

Nutzenbewertung



Gemeinsamer
Bundesausschuss

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

**Bewertung von Arzneimitteln für seltene Leiden
nach § 35a Absatz 1 Satz 11 i.V.m.**

5. Kapitel § 12 Nr. 1 Satz 2 Verfo

Wirkstoff: Onasemnogen-Abeparvovec

Datum der Veröffentlichung: 1. Oktober 2020

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 Aufbau und Design der eingeschlossenen Studien	13
2.2.1 Zulassungsstudien	13
2.2.2 Externe Kontrollen.....	33
2.3 Endpunkte.....	42
2.3.1 Mortalität.....	43
2.3.2 Morbidität.....	46
2.3.3 Lebensqualität	55
2.3.4 Sicherheit	56
2.3.5 Übersicht der Vergleichbarkeit der Endpunkte in den Studien CL-101 und CL-303 mit den externen Kontrollen (PNCr, NeuroNEXT, ENDEAR).....	58
2.4 Statistische Methoden	63
2.4.1 Studie CL-303	63
2.4.2 Studie CL-101	66
2.4.3 Studien CL-302 und CL-304	68
2.4.4 Indirekter Vergleich der Studien CL-303 und CL-101 vs. externe Kontrolle (PNCr, NeuroNEXT, ENDEAR)	68
2.5 Verzerrungspotential auf Studien- und Endpunktebene	71
3 Ergebnisse der eingeschlossenen Studien.....	73
3.1 Studiencharakteristika und Studienmedikation	73
3.2 Mortalität	78
3.3 Morbidität	84
3.4 Lebensqualität	93
3.5 Sicherheit.....	93

4	Diskussion der Methodik und Ergebnisse	101
4.1	Zulassungsstatus und Zulassungspopulation von Onasemnogen-Abeparvovec (Zolgensma®)	101
4.2	Design und Methodik der Studie	101
4.3	Externe Kontrollen (PNCR, NeuroNEXT, ENDEAR)	103
4.4	Mortalität	108
4.5	Morbidität	111
4.6	Lebensqualität	113
4.7	Sicherheit	113
5	Anforderung an eine qualitätsgesicherte Anwendung	116
6	Zusammenfassung der Nutzenbewertung	117
	Referenzen	119
	Anhang 1	122
	Anhang 2	123
	Anhang 3	124

Tabellenverzeichnis

Tabelle 1:	Spezifische Verpflichtung zum Abschluss von Maßnahmen nach der Zulassung unter „Besonderen Bedingungen“ [15]	10
Tabelle 2:	Übersicht über die Studienbasis	11
Tabelle 3:	Charakterisierung der Studie CL-303 (STR1VE-US).....	13
Tabelle 4:	Wesentliche Protokolländerungen Studie CL-303 (STR1VE-US)	16
Tabelle 5:	Charakterisierung der Studien CL-101 (START) und LT-001 (Langzeit-Extensionsstudie von Studie CL-101)	17
Tabelle 6:	Wesentliche Protokolländerungen Studie CL-101	20
Tabelle 7:	Charakterisierung der Intervention in der Studie CL-303.....	22
Tabelle 8:	Charakterisierung der Intervention in der Studie CL-101.....	23
Tabelle 9:	Charakterisierung der Studien CL-302 (STR1VE-EU) und CL-304.....	24
Tabelle 10:	Charakterisierung der externen Kontrollen PNCR, NeuroNEXT und ENDEAR (BSC)	38
Tabelle 11:	Zusammenfassung der Endpunktbewertung.....	42
Tabelle 12:	Standardübersicht Entwicklung motorischer Meilensteine bis zum 60. Lebensmonat	49
Tabelle 13:	Vergleichbarkeit der Endpunkte der Studien CL-303 und CL-101 mit den externen Vergleichskohorten (PNCR, NeuroNEXT, ENDEAR (BSC)).....	58
Tabelle 14:	Relevante Analysepopulationen in den für die Nutzenbewertung berücksichtigten Studien	70
Tabelle 15:	Allgemeine Angaben für die Studien CL-101, CL-303, CL-302, CL-304 und LT-001.....	74
Tabelle 16:	Charakterisierung der Studienpopulationen zu Studienbeginn: Studien CL-101, CL-303, CL-302, CL-304 und LT-001 – Sicherheitspopulationen.....	75
Tabelle 17:	Ergebnisse zum Endpunkt „Mortalität“ in den Studien CL-303 und CL-101 und den externen Kontrollstudien (PNCR, NeuroNEXT, ENDEAR) (Alter: 13–14 Monate).....	79
Tabelle 18:	Ergebnisse zum Endpunkt „Mortalität“ in den Studien CL-101+CL-303 im Vergleich zu den externen Kontrollstudien PNCR+NeuroNEXT (Alter: 13–14 Monate).....	79
Tabelle 19:	Ergebnisse zum Endpunkt „Mortalität“ in den Studien CL-101+CL-303 im Vergleich zur externen Kontrollstudie ENDEAR (Alter: 13–14 Monate)	80
Tabelle 20:	Ergebnisse zum Endpunkt „Überleben“ (Tod oder dauerhafte Beatmung) in den Studien CL-101 und CL-303.....	81
Tabelle 21:	Ergebnisse zum Endpunkt „Überleben“ (Tod oder dauerhafte Beatmung) in den Studien CL-303 und CL-101 und den externen Kontrollen (PNCR, NeuroNEXT, ENDEAR) (Alter: 13–14 Monate)	82
Tabelle 22:	Ergebnisse zum Endpunkt „Überleben“ (Tod oder dauerhafte Beatmung) in den Studien CL-101+CL-303 im Vergleich zur externen Kontrolle PNCR+NeuroNEXT (Alter: 13–14 Monate).....	83

