur an qualifizierten Zentren erfolgen, bei denen das Personal ein verpflichtendes Schulungsprogramm zum Einsatz des Präparats erhalten hat.</p>
<b>Ort und Zeitraum der Durchführung</b>	<p><b>Ort der Durchführung</b></p> <ul style="list-style-type: none"> <li>• Durchführung an 26 Studienzentren<sup>1)</sup> (4 in Deutschland) in 15 Ländern (Österreich, Belgien, Brasilien, Tschechien, Deutschland, Dänemark, Spanien, Frankreich, England, Kroatien, Israel, Italien, Niederlande, Polen, Slowakei)</li> </ul> <p><b>Zeitraum der Durchführung</b></p> <ul style="list-style-type: none"> <li>• Datum der Studieninitiierung: 18. Dezember 2019</li> <li>• Geplantes Studienende: Dezember 2029</li> </ul> <p><b>Datenschnitte</b></p> <ul style="list-style-type: none"> <li>• 1. Datenschnitt: 31. August 2020</li> <li>• Jährliche Datenschnitte geplant</li> <li>• Aktueller Datenschnitt: 31. August 2021</li> <li>• Finaler Studienbericht geplant zu Juni 2030</li> </ul>
<b>Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll</b>	<p><b>Primärer Endpunkt</b></p> <ul style="list-style-type: none"> <li>• Sicherheit: Erhebung aller UE von besonderem Interesse, anderer UE und SUE</li> </ul> <p><b>Sekundäre Endpunkte</b></p> <ul style="list-style-type: none"> <li>• Untersuchung der Sicherheit und Verträglichkeit von Voretigen Neparvec bei Schwangerschaft (Patientinnen sowie Partnerinnen von Patienten)</li> <li>• Ophthalmologische Untersuchungen <ul style="list-style-type: none"> <li>○ Sehschärfe (visual acuity (VA)), gemessen als „best corrected visual acuity“ (BCVA)</li> <li>○ Gesichtsfeld-Untersuchung (visual field (VF)), gemessen mittels Goldmann-Perimetrie und Humphrey-Perimetrie</li> <li>○ Lichtempfindlichkeit gemessen mittels Vollfeld-Lichtempfindlichkeit-Schwellenwert-Test (FST)</li> <li>○ Optische Kohärenztomographie (OCT)</li> </ul> </li> </ul> <p><b>Weitere Endpunkte</b></p> <ul style="list-style-type: none"> <li>• Fragebogen zu patientenberichteten Endpunkten: adaptierter VFQ-25 bzw. durch einen validierten Fragebogen ersetzt</li> <li>• Burden Index of Caregivers (BIC)</li> </ul>
<b>Subgruppenanalysen</b>	<p>Präspezifiziert:</p> <ul style="list-style-type: none"> <li>• Keine</li> </ul> <p>Post hoc im Studienbericht gemäß Auftreten des folgenden UE im Laufe der Studie:</p> <ul style="list-style-type: none"> <li>• Chorioretinale Atrophie</li> </ul>

<sup>1)</sup> Es werden im Dossier Modul 4 sowohl 26 Studienzentren als auch 39 Studienzentren genannt.

Abkürzungen: BCVA: best corrected visual acuity; BIC: Burden Index of Caregivers; FST: Full-Light-Sensitivity-Test; OCT: Optische Kohärenztomographie; (S)UE: (Schwerwiegende) unerwünschte Ereignisse, VFQ-25: National Eye Institute Visual Functioning Questionnaire – 25

## Protokolländerungen

### Registerstudie 401

Das Studienprotokoll wurde nach Version 01 (11. Juli 2019) einmal geändert. In Version 01-CH.01 (12. Juni 2020) wurden lokale Änderungen für die Schweiz bezüglich einer Spezifizierung der ophthalmologischen Beurteilungsmethoden für die Studienzentren vorgenommen.