Tabelle 23: Ergebnisse zum Endpunkt „Überleben (Tod oder dauerhafte Beatmung) in den Studien CL-101+CL-303 im Vergleich zur externen Kontrolle ENDEAR (Alter: 13–14 Monate).....	83
Tabelle 24: Ergebnisse zum Endpunkt „Dauerhafte Beatmung“ in den Studien CL-303 und CL-101 im Vergleich zur externen Kontrolle PNCR (Alter: 13–14 Monate)	84
Tabelle 25: Ergebnisse zum Endpunkt „Intubation“ in den Studien CL-101 und CL-303 im Vergleich zur externen Kontrolle NeuroNEXT (Alter: 13–14 Monate)	85
Tabelle 26: Ergebnisse zum Endpunkt „Dauerhafte Beatmung“ in den Studien CL-303 und CL-101 im Vergleich zur externen Kontrolle ENDEAR (Alter: 13–14 Monate) ..	85
Tabelle 27: Ergebnisse zum Endpunkt „Erreichen motorischer Meilensteine“ in den Studien CL-303 und CL-101 (gemäß zentralem Review)	87
Tabelle 28: Ergebnisse zum Endpunkt „Unabhängiges Sitzen (≥ 30 Sekunden)“ in den Studien CL-101 und CL-303.....	88
Tabelle 29: Ergebnisse zum Endpunkt „Bayley-Skala – Grobmotorik“ in der Studie CL-303.....	88
Tabelle 30: Ergebnisse zum Endpunkt „Bayley-Skala – Feinmotorik“ in der Studie CL-303.....	89
Tabelle 31: Ergebnisse zum Endpunkt „CHOP INTEND ¹⁾ “ in der Studie CL-303	90
Tabelle 32: Ergebnisse zum Endpunkt „CHOP INTEND“ in der Studie CL-101.....	91
Tabelle 33: Ergebnisse zum Endpunkt „CHOP-INTEND-Schwellenwerte“ in den Studien CL-303 und CL-101	93
Tabelle 34: Ergebnisse der Nebenwirkungen: Zusammenfassung der UE in den Studien CL-303, CL-101, CL-302, CL-304 und LT-001 – Sicherheitspopulation.....	94
Tabelle 35: UE mit Inzidenz ≥10 % in den Studien CL-303 und CL-101 – Sicherheitspopulation	95
Tabelle 36: Schwere UE (Grad ≥ 3) in den Studien CL-303 und CL-101 – Sicherheitspopulation	97
Tabelle 37: SUE mit Inzidenz ≥ 5 % in den Studien CL-303 und CL-101 – Sicherheitspopulation	98
Tabelle 38: (Prä-)Spezifizierte unerwünschte Ereignisse von besonderem Interesse in der Studie CL-303 – Sicherheitspopulation	99
Tabelle 39: Ergebnisse zum Endpunkt „Hospitalisierung“ in den Studien CL-101 und CL-303	100
Tabelle 40: Zusammenfassende Darstellung der Ergebnisse für die Studien CL-101 und CL-303 (finaler Datenschnitt).....	117
Tabelle 41: Charakterisierung der Population mit SMA Typ 1 der PNCR-Kohorte	122
Tabelle 42: Charakterisierung der Population der NeuroNEXT-Kohorte.....	123
Tabelle 43: Charakterisierung der Studienpopulationen CL-101 und CL-303 mit den externen Kontrollen (PNCR, NeuroNEXT, ENDEAR (BSC)).....	125

Abbildungsverzeichnis

Abbildung 1: Studienverlauf CL-303 (STR1VE-US) nach Studienperioden	13
Abbildung 2: Studienverlauf CL-101 nach Studienperioden	17
Abbildung 3: Studienverlauf CL-302 (STR1VE-EU) nach Studienperioden	25
Abbildung 4: Studienverlauf CL-304 – Kohorte mit 2 SMN2-Genkopien nach Studienperioden ..	24
Abbildung 5: Studienverlauf CL-304 – Kohorte mit 3 SMN2-Genkopien nach Studienperioden ..	25
Abbildung 6: Patientenfluss im PNCR-Datensatz.....	34
Abbildung 7: Patientenfluss im NeuroNEXT-Datensatz	36
Abbildung 8: Altersabhängige Entwicklungsstufen Säuglinge, gemäß WHO.....	51
Abbildung 9: CHOP INTEND – Werte-Entwicklung in Studie CL-101 nach Kohorten.....	92

Abkürzungsverzeichnis

AESI	Unerwünschte Ereignisse von besonderem Interesse
AIMS	Alberta Infant Motor Scale
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
BiPAP	Bi-level Positive Airway Pressure
BSC	Best Supportive Care
CHOP INTEND	The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
CTCAE	Common Terminology Criteria for Adverse Events
DSMB	Data Safety Monitoring Board
EMA	European Medicines Agency
EPAR	European Public Assessment Report
FAS	Full Analysis Set
FDA	U. S. Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
GMS	Gross Motor Skills
HINE	Hammersmith Infant Neurological Examination
ITT	Intention-to-Treat
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KI	Konfidenzintervall
LS	Least Square
MedDRA	Medical Dictionary for Regulatory Activities
MMRM	Mixed Model of Repeated Measures
NeuroNEXT	National Network for Excellence in Neuroscience Clinical Trials
PNCR	Pediatric Neuromuscular Clinical Research Network for SMA
pU	pharmazeutischer Unternehmer
RR	Relatives Risiko
RSV	Respiratory Syncytical Virus
SAP	Statistischer Analyseplan
SD	Standardabweichung
SE	Standardfehler
SGB	Sozialgesetzbuch
SMA	Spinale Muskelatrophie (Spinal Muscular Atrophy)
SMN	Survival Motor Neuron
SUE	Schwerwiegende/s unerwünschte/s Ereignis/se
TIMPSI	Test of Infant Motor Performance Screening Items

UE	Unerwünschte/s Ereignis/se
VerfO	Verfahrensordnung des G-BA
vg	Vektorgenom/e
WHO	World Health Organization
WHO MGRS	World Health Organization Multicenter Growth Reference Study

Hintergrund

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

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

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

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

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Onasemnogen-Abeparvovec in seiner Sitzung am 21. September 2020 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 29. Juni 2020 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

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

Die Nutzenbewertung wird am 1. Oktober 2020 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der G-BA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Fragestellung

Onasemnogen-Abeprarovovec (Zolgensma®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Abs. 1 Satz 11 1. Halbs. SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der G-BA bestimmt bei Orphan Drugs, die einen Umsatz von 50 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. In die Bewertung werden gemäß Zulassungsstatus Patientinnen und Patienten in folgender Indikation eingeschlossen:

- 5q13-assoziierte spinaler Muskelatrophie (SMA) mit biallelischer Deletion des SMN1-Gens und entweder einer klinischen Diagnose SMA Typ 1 oder bis zu drei Kopien des SMN2-Gens.

Onasemnogen-Abeprarovovec (Zolgensma®) wurde unter „Besonderen Bedingungen“ zugelassen.

Tabelle 1: Spezifische Verpflichtung zum Abschluss von Maßnahmen nach der Zulassung unter „Besonderen Bedingungen“ [15]