## A 2.1 Endpunkte

In diesem Kapitel wird die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Alle Endpunkte der eingeschlossenen Studie (siehe Tabelle 37) wurden einer Prüfung unterzogen. Tabelle 38 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

*Tabelle 38: Zusammenfassung der Endpunktbewertung der Studie 401*

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Mortalität <sup>1)</sup>	Mortalität	Ja	Ja
Unerwünschte Ereignisse <sup>2)</sup>	Sicherheit	Ja	Ja

<sup>1)</sup> Über die Sicherheit erfasst

<sup>2)</sup> Primärer Endpunkt der Studie 401

Im Dossier werden für die Studie 401 die Endpunkte Lichtempfindlichkeit, Perimetrie nach Humphrey und/oder Goldmann und Sehschärfe nicht dargestellt, da keine standardisierte Erhebung, sondern eine Erhebung nach lokalen Vorgaben der Endpunkte in den Studienzentren vorgesehen ist. Der Ausführung des pU wird gefolgt und auf eine Darstellung der Endpunkte in der Kategorie Morbidität und Lebensqualität wird daher verzichtet. Zudem beträgt die Beobachtungszeit der einarmigen Studie derzeit im Median nur 0,65 Jahre und die Auswertungen erfolgten ausschließlich bezogen auf die behandelten Augen, nicht bezogen auf Personen, die eine Behandlung erhalten haben.

### A 2.1.1 Mortalität

Der Endpunkt Mortalität wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

In der Studie 401 wurden Todesfälle über die Sicherheit erhoben.

##### *Bewertung*

Die Operationalisierung ist 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

Die Erhebung wird als valide angesehen.

### A 2.1.2 Sicherheit

Der Endpunkt Sicherheit wird in der Nutzenbewertung berücksichtigt.

#### Operationalisierung

##### *Beschreibung*

In der Studie wurden unerwünschte Ereignisse von substanzspezifischem Interesse (AESI) sowie alle anderen unerwünschten Ereignisse (UE und SUE) erfasst. Die Kodierung erfolgt gemäß MedDRA nach SOC und PT sowie nach einer Unterteilung in okuläre und nicht-okuläre UE. Bei Teilnehmenden, die vor Einschluss in die Studie die Intervention bereits erhalten haben, sollen möglichst vollständig die im Protokoll vorgeschlagenen Mindestinformationen zu Baseline / vor der Behandlung und über die Behandlung retrospektiv erfasst werden.

Um zwischen den Behandlungsgruppen vergleichbare Beobachtungszeiträume zugrunde zu legen, wurden die UE jeweils für 2 Perioden ausgewertet:

1. Tag der ersten Injektion bis 1 Jahr nach der zweiten Injektion
2. Tag der ersten Injektion bis 5 Jahre nach der zweiten Injektion

Für die AESI war eine standardisierte Erhebung mittels eines Fragebogens vorgesehen. Als AESI sind definiert:

- Okuläre UE
- Mangelnde Wirksamkeit und/oder Abnahme der Wirksamkeit über die Zeit
- Übertragung durch oder an Dritte („third party transmission“)
- Immunantwort des Hosts („Host immune response“)
- Entwicklung oder Exazerbation onkologischer, hämatologischer, neurologischer oder autoimmuner Krankheiten

Weitere Sicherheitsdaten werden mit nicht-direktiver Befragung der Teilnehmenden, Dokumentation während oder zwischen Visiten, durch körperliche Untersuchung oder Labortests erhoben.

Als schwerwiegendes unerwünschtes Ereignis (SUE) waren UE definiert, die

- zum Tod
- zu einem lebensbedrohlichen Ereignis
- zu einem stationären Krankenhausaufenthalt oder zu einem verlängerten Krankenhausaufenthalt mit Ausnahme von
  - Routinebehandlung oder Überwachung bezüglich der Indikation, die nicht mit einer Verschlechterung der Erkrankung assoziiert ist. Dies beinhaltet eine Hospitalisierung vor oder nach Injektion aufgrund einer Routine-Überwachung gemäß lokalen Vorgaben
  - Geplante oder vorhersehbare Eingriffe für eine bereits bestehende Erkrankung, die nicht in Zusammenhang mit der Indikation stehen und die keine Verschlechterung seit Applikation des Prüfpräparats bedeutet
  - Soziale Gründe und Kurzzeitpflegen, wenn sich der Gesundheitszustand des Teilnehmenden nicht verschlechtert
- zu einer langanhaltenden oder signifikanten Unfähigkeit oder erheblichen Beeinträchtigung der Ausführung von alltäglichen Aktivitäten oder
- zu einer angeborenen Anomalie bzw. Geburtsfehler führten.
- als medizinisch signifikant angesehen werden: u. a. Ereignisse, die den/die Patienten/Patientin gefährden; medizinische, chirurgische Interventionen zur Prävention eines Ereignisses (definiert als SUE) zu ermöglichen oder ein Ereignis, das

nicht den oben genannten Definitionen eines SUE entspricht und eine Krankenhauseinweisung bzw. ambulante Notfallbehandlung erforderlich macht.