Beschreibung	Fällig am
Wirksamkeitsprüfung nach der Zulassung (PAES): Um die Wirksamkeit, Sicherheit und Verträglichkeit einer Einzeldosis Zolgensma bei Patientinnen und Patienten unter 6 Monaten mit SMA Typ 1 mit einer oder zwei SMN2-Genkopien zu bestätigen, soll der Inhaber der Genehmigung für das Inverkehrbringen (MAH) Abschlussdaten zu der Studie AVXS-101-CL-303 vorlegen.	<u>Abschlussergebnisse:</u> bei der ersten jährlichen Verlängerung
Wirksamkeitsprüfung nach der Zulassung (PAES): Um die Wirksamkeit, Sicherheit und Verträglichkeit einer Einzeldosis Zolgensma bei Patientinnen und Patienten unter 6 Monaten mit SMA Typ 1 mit einer oder zwei SMN2-Kopien zu bestätigen, soll der Inhaber der Genehmigung für das Inverkehrbringen (MAH) Interimsdaten und Abschlussdaten zu der Studie AVXS-101-CL-302 vorlegen.	<u>Zwischenergebnisse:</u> bei jeder jährlichen Verlängerung <u>Abschlussergebnisse:</u> August 2021
Wirksamkeitsprüfung nach der Zulassung (PAES): Zur Bestätigung der Wirksamkeit, Sicherheit und Verträglichkeit einer Einzeldosis Zolgensma bei präsymptomatischen SMA-Patientinnen und -Patienten mit genetischer Diagnose, die zum Zeitpunkt der Behandlung 6 Wochen alt oder jünger sind und die eine biallelische Deletion von SMN1 und 2 oder 3 Kopien von SMN2 aufweisen, soll der Inhaber der Genehmigung für das Inverkehrbringen (MAH) Interimsdaten und Abschlussdaten zu der Studie AVXS-101-CL-304 vorlegen.	<u>Zwischenergebnisse:</u> bei jeder jährlichen Verlängerung <u>Abschlussergebnisse:</u> August 2026

Abkürzungen: MHA: Marketing Authorisation Holder; SMN: Survival Motor Neuron; PAES: Post-Authorisation Efficacy Study; SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron.

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 2: Übersicht über die Studienbasis

Studienname (Nummer)	Studie mit Dossier vom pU eingereicht (Ja/Nein)	Studie vom pU als relevant für Ableitung des Zusatznutzens erachtet (Ja/Nein)	Studie relevant für die Nutzenbewertung (Ja/Nein)	Ausschlussgrund
Studien zum Wirkstoff Onasemnogen-Abeparvovec¹⁾				
CL-303 ²⁾	Ja	Ja	Ja	-
CL-101 ³⁾	Ja	Ja	Ja	-
CL-302 ⁴⁾	Ja	Ja ⁵⁾	Ja	-
CL-304 ⁶⁾	Ja	Ja ⁵⁾	Ja	-
CL-306	Nein	Nein	Nein	Laufende, offene Phase-III-Studie in Japan, Südkorea und Taiwan bei Patientinnen und Patienten mit SMA Typ 1 und ein oder zwei SMN2-Genkopien. Bisher wurde nur eine Person in die Studie eingeschlossen. Es sind aktuell keine verwertbaren Daten verfügbar. Zudem ist die Übertragbarkeit auf den deutschen Versorgungskontext nicht gegeben.
LT-001 ⁷⁾	Ja	Ja ⁵⁾	Ja	-
LT-002	Nein	Nein	Nein	Langzeit-Sicherheits-/Wirksamkeitsstudie (Nachbeobachtung bis 15 Jahre) der Studien CL-303, CL-102, CL-302, CL-304, CL-306. Aktuell liegen nur Studienprotokolle vor. Es sind aktuell keine verwertbaren Daten verfügbar.
CL-102	Nein	Nein	Nein	Onasemnogen-Abeparvovec wurde in der Studie nicht zulassungskonform angewendet (intrathekale Applikation).
RG-001 ⁸⁾	Nein	Nein	Nein	Publikationstyp: kein Studienbericht (Ergebnisse, SAP) für Registerstudie verfügbar. Aktuell liegen nur Studienprotokolle vor.
Studien zu externen Kontrollen				
PNCR ⁹⁾	Ja	Ja	Ja	-
NeuroNEXT ⁹⁾	Ja	Ja	Ja	-
ENDEAR	Ja	Ja	Ja	-

Studiename (Nummer)	Studie mit Dossier vom pU ein- gereicht (Ja/Nein)	Studie vom pU als rele- vant für Ableitung des Zusatz- nutzens erachtet (Ja/Nein)	Studie relevant für die Nutzen- bewertung (Ja/Nein)	Ausschlussgrund
Finkel et al. (2014 b) ¹⁰⁾	Ja	Ja	Nein	Publikationstyp: Explorative, unkontrollierte, monozentrische, nicht- interventionelle Studie (N = 7 SMA Typ 1); in der Publikation finden sich keine detaillierten Angaben zum Patientenkollektiv; keine Effektschätzer berechnet, rein qualitativer Vergleich.

¹⁾ In Studie CL-101 wurde das Produktverfahren A verwendet. In den Studien CL-303, CL-302 und CL-304 wurde das Produktverfahren B verwendet [16].

²⁾ Zulassungsrelevante Studie gemäß EPAR [16].

³⁾ Gemäß EPAR konnte, basierend auf der Qualitätsbeurteilung, die Vergleichbarkeit von Produktverfahren A und B nicht festgestellt werden, weshalb Studie CL-101 als supportiv betrachtet wird [16].

⁴⁾ Studie CL-302 ist eine noch laufende, einarmige, multizentrische Phase-III-Studie. Die Rekrutierung war zum Datenschnitt vom Mai 2019 bereits abgeschlossen (N = 33). Das geschätzte Studienende ist für das 3. Quartal 2020 geplant (nähere Informationen siehe Tabelle 9).

⁵⁾ Die laufenden Studien werden vom pU teilweise berücksichtigt.

⁶⁾ Studie CL-304 ist eine noch laufende, einarmige, multizentrische Phase-III-Studie. In Studie CL-304 wurden 30 Personen eingeschlossen, davon haben 14 zwei SMN2-Genkopien (Kohorte 1, N = 14), 15 drei SMN2-Genkopien (Kohorte 2, N = 15) und ein/e Patient/in vier SMN2-Genkopien. Das Studienende ist für das 4. Quartal 2020 (2 SMN2-Genkopien) und 2. Quartal 2021 (3 SMN2-Genkopien) geplant (nähere Informationen siehe Tabelle 9).

⁷⁾ Studie LT-001 ist eine Langzeit-Sicherheitsstudie, in der Patientinnen und Patienten der Studie CL-101 die Möglichkeit hatten, weiterbeobachtet zu werden. Geplant ist ein Follow-up von bis zu 15 Jahren.

⁸⁾ Registerstudie (RESTORE).

⁹⁾ Für die Studien CL-101 und CL-303 wurde ein Report (RPT-806) zum Vergleich mit dem natürlichen Verlauf der PNCR- und NeuroNEXT-Studie eingereicht [7].

¹⁰⁾ Bei Finkel et al. (2014 b) handelt um eine Studie zur Bewertung von Untersuchungsmethoden zur Erfassung der Funktion der respiratorischen Muskulatur [20].

Abkürzungen: EPAR: European Public Assessment Report; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; SMA: Spinale Muskelatrophie; pU: pharmazeutischer Unternehmer; SAP: Statistischer Analyseplan; SMN: Survival Motor Neuron.

Die vom pU ausgewählten und herangezogenen externen Kontrollen – Pediatric Neuromuscular Clinical Research Network for SMA (PNCR), National Network for Excellence in Neuroscience Clinical Trials (NeuroNEXT) und ENDEAR – werden in der Nutzenbewertung ergänzend berücksichtigt. Die vorgeschlagenen Vergleichspopulationen bilden weitestgehend die von der Zulassung umfasste Patientenpopulation ab, allerdings besteht eine Vielzahl an methodischen Limitationen, die in den nachfolgenden Kapiteln adressiert werden. Nähere Informationen zu den indirekten Vergleichen finden sich in den Abschnitten 2.2.2., 2.3.5, 2.4.4 und 2.5.

Zur Nutzenbewertung für Onasemnogen-Abeparvovec wurden folgende Studien und Daten herangezogen

- Nutzendossier zu Onasemnogen-Abeparvovec [8].
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [16].
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie CL-101 [6].
- Studienbericht inkl. Studienprotokoll und SAP der Studie CL-303 [5].
- Interimsanalysen und Studienprotokoll der Studie CL-302 [2].
- Interimsanalysen und Studienprotokoll der Studie CL-304 [3].

- Interimsanalysen und Studienprotokoll der Studie LT-101 [4].
- Informationen zur PNCR- und NeuroNEXT-Studie [7,18,25,26].
- Informationen zur ENDEAR-Studie [12,19].

2.2 Aufbau und Design der eingeschlossenen Studien

Die Unterlagen zur Nutzenbewertung für Onasemnogen-Abeprarvovec basieren auf den Zulassungsstudien CL-303 (STR1VE-US) und CL-101 (START). Ergänzend dargestellt werden die Studien LT-001 (Langzeit-Sicherheitsstudie von CL-101), CL-302 (STR1VE-EU) und CL-304 (SPR1NT). Die Studien und die Intervention werden in den Tabelle 3 bis Tabelle 9 charakterisiert. Darüber hinaus werden vom pU Daten zum natürlichen Verlauf der Erkrankung (PNCR, NeuroNEXT und ENDEAR) vorgelegt, siehe. Tabelle 10.

2.2.1 Zulassungsstudien

Tabelle 3: Charakterisierung der Studie CL-303 (STR1VE-US)

Charakteristikum	Beschreibung
Design	<p>Die Studie CL-303 (STR1VE-US) ist eine einarmige, offene, multizentrische Phase-III-Studie. Das Ziel der Studie besteht darin, die Wirksamkeit und Sicherheit von Onasemnogen-Abeprarvovec als einmalige intravenöse Injektion bei symptomatischen und präsymptomatischen Kindern mit SMA Typ 1 und ein oder zwei SMN2-Genkopien zu untersuchen.</p> <p>Die Studie gliedert sich in 3 Studienperioden: Screening, Genersatztherapie und Nachbeobachtung. Während der Screeningperiode (Tag -30 bis Tag -2) erfolgte die Überprüfung der Eignung für den Einschluss in die Studie. Nach Bestätigung der Eignung schloss sich die Periode der stationären Genersatztherapie (Tag -1 bis Tag 3) an. An Tag -1 wurden die Patientinnen und Patienten zur Vorbehandlung in ein Krankenhaus eingewiesen und erhielten an Tag 1 eine einmalige intravenöse Infusion Onasemnogen-Abeprarvovec über etwa 30–60 Minuten. Daran anschließend verblieben die Patientinnen und Patienten 48 Stunden in stationärer Sicherheitsüberwachung. Die Entlassung erfolgte 48 Stunden nach der Genersatztherapie nach Einschätzung des ärztlichen Prüfpersonals. Während der ambulanten Nachbeobachtungszeit (Tag 4 bis zum Ende der Studie [Alter 18 Monate]) erfolgten regelmäßige Studienvisiten zur Wirksamkeits- und Sicherheitsbeurteilung.</p> <div style="text-align: center;">  <p>SCREENING PERIOD GENE REPLACEMENT THERAPY PERIOD (In-patient) FOLLOW-UP PERIOD^a (Outpatient)</p> <p>Days -30 to -2 Days -1, 1 to 3 Days 7, 14, 21, 30 Monthly to End of Study at 18 Months of Age (or ET)</p> </div> <p><i>Abbildung 1: Studienverlauf CL-303 (STR1VE-US) nach Studienperioden (Quelle: Studienbericht)</i></p> <p>Die Studie ist beendet (Dezember 2019). Im Anschluss an die Studie hatten die Patientinnen und Patienten die Möglichkeit in die Langzeitbeobachtungsstudie (LT-002) überzugehen (Nachbeobachtung bis zu 15 Jahre).</p>

Charakteristikum	Beschreibung
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • SMA Typ 1 mit biallelischer Mutation des SMN1-Gens (Deletion oder Punktmutation) und ein oder zwei SMN2-Genkopien (<i>Hinweis: Personen mit der genetischen Modifikation c.859G>C waren <u>nicht</u> ausgeschlossen</i>). • Alter < 6 Monate (< 180 Tage) zum Zeitpunkt des Gentransfers. • Schlucktest vor Therapiebeginn. • Impfstatus entsprechend den aktuellen Empfehlungen (umfasst auch Palivizumab-Prophylaxe (auch bekannt als Synagis) zur Verhinderung von RSV-Infektionen). • Eltern/Erziehungsberechtigte sind willens und fähig, den Prozess der informierten Einverständniserklärung abzuschließen und den Anforderungen der Studie und den geplanten Visiten nachzukommen. <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Vorherige geplante oder erwartete Skoliose-Operation in einem Alter < 18 Monaten. • Die Sättigung der Pulsoximetrie < 96 % beim Screening, während die Patientin / der Patient wach ist oder schläft, ohne zusätzlichen Sauerstoff oder Atemunterstützung, oder für Höhen > 1.000 m, Sauerstoffsättigung < 92 % wach oder schläft, ohne zusätzlichen Sauerstoff oder Atemunterstützung. Die Sättigung der Pulsoximetrie kann nach dem Screening auf < 96 % sinken, vorausgesetzt, dass die Sättigung nicht um ≥ 4 Prozentpunkte abnimmt. • Tracheostomie oder gegenwärtiger Gebrauch oder Bedarf an nicht-invasiver Beatmungsunterstützung im Durchschnitt ≥ 6 Stunden/Tag während der 7 Tage vor der Screening-Visite oder ≥ 6 Stunden/Tag im Durchschnitt während der Screeningperiode oder Bedarf an Beatmungsunterstützung im Wachzustand während der 7 Tage vor dem Screening oder zu irgendeinem Zeitpunkt während der Screeningperiode vor der Verabreichung. • Patientinnen/Patienten mit Anzeichen von Aspiration, nicht verdickte Flüssigkeiten zu vertragen, basierend auf einem formalen Schlucktest, der als Teil des Screenings durchgeführt wurde. Patientinnen/Patienten mit einer Gastrostomiekanüle, die den Schlucktest bestehen, dürfen in die Studie aufgenommen werden. • Patientinnen/Patienten, deren Gewicht im Verhältnis zum Alter unter dem dritten Perzentil liegt, basierend auf den Child Growth Standards der WHO. • Aktive virale Infektion (inklusive HIV oder Seropositivität auf Hepatitis B oder C oder bekannte Infektion mit Zika-Virus). • Schwerwiegende Erkrankung (nicht der Atemwege) mit Bedarf für systemische Behandlung und/oder Hospitalisierung innerhalb der 2 Wochen vor Screening. • Infektion der oberen oder unteren Atemwege, die medizinische Intervention, Aufmerksamkeit oder eine Zunahme der unterstützenden Pflegemaßnahmen erfordern, innerhalb der 4 Wochen vor Screening. • Schwere nicht-pulmonale/respiratorische Infektion (z. B. Pyelonephritis oder Meningitis) innerhalb von 4 Wochen vor Verabreichung der Genersatztherapie oder Begleiterkrankungen, die nach Meinung des ärztlichen Prüfpersonals unnötige Risiken für die Genersatztherapie mit sich bringen, wie z. B. Diabetes mellitus, Symptomatische Kardiomyopathie. • Bekannte Allergien oder Überempfindlichkeit gegenüber Prednisolon oder anderen Glukokortikosteroiden oder deren weiteren Inhaltsstoffen. • Gleichzeitige Anwendung eines oder mehrerer der folgenden Arzneimittel: Medikamente zur Behandlung von Myopathie oder Neuropathie, Diabetes mellitus, andauernde immunsuppressive Therapie, Plasmapherese, Immunmodulatoren wie Adalimumab oder immunsuppressive Therapie (z. B. Kortikosteroide, Cyclosporin, Tacrolimus, Methotrexat, Cyclophosphamid, IV-Immunglobulin, Rituximab) innerhalb der letzten drei Monate vor Gentherapie. • Anti-AAV9-Antikörpertiter > 1:50 bestimmt über ELISA. Bei einem Anti-AAV9-Antikörpertiter > 1:50, kann innerhalb von 30 Tagen während der Screening-Phase