Eine studienindividuelle Schweregradeinteilung wurde vom pU sofern möglich folgendermaßen vorgenommen:

- Mild: Symptome sind für die/den Teilnehmende/n kaum spürbar und haben keinen Einfluss auf ihre/seine Leistung oder Funktion.
- Moderat: Symptome mit ausreichendem Schweregrad, sodass Teilnehmende Beschwerden haben bzw. in der Verrichtung von täglichen Aktivitäten eingeschränkt sind. Eine Behandlung der Symptome könnte notwendig sein.
- Schwer: Symptome verursachen schwere Beschwerden. Eine Behandlung der Symptome und/oder Hospitalisierung kann erfolgen.
- Lebensgefährlich: Das Auftreten eines Ereignisses kann zu unmittelbarer Todesgefahr führen.
- Tod: Das Auftreten eines Ereignisses führt zum Tod.

#### *Bewertung*

Die Operationalisierung ist überwiegend nachvollziehbar. Es bleibt unklar, wie eine möglichst vollständige Erhebung von UE bei Teilnehmenden, die vor Einschluss in die Studie das Prüfpräparat erhalten haben, gewährleistet werden soll. Eine Erhebung von Endpunkten war nach lokalen Vorgaben vorgesehen und es wurden nur jährliche Visiten empfohlen.

#### Patientenrelevanz:

Das Auftreten von unerwünschten Ereignissen wird als patientenrelevant angesehen.

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“.

#### Validität:

##### *Bewertung*

Die Erhebung wird überwiegend als valide angesehen. Zu beachten ist, dass bei Teilnehmenden, die das Prüfpräparat vor Einschluss in die Studie erhalten haben, Daten retrospektiv erfasst werden sollen, wodurch die Gefahr einer nicht vollständigen Erhebung besteht.

*Tabelle 39: Erhebungszeitpunkte der ergänzend dargestellten Endpunkte in der Registerstudie 401*

Studiensite Endpunkt	Vor Behandlung/ Baseline <sup>1)</sup>	Nach Behandlung <sup>1)</sup>	Jährliche Visite bis Jahr 5 <sup>2)3)</sup>
Sicherheit		x	x

<sup>1)</sup> Bei Teilnehmenden, die die Intervention vor Einschluss in die Studie bereits erhalten haben, sollen möglichst vollständig die im Protokoll vorgeschlagenen Mindestinformationen zu Baseline / vor der Behandlung und über die Behandlung retrospektiv erfasst werden. Bei Teilnehmenden, die erst nach Einschluss die Intervention erhalten haben, können Informationen zu ophthalmologischen Untersuchungen bis zu 3 Monate vor Behandlung als Baseline erfasst werden.

<sup>2)</sup> Empfohlene Mindestanzahl an Visiten; Teilnehmende sollen bis zu 5 Jahre nach Applikation nachbeobachtet werden.

<sup>3)</sup> Oder letzte Studiensite vor Ausscheiden aus der Studie.

## A 2.2 Statistische Methoden

### Relevante Änderungen in den Versionen des Statistischen Analyseplans

Es existieren 2 Versionen des SAP. Die erste Version ist datiert auf den 14. August 2017, die zweite Version auf den 19. Oktober 2021. Es wurden in dieser Version des SAP neue Auswertungen zu Patientencharakteristika, Dosierung und Wirksamkeit für eine Subpopulation, bei der das TEAE „chorioretinale Atrophie“ während der Studie beobachtet wurde, aufgeführt.

### Analysepopulationen

- Alle eingeschlossenen Patientinnen und Patienten (N = 106)
- Full Analysis Set (FAS): Alle eingeschlossenen Teilnehmenden, bei denen das Prüfpräparat mindestens einmal appliziert wurde.

### Datenschnitte

Es waren jährlich Datenschnitte vorgesehen. Der erste Datenschnitt ist vom 31. August 2020 und der zweite und aktuelle Datenschnitt ist vom 31. August 2021.

### Präspezifizierung geplanter und durchgeführter Analysen

Als Baseline war der Tag der Injektion ins erste Auge definiert. Es war eine deskriptive Auswertung vorgesehen.

### Fehlende Werte und Daten-Imputation

Es waren keine Imputationen vorgesehen.