Charakteristikum	Beschreibung
	<p>erneut getestet werden (wenn Titer \leq 1:50 kann Person in die Studie eingeschlossen werden).</p> <ul style="list-style-type: none"> • Klinisch pathologische Laborwerte: GGT, ALT, AST $>$ 3 x ULN, Bilirubin \geq 3,0 mg/dl, Kreatinin \geq 1,0 mg/dl, Hgb $<$ 8 oder $>$ 18 g/dl; WBC $>$ 20.000 pro ccm). • Kürzliche Teilnahme an einer klinischen Studie zur Behandlung der SMA (Ausnahme: Beobachtungsstudien oder nicht-interventionelle Studien) oder Behandlung mit einer Studienmedikation oder einem kommerziellen Produkt mit dem Ziel die SMA zu behandeln (z. B. Nusinersen, Valproinsäure) zu irgendeinem Zeitpunkt vor Screening. Orale Beta-Agonisten mussten mindestens 30 Tage vor Behandlung abgesetzt werden. • Erwartung von größeren chirurgischen Eingriffen während der Studienperiode (z. B. Operation an der Wirbelsäule oder Tracheostomie). • Gestationsalter bei der Geburt $<$ 35 Wochen.
Intervention und Zahl der Patienten	<p>Screening: N = 26 Studieneinschluss und -behandlung: N = 22 Alle Patientinnen und Patienten erhielten einmalig Onasemnogen-Abepravovec $1,1 \times 10^{14}$ vg/kg. Studie beendet: N = 19¹⁾ (Patientinnen und Patienten, die nicht aus der Studie ausgeschieden sind und eine Visite zu Monat 18 hatten).</p>
Ort und Zeitraum der Durchführung	<p>Die Studie wurde an 12 Zentren in den USA durchgeführt. Einschluss erste/r Patient/in: 24.10.2017 Letzte/r Patient/in abgeschlossen: 12.11.2019</p>
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>(Ko-)Primärer Endpunkt</p> <ul style="list-style-type: none"> • Erreichen des Entwicklungs-Meilensteins „Sitzen ohne Unterstützung für mindestens 30 Sekunden“ (definiert entsprechend der WHO Developmental Milestones) im Alter von 18 Monaten. • Überleben im Alter von 14 Monaten. Überleben ist definiert als das Nicht-Eintreten von (a) Tod oder (b) dauerhafter Beatmung, definiert als Tracheostomie oder der Notwendigkeit respiratorischer Unterstützung an \geq 16 Stunden pro Tag (durch nicht-invasive ventilatorische Unterstützung) für einen Zeitraum von \geq 14 aufeinanderfolgenden Tagen in der Abwesenheit einer akuten, reversiblen Erkrankung, aufgenommen perioperative Beatmung. <p>(Ko-)Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Gedeihfähigkeit definiert als das Erreichen folgender Kriterien im Alter von 18 Monaten: <ul style="list-style-type: none"> ○ Keine mechanische Unterstützung bei der Nahrungsaufnahme, z. B. durch eine Sonde. ○ Fähigkeit dünnflüssige Nahrung aufzunehmen, demonstriert über einen formalen Schlucktest. ○ Erhalt des Gewichts ($>$ 3. Perzentil für Alter und Geschlecht). • Frei von ventilatorischer Unterstützung in einem Alter von 18 Monaten. <p>Explorative Endpunkte</p> <ul style="list-style-type: none"> • Erreichen des Entwicklungs-Meilensteins „Kopfkontrolle“ (ohne Unterstützung). • Erreichen des Entwicklungs-Meilensteins „Rollen“. • Erreichen des Entwicklungs-Meilensteins „Sitzen ohne Unterstützung“ ($>$ 10 Sekunden, Definition der WHO Developmental Milestones). • Erreichen des Entwicklungs-Meilensteins „Krabbeln“. • Erreichen des Entwicklungs-Meilensteins „Hochziehen in eine stehende Position“. • Erreichen des Entwicklungs-Meilensteins „Stehen mit Unterstützung“. • Erreichen des Entwicklungs-Meilensteins „Freies Stehen“. • Erreichen des Entwicklungs-Meilensteins „Gehen mit Unterstützung“. • Erreichen des Entwicklungs-Meilensteins „Freies Gehen“.