## A 2.3 Ergebnisse

Es werden die Studiencharakteristika und Ergebnisse zur Sicherheit der Registerstudie 401 dargestellt.

### A 2.3.1 Studiencharakteristika und Studienmedikation

In Tabelle 40 sind die allgemeinen Angaben zur Registerstudie 401 dargestellt. Die mediane Beobachtungsdauer zwischen Injektion ins erste Auge und Datenschnitt (31. August 2021) lag bei 238 Tagen (min; max: 14; 833) bzw. 0,65 Jahren (0,0; 2,3); bei 77 % der Teilnehmenden lag die Beobachtungsdauer in der Studie < 1 Jahr. 106 Personen wurden eingeschlossen, von denen 3 zum Zeitpunkt des Datenschnitts keine Behandlung erhielten. Insgesamt erhielten 41 % der Teilnehmenden eine Behandlung vor Studieneinschluss, 18 % nur in ein Auge. Gründe für die fehlende Applikation ins zweite Auge vor Studieneinschluss konnten nicht identifiziert werden.

Die mittlere Zeit zwischen Erst- und Zweitinjektion in Tagen betrug 37,1 (SD 82,9), während die mediane Zeit bei 7 Tagen (7; 476,0) lag. Weitere Angaben zu den Beobachtungszeiten, wie die Zeit zwischen der Injektion ins erste Auge und Studieneinschluss für Teilnehmende, die vor Einschluss in die Registerstudie das Prüfpräparat erhielten, bzw. für Teilnehmende, die das Prüfpräparat erst nach Einschluss erhielten, die mediane Beobachtungszeit zwischen Einschluss ins Register und Zeitpunkt der ersten Injektion, konnten nicht identifiziert werden.

Angaben zu Vortherapie und Begleitmedikation wurden vom pU für jede Person individuell berichtet. Auf eine Darstellung wird daher verzichtet.

Tabelle 40: Registerstudie 401: Allgemeine Angaben; Datenschnitt vom 31. August 2021

<b>Registerstudie 401</b>	<b>Alle eingeschlossenen Personen N = 106</b>
Anzahl eingeschlossener Personen in die Studie, n (%) FAS-Population <sup>1)</sup>	106 (100,0) 103 (97,2)
Abbruch der Studie bis zum Datenschnitt, n (%)	0 (0)
Mediane Beobachtungsdauer zwischen Injektion ins erste Auge und Datenschnitt, Tage (min; max)	238 (14; 833)
Mediane Beobachtungsdauer zwischen Injektion ins erste Auge und Datenschnitt, Jahre (min; max)	0,65 (0,0; 2,3)
Mediane Beobachtungsdauer zwischen Injektion ins erste Auge und Einschluss in die Studie, Jahre (min; max)	k. A.
Mediane Beobachtungsdauer zwischen Einschluss in die Studie und Injektion ins 1. Auge Jahre (min; max)	k. A.
Beobachtungsdauer in der Studie, n (%) <sup>2)</sup>	
< 1 Jahr	79 (76,7)
1 Jahr bis < 2 Jahre	17 (16,5)
2 Jahre bis < 3 Jahre	7 (6,8)
Zeit zwischen Erst- und Zweitinjektion in Tagen	
Mittelwert (SD)	37,1 (82,9)
Median (min; max)	7 (7; 476,0)
Behandlungsstatus zum Zeitpunkt des Einschlusses in die Studie, n (%)	
Nicht behandelt vor Studieneinschluss	63 (59,4)
Behandelt vor Studieneinschluss	43 (40,6)
Behandlung an einem Auge	19 (17,9)
Behandlung an beiden Augen	24 (22,6)

<sup>1)</sup> Alle eingeschlossenen Teilnehmenden, bei denen das Prüfpräparat mindestens einmal appliziert wurde.

<sup>2)</sup> FAS-Population

Abkürzungen: FAS: Full Analysis Set; k. A.: keine Angabe; SD: Standardabweichung

In Tabelle 41 sind die Charakteristika der Population der Registerstudie 401 dargestellt. Es wurden Teilnehmende aus dem europäischen Raum, überwiegend aus Deutschland (23 %) und Frankreich (20 %), eingeschlossen.