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> • Motorische Funktion mittels fein- und grobmotorischer Domänen der Bayley-Skala. • Grobmotorik bestimmt über eine Verbesserung des CHOP-INTEND-Wertes. • Charakterisierung der physiologischen Wirksamkeit durch Verbesserung der Amplitude des peronealen Nervenkomplex-Motorischen Aktionspotentials (CMAP). • Alter, in dem „Freies Sitzen“ (30 Sekunden) erstmalig erreicht wurde. <p>Sicherheit</p> <ul style="list-style-type: none"> • Bewertung der Sicherheit der Behandlung mit Onasemnogen-Abeprarvovec bei Patientinnen und Patienten mit SMA Typ 1. • Bestimmung der Sicherheit von Onasemnogen-Abeprarvovec basierend auf der Entwicklung von inakzeptabler Toxizität definiert als das Auftreten jedweder unerwarteter behandlungsassoziierter UE CTCAE-Grad 3 oder höher.
Subgruppenanalysen	<p>A priori geplante Subgruppen (SAP, Version 2.0): Zwei Subgruppen basierend auf CHOP-INTEND-Wert zu Baseline:</p> <ul style="list-style-type: none"> • Patientinnen/Patienten mit biallelen Deletionsmutationen und 2 SMN2-Genkopien mit einem CHOP-INTEND-Baseline-Wert < 20 Punkte. • Patientinnen/Patienten mit biallelen Deletionsmutationen und 2 SMN2-Genkopien mit einem CHOP-INTEND-Baseline-Wert ≥ 20 Punkte <p>Zwei Subgruppen für Patientinnen und Patienten mit einem CHOP-INTEND-Baseline-Wert ≥ 20 Punkte:</p> <ul style="list-style-type: none"> • Symptomatische Patientinnen/Patienten zu Baseline. • Asymptomatische Patientinnen/Patienten zu Baseline. • Subgruppe abhängig vom Alter bei Dosierung (Alter an Tag 1 ≤ 4 Monate oder > 4 Monate), abhängig von der Anzahl der SMN2-Genkopien (1 Kopie, 2 Kopien) und abhängig vom Vorhandensein/Fehlen des Genmodifikators SMN2 c.859G>C.

¹⁾ N = 3 haben die Studie nicht beendet: Todesfall N = 1, UE N = 1, Rückzug der Einverständniserklärung N = 1.

Abkürzungen: ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CMAP: Compound Muscle Action Potential; CTCAE: Common Terminology Criteria for Adverse Events; ELISA: Enzyme-linked Immunosorbent Assay; GGT: Gamma-Glutamyl-Transferase; Hgb: Hämoglobin; RSV: Respiratory Syncytical Virus; SAP: Statistischer Analyseplan; SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron; UE: Unerwünschte/s Ereignis/se; ULN: Obere Grenze des Normwerts; vg: Vektorgenom/e; WBC: Weiße Blutkörperchen; WHO: World Health Organization.

Protokolländerungen

Insgesamt gibt es vier Versionen des Studienprotokolls (Originalversion 1.0 vom 29. März 2017) der Studie CL-303. Es liegen keine Informationen zu eingeschlossenen Patientinnen und Patienten je Protokollversion vor. Nachfolgend sind die für die Nutzenbewertung relevantesten Änderungen in Tabelle 4 aufgelistet.

Tabelle 4: Wesentliche Protokolländerungen Studie CL-303 (STR1VE-US)

Protokollversion	Wesentliche Änderung
Version 2.0 Amendment 1 15. Oktober 2017	Es wurde ergänzt, dass mindestens 15 Personen, die die ITT-Kriterien erfüllten, in die Studie eingeschlossen werden sollen.
Version 3.0 Amendment 2 21. Dezember 2017	Ein zusätzlicher Zeitpunkt für die CHOP-INTEND-Bewertung (Tag -1) wurde ergänzt.
Version 4.0 Amendment 3 4. Oktober 2018	<ul style="list-style-type: none"> • Aufnahme von zusätzlichem kardialen Monitoring zur Patientensicherheit. • Standardisierung von Infusionszeiten.

Abkürzungen: CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; ITT: Intention-to-Treat.

Tabelle 5: Charakterisierung der Studien CL-101 (START) und LT-001 (Langzeit-Extensionsstudie von Studie CL-101)

Charakteristikum	Beschreibung
Design	<p>Die Studie CL-101 (START) ist eine einarmige, offene, monozentrische Phase-I-Studie. Das primäre Ziel der Studie besteht darin, die Sicherheit verschiedener Dosierungen von Onasemnogen-Abeprarvovec als einmalige intravenöse Injektion bei Patientinnen und Patienten mit SMA Typ 1 und zwei SMN2-Genkopien zu untersuchen.</p> <p>Die Studie umfasst eine Screeningperiode, eine stationäre Periode zum Gentransfer sowie eine Nachbeobachtungsperiode. Während der Screeningperiode (Tag -30 bis Tag -2) erfolgte die Überprüfung der Eignung für den Einschluss in die Studie. Nach Bestätigung der Eignung schloss sich die Periode der stationären Gentransfers (Tag -1 bis Tag 2) an. An Tag -1 wurden die Patientinnen und Patienten zur Vorbehandlung in ein Krankenhaus eingewiesen und erhielten an Tag 0 eine einmalige intravenöse Infusion von Onasemnogen-Abeprarvovec über etwa 60 Minuten. Anschließend folgte eine zweijährige Nachbeobachtungsphase für die Bewertung der Sicherheit. Im Anschluss daran konnten die Patientinnen und Patienten in Studie LT-001 (Langzeit-Sicherheitsstudie) übergehen.</p>  <p>Note: The long-term follow-up portion of the study is being conducted under Protocol Number AVXS-101-LT-001.</p> <p><i>Abbildung 2: Studienverlauf CL-101 nach Studienperioden (Quelle: Studienbericht)</i></p> <p>Die Studie wurde mit potentiell vier verschiedenen Kohorten (sequenziell) geplant¹⁾:</p> <ul style="list-style-type: none"> • Kohorte 1 (niedrige Dosis): $6,7 \times 10^{13}$ vg/kg Onasemnogen-Abeprarvovec • Kohorte 2A (mittlere Dosis): $2,0 \times 10^{14}$ vg/kg Onasemnogen-Abeprarvovec • Kohorte 2B (mittlere Dosis): $2,0 \times 10^{14}$ vg/kg Onasemnogen-Abeprarvovec • Kohorte 3 (hohe Dosis): $3,3 \times 10^{14}$ vg/kg Onasemnogen-Abeprarvovec <p>In der Studie wurde keine Dosis-Eskalation bis zu Kohorte 3 durchgeführt. Kohorten 2A und 2B wurden zusammengefasst. Die Dosierung von $2,0 \times 10^{14}$ vg/kg (bzw. $1,1 \times 10^{14}$ vg/kg) entspricht der therapeutischen Dosierung und einer fachinformationskonformen Anwendung von Onasemnogen-Abeprarvovec.²⁾</p> <p>Die Studie CL-101 wurde nach einer geplanten Laufzeit von 24 Monaten beendet (14.12.2017). Im Anschluss konnten die Patientinnen und Patienten zur kontinuierlichen Sicherheitsüberwachung über einen Zeitraum von bis zu 15 Jahren in die Langzeit-Sicherheitsstudie LT-001 übergehen, die unter separatem Protokoll läuft. Es erfolgen jährliche Studienvisiten über einen Zeitraum von 5 Jahren und jährliche Telefonkontakte über einen Zeitraum von 10 Jahren. Das Ende der Studie LT-001 ist für 2033 vorgesehen.</p>
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • Alter ≤ 6 Monate zum Zeitpunkt des Gentransfers und SMA Typ 1 mit folgenden Merkmalen: <ul style="list-style-type: none"> ○ Biallelische Mutation des SMN1-Gens (Deletion oder Punktmutation) und zwei SMN2-Genkopien. ○ Auftreten der ersten Symptome in einem Alter zwischen 0 und 6 Monaten. ○ Hypotonie, festgestellt über eine klinische Untersuchung mit Hinauszögerung des Erlernens motorischer Fähigkeiten, schwacher Kontrolle des Kopfes, runde Haltung der Schulter und übersteigerte Beweglichkeit der Gelenke.

Charakteristikum	Beschreibung
	<p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Aktive virale Infektion (beinhaltet HIV oder seropositiv für Hepatitis B oder C). • Verwendung von invasiver Beatmung (Tracheostomie mit positivem Druck) oder < 95 % Sauerstoffsättigung als Ergebnis der Pulsoximetrie zum Screening. Patientinnen/Patienten können nach Entscheidung des Arztes oder des Studienteams nicht-invasive Beatmung (BiPAP) für weniger als 16 Stunden pro Tag erhalten. • Bestehende Erkrankung, die nach Einschätzung des ärztlichen Prüfpersonals ein unnötiges Risiko für den Gentransfer darstellt. • Gleichzeitige Behandlung mit einem der folgenden Arzneimittel: Arzneimittel zur Behandlung von Myopathie oder Neuropathie, Arzneimittel zur Behandlung von Diabetes mellitus, andauernde immunsuppressive Behandlung oder Behandlung mit Immunsuppressiva innerhalb der letzten 3 Monate vor Beginn der Studie (z. B. Kortikosteroide, Cyclosporine, Tacrolimus, Methotrexat, Cyclophosphamid, intravenöse Immunglobuline, Rituximab). • Patientinnen/Patienten mit Anti-AAV9-Antikörpertiter > 1:50 bestimmt über ELISA. • Abnorme Laborwerte, die als klinisch bedeutsam eingestuft werden (GGT > 3 x ULN; Bilirubin ≥ 3,0 mg/dl; Kreatinin ≥ 1,8 mg/dl; Hgb < 8 oder > 18 g/dl; WBC > 20.000 per mm³). • Patientinnen/Patienten mit Anzeichen von Aspiration, basierend auf einem Schlucktest, und die nicht bereit sind eine alternative Methode zur oralen Ernährung zu verwenden. • Patientinnen/Patienten mit genetischer Modifikation c.859G>C im SMN2-Gen (ggf. milder Phänotyp). <p><u>LT-001</u> Einschlussvoraussetzung war die vorherige Teilnahme an der Studie CL-101.</p>
Intervention und Zahl der Patienten	<p>Screening: N = 16 Studieneinschluss und -behandlung: N = 15</p> <p>a) Kohorte 1: 6,7 x 10¹³ vg/kg Onasemnogen-Abeparvovec: N = 3 b) Kohorte 2: 2,0 x 10¹⁴ vg/kg Onasemnogen-Abeparvovec: N = 12²⁾ c) Kohorte 3: 3,3 x 10¹⁴ vg/kg Onasemnogen-Abeparvovec: nicht durchgeführt</p> <p>Einschluss in Langzeit-Sicherheitsstudie LT-001: N = 13 (N = 3 aus Kohorte 1 [nicht zulassungskonform gemäß Fachinformation] und N = 10 aus Kohorte 2).</p>
Ort und Zeitraum der Durchführung	<p>Die Studie wurde an einem Zentrum in den USA durchgeführt. Einschluss erste/r Patient/in: 13.05.2014 Letzte/r Patient/in abgeschlossen: 14.12.2017</p>
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkt Sicherheit: Inzidenz von jeglichem nicht vorhergesehenen SUE von Grad 3 oder höher, dass mit der Behandlung in Zusammenhang steht, zu klinischen Symptomen führt und medizinische Behandlung erfordert (innerhalb von 2 Jahren).</p> <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Primärer Wirksamkeitsendpunkt: Überleben, definiert als Zeit von der Geburt bis a) Tod <u>oder</u> b) dauerhafte Beatmung, definiert als Notwendigkeit einer invasiven Beatmungsunterstützung ≥ 16 Stunden pro Tag kontinuierlich für ≥ 14 Tage, bei Abwesenheit einer akuten reversiblen Erkrankung, ausgenommen perioperative Beatmung. • Veränderung des CHOP-INTEND-Wertes im Vergleich zu Baseline. • Verbesserung der motorischen Funktion und Muskelkraft bestimmt über das Erreichen motorischer Meilensteine einschließlich, aber nicht beschränkt auf, der Fähigkeit selbstständig zu sitzen und sich ohne Hilfe umzudrehen.

Charakteristikum	Beschreibung
	<p>Explorative Endpunkte</p> <ul style="list-style-type: none"> • ACTIVE-mini-Bewertung der Bewegungsfähigkeit. • Funktionsmessungen mittels CHOP INTEND. • Bayley-Skala (kognitive, sprachliche und motorische Funktion (Grob- und Feinmotorik), bei Erreichen einer Punktzahl von ≥ 60 von 64). • Motoneuron-Funktion mittel CMAP und MUNE. • Pathologischer Status der Muskeln (EIM). • Alter, in dem signifikante motorische Meilensteine erreicht werden mittels standardisierten „Motor Milestone Development Survey“. • Überzeugende, nachweisbare, dokumentierte Beweise für die Wirksamkeit, die durch Veränderungen der funktionellen Fähigkeiten bestimmt werden, wie sie während der Videoaufnahmen während der Studienvsiste und/oder von Patientinnen/ Patienten / Eltern / Erziehungsberechtigten zu Hause erfasst werden. <p>Darüber hinaus wurden zu den genannten explorativen Endpunkten gemäß Studienprotokoll weitere explorative Endpunkte post hoc im Studienbericht erfasst:</p> <ul style="list-style-type: none"> • Gedeihfähigkeit. • Frei von ventilatorischer Unterstützung. • Erreichen von Schwellenwerten des CHOP INTEND. • Grobmotorische Fähigkeiten. • funktionelles unabhängiges Sitzen. • Bayley-Skala. • Verwendung von nicht-oraler Ernährungsunterstützung. • Hospitalisierung während der Studie. • Vorkommen und Gründe für invasive Beatmungsunterstützung.
Subgruppenanalysen	Keine

¹⁾ Im Studienprotokoll wird ein schrittweises Vorgehen beschrieben: Eine hohe Dosierung (Kohorte 3: $3,3 \times 10^{14}$ vg/kg Onasemnogen-Abeparvovec, verabreicht in eine periphere Vene als einmalige Infusion über einen venösen Katheter) war vorgesehen, wenn in Kohorte 2(B) eine Wirksamkeit beobachtet wurde, die jedoch für eine erfolgreiche Therapie als nicht ausreichend betrachtet wurde. Die letzte mögliche Dosis eskalation in Kohorte 3 wurde nicht durchgeführt. 12 Patientinnen und Patienten sind mit der mittleren (therapeutischen) Dosis behandelt worden.