Tabelle 41: Registerstudie 401: Charakterisierung der Studienpopulation

<b>Registerstudie 401</b>	<b>Alle eingeschlossenen Personen N = 106</b>
<i>Alter (Jahre)</i>	
Mittelwert (SD)	19,6 (10,8)
Median (min; max)	18,5 (2; 51)
<i>Altersgruppe (Jahre), n (%)</i>	
2–11	31 (29,2)
12–18	21 (19,8)
18–65	54 (50,9)
<i>Geschlecht, n (%)</i>	
Männlich	50 (47,2)
Weiblich	55 (51,9)
Unbekannt	1 (0,9)

<b>Registerstudie 401</b>	<b>Alle eingeschlossenen Personen N = 106</b>
<i>Abstammung (genetisch), n (%)</i>	
Weiß	74 (69,8)
Schwarz oder Afro-Amerikanisch	4 (3,8)
Asiatisch	4 (3,8)
Unbekannt	24 (22,6)
<i>Region, n (%)</i>	
Lateinamerika	1 (0,9)
Europa	105 (99,1)
<i>Anzahl an teilnehmenden Ländern</i>	15
<i>Länder (mind. 10 % der Teilnehmenden), n (%)</i>	
Deutschland	24 (22,6)
Dänemark	11 (10,4)
Frankreich	21 (19,8)
Vereinigtes Königreich	12 (11,3)
Italien	15 (14,2)

Abkürzungen: SD: Standardabweichung

### Exposition mit der Studienmedikation

Angaben zur Exposition wurden überwiegend in Bezug zu den behandelten Augen, jedoch nicht pro Patient/in vorgenommen. Bis zum aktuellen Datenschnitt (31. August 2021) erhielten 80 Teilnehmende (78 %) eine bilaterale Behandlung, während 22 % eine Behandlung in nur ein Auge erhielten. Insgesamt erhielten 158/183 der behandelten Augen (86 %) das empfohlene Volumen von 300 µl. 22 Augen (12 %) erhielten weniger als 300 µl, während 1 Auge mehr als 300 µl erhielt. Das mediane Volumen betrug 300 µl (min; max: 80; 350).

Es konnten Abweichungen vom Standardverfahren beobachtet werden. Unter anderem wurde bei 23,5 % der behandelten Augen (N = 183) entgegen der empfohlenen Applikation ein automatisiertes Injektionssystem verwendet.

#### **A 2.3.2 Mortalität**

Todesfälle wurden über die Sicherheit erfasst. Es wurden bis zum Datenschnitt vom 31. August 2021 keine Todesfälle berichtet.

#### **A 2.3.3 Morbidität**

Es wurden keine geeigneten Daten vorgelegt.

#### **A 2.3.4 Lebensqualität**

Es wurden keine geeigneten Daten vorgelegt.

#### **A 2.3.5 Sicherheit**

In Tabelle 42 sind die zusammenfassenden UE der Registerstudie 401 dargestellt.

Tabelle 42: Registerstudie 401: Ergebnisse der Nebenwirkungen (Datenschnitt: 31. August 2021)

Personen mit mindestens einem ...	FAS-Population <sup>1)2)</sup> N = 103 <sup>3)</sup> n (%)
UE	39 (37,9)
Okuläre UE	35 (34,0)
Nicht-okuläre UE	8 (7,8)
Schwere UE	1 (1,0)
Okuläre UE	0 (0,0)
Nicht-okuläre UE	1 (1,0)
SUE	3 (2,9)
Okuläre UE	2 (1,9)
Nicht-okuläre UE	1 (1,0)
UE, das zum Abbruch der Studienmedikation führte	0 (0,0)

<sup>1)</sup> FAS-Population: Alle Teilnehmenden, die das Prüfpräparat erhalten haben.

<sup>2)</sup> Die mediane Beobachtungszeit betrug vom Tag der ersten Injektion bis Datenschnitt (31. August 2021) 238 Tage (min; max: 14; 833).

<sup>3)</sup> Bei 41 % der Teilnehmenden erfolgte die Applikation des Prüfpräparats vor Studienbeginn. Gemäß Studienprotokoll sollte retrospektiv eine Erfassung der Daten erfolgen. Jedoch bleibt unklar, inwieweit eine vollständige Erfassung der UE möglich war.

Abkürzungen: FAS: Full Analysis Set; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis

### Unerwünschte Ereignisse

Es traten insgesamt nur wenige okuläre UE bis zum Datenschnitt (31. August 2021) auf. Es konnten keine nicht-okulären UE mit einer Inzidenz  $\geq 10,0$  % bis zum Datenschnitt beobachtet werden.

Tabelle 43: Registerstudie 401: Okuläre UE mit Inzidenz  $\geq 10,0$  %, FAS-Population (Datenschnitt: 31. August 2021)

MedDRA-Systemorganklasse Preferred Term	FAS-Population <sup>1)2)</sup> N = 103 <sup>3)</sup> n (%)
<b>Augenerkrankungen</b>	27 (26,2)
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>	12 (11,7)
Injektionsstelle atrophisch	12 (11,7)

<sup>1)</sup> FAS-Population: Alle Teilnehmenden, die das Prüfpräparat erhalten haben.

<sup>2)</sup> Die mediane Beobachtungszeit betrug vom Tag der ersten Injektion bis Datenschnitt (31. August 2021) 238 Tage (min; max: 14; 833).

<sup>3)</sup> Bei 41 % der Teilnehmenden erfolgte die Applikation des Prüfpräparats vor Studienbeginn. Gemäß Studienprotokoll sollte retrospektiv eine Erfassung der Daten erfolgen. Jedoch bleibt unklar, inwieweit eine vollständige Erfassung der UE möglich war.

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

### Schwere UE

Es konnten keine Angaben zu schweren UE auf SOC- und PT-Ebene identifiziert werden.

### Schwerwiegende unerwünschte Ereignisse

Auf SOC- und PT-Ebene traten SUE nur bei jeweils einer Person auf: SOC „Augenerkrankungen“, PT „Augenentzündung“, SOC „Untersuchungen“, PT „intraokulärer



Druck erhöht“, SOC „Psychiatrische Erkrankungen“, PT „Aggression“, PT „Agitiertheit“, PT „Delirium“.

### Unerwünschte Ereignisse von besonderem Interesse

In Tabelle 44 sind die okulären UE von besonderem Interesse für die Registerstudie bis zum Datenschnitt vom 31. August 2021 dargestellt. Es traten keine nicht-okulären UE von besonderem Interesse bis zum aktuellen Datenschnitt auf.

*Tabelle 44: Registerstudie 401: Okuläre UE von besonderem Interesse<sup>1)</sup>, FAS-Population;  
Datenschnitt: 31. August 2021*

<b>MedDRA-Systemorganklasse</b> Preferred Term	<b>FAS-Population<sup>2)3)</sup></b> <b>N = 103<sup>4)</sup></b> <i>n (%)</i>
<b>Ausdünnung der Fovea</b>	4 (3,9)
Foveale Degeneration	4 (3,9)
Makulanarbe	1 (1,0)
<b>Erhöhter intraokulärer Druck</b>	5 (4,9)
Intraokulärer Druck erhöht	5 (4,9)
<b>Intraokuläre Entzündung und/oder Infektion im Zusammenhang mit der Verabreichung</b>	7 (6,8)
Augenentzündung	3 (2,9)
Vitritis	4 (3,9)
<b>Makulaloch</b>	1 (1,0)
Makulaloch	1 (1,0)
<b>Makulopathie (z. B. epiretinale Membran, Makula Pucker)</b>	2 (1,9)
Makulafibrose	1 (1,0)
Retinoschisis	1 (1,0)
<b>Netzhaut einriss</b>	<b>2 (1,9)</b>
Netzhaut einriss	2 (1,9)

<sup>1)</sup> Erhebung über einen standardisierten Fragebogen.

<sup>2)</sup> FAS-Population: Alle Teilnehmenden, die das Prüfpräparat erhalten haben.

<sup>3)</sup> Die mediane Beobachtungszeit betrug vom Tag der ersten Injektion bis Datenschnitt (31. August 2021) 238 Tage (min; max: 14; 833).

<sup>4)</sup> Bei 41 % der Teilnehmenden erfolgte die Applikation des Prüfpräparats vor Studienbeginn. Gemäß Studienprotokoll sollte retrospektiv eine Erhebung der Daten erfolgen. Jedoch bleibt unklar, inwieweit eine vollständige Erfassung der UE rückwirkend möglich war.

Abkürzungen: FAS: Full Analysis Set, UE: Unerwünschtes Ereignis.

### Weitere Analysen: Chorioretinale Atrophie

Bei der Auswertung zum Zwischenbericht 2020 wurden vermehrt chorioretinale Atrophien beobachtet. Von den 103 Personen, die das Prüfpräparat erhalten haben, trat das UE bei 13 Personen (12,6 %) in mindestens einem Auge auf.