²⁾ Die Angabe in der Fachinformation von $1,1 \times 10^{14}$ vg/kg entspricht der in Studie CL-101 verwendeten therapeutischen Dosierung. Die verschiedenen Konzentrationsangaben der Kohorten 2A und 2B beziehen sich ausschließlich auf eine unterschiedliche Messmethode (qPCR vs. ddPCR).

Abkürzungen: ACTIVE-mini: Ability Captured Through Interactive Video Evaluation-mini; BiPAP: Bi-level Positive Airway Pressure; CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorder; CMAP: Compound Muscle Action Potential; ddPCR: droplet digital Polymerase Chain Reaction; EIM: Electrical Impedance Myography; ELISA: Enzyme-linked Immunosorbent Assay; GGT: Gamma-Glutamyl-Transferase; Hgb: Hämoglobin; MUNE: Motor Unit Number Estimation; qPCR: quantitative Polymerase Chain Reaction; SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron; SUE: Schwerwiegende/s Unerwünschte/s Ereignis/se; ULN: Obere Grenze des Normwerts; vg: Vektorgenom/e; WBC: Weiße Blutkörperchen.

Protokolländerungen

Das Originalprotokoll (Version 6.2, 4. April 2014) wurde 9 Änderungen unterzogen. Vorherige Versionen waren Protokollentwürfe (Version 1: 5. Oktober 2012, Version 6.1: 13. März 2014) und entstanden vor Einschluss der Patientinnen und Patienten in die Studie. Die früheren Protokollversionen stellen eine Abstimmung zwischen dem Nationwide Children's Hospital und der FDA (U. S. Food and Drug Administration) dar, die Dokumente liegen dem pU nicht vor. Die nachfolgend genannten Protokolländerungen beinhalten die wesentlichen Änderungen sowie die Anzahl der eingeschlossenen Personen zur jeweiligen Protokollversion (siehe Tabelle 6). Die letzte Protokollversion 15.0 (21. November 2016) wurde aufgrund einer Rückmeldung der FDA zurückgenommen.

Danach sollten die in der Protokollversion 14.0 (21. April 2016) dargelegten Ziele und Endpunkte beibehalten werden.

Tabelle 6: Wesentliche Protokolländerungen Studie CL-101

Protokollversion	Protokolländerung
Originalversion 6.2 4. April 2014 Eingeschlossen: Kohorte 1: N = 1 Kohorte 2: N = 0 Gesamt: N = 1	-
Version 7.0 19. Mai 2014 Eingeschlossen: k. A.	Keine relevanten Änderungen.
Version 8.0 24. Juli 2014 Eingeschlossen: Kohorte 1: N = 2 Kohorte 2: N = 0 Gesamt: N = 3	Keine relevanten Änderungen.
Version 9.0 8. Oktober 2014 Eingeschlossen: Kohorte 1: N = 0 Kohorte 2: N = 2 Gesamt: N = 5	<ul style="list-style-type: none"> • Mittlere Dosis (Kohorte 2) mit $2,0 \times 10^{14}$ vg/kg wurde ergänzt. • Hohe Dosis mit $3,3 \times 10^{14}$ vg/kg wurde zu Kohorte 3 verschoben.
Version 10.0 20. Januar 2015 Eingeschlossen: Kohorte 1: N = 0 Kohorte 2: N = 1 Gesamt: N = 6	Änderung des sekundären Zielkriteriums: Zeit von der Geburt bis Tod oder Zeit bis zur dauerhaften Beatmung ≥ 16 Stunden kontinuierlich für ≥ 14 Tage, in Abwesenheit einer akuten reversiblen Erkrankung.
Version 11.0 12. März 2015 Eingeschlossen: Kohorte 1: N = 0 Kohorte 2: N = 3 Gesamt: N = 9	<ul style="list-style-type: none"> • Änderung des primären Zielkriteriums: Untersuchung der Sicherheit basierend auf der Beobachtung einer nicht-tolerierbaren Toxizität definiert als das unerwartete Auftreten von UE von Grad 3 oder höher, die mit der Behandlung in Verbindung gebracht wurden. • Ersetzen der UE-Schweregrad-Klassifizierung durch CTCAE-Klassifizierungen (v4.03) (Grade 1 bis 5). • Der Schweregrad der Lebertoxizität wird anhand der NIH-Richtlinie für die Einstufung des Schweregrades bei medikamenteninduzierten Leberschäden (Grade 1 bis 5) bewertet, einschließlich der Bewertung der Assoziation oder des Zusammenhangs mit dem Prüfpräparat.
Version 12.0 4. Mai 2015 Eingeschlossen: k. A.	<ul style="list-style-type: none"> • Aufnahme der Bayley-Skala (Grob- und Feinmotorik) als explorativer Endpunkt für Patientinnen und Patienten, die einen CHOP-INTEND-Wert von ≥ 60 erreichen. • Entfernung der Einstufung des Schweregrades der Lebertoxizität, die in das Protokoll v11.0 aufgenommen wurde.

Protokollversion	Protokolländerung
Version 13.0 10. Juni 2015 Eingeschlossen: Kohorte 1: N = 0 Kohorte 2: N = 2 Gesamt: N = 11	<ul style="list-style-type: none"> • Kohorte 3 entfernt und Kohorte 2 in Kohorte 2A umbenannt. • Erweiterung der Rekrutierung um 6 zusätzliche Patientinnen und Patienten (mittlere Dosis) nach Gentransfer der ersten 9 Patientinnen und Patienten.
Version 13.0 Addendum 24. Juni 2015 Eingeschlossen: Kohorte 1: N = 0 Kohorte 2: N = 4 Gesamt: N = 15	<ul style="list-style-type: none"> • Erweiterte Stichprobengröße für Kohorte 2B ($2,0 \times 10^{14}$ vg/kg) und Hinzufügen der Kohorte 3 ($3,3 \times 10^{14}$ vg/kg), jeweils um 3 Patientinnen und Patienten. • Das Alter der Patientinnen und Patienten am Tag der Infusion des Prüfpräparates wurde verringert auf 6 Monate oder jünger für Kohorte 2B und 3¹⁾.
Version 14.0 21. April 2016 Eingeschlossen: Rekrutierung abgeschlossen	<ul style="list-style-type: none"> • Hinzufügen weiterer sekundärer Zielkriterien: CHOP-INTEND-Veränderung gegenüber dem Baseline-Wert und Verbesserung der motorischen Funktion und Muskelkraft gezeigt über das Erreichen von motorischen Meilensteinen inklusive, aber nicht ausschließlich, dem freien Sitzen. • Einfügen der ersten Wirksamkeitsanalyse zu einem Zeitpunkt, an dem alle Patientinnen und Patienten ein Alter von 13,6 Monaten erreicht haben, und finales Sicherheits-Follow-up 24 Monate nach Gentransfer. • Videoaufnahmen von CHOP-INTEND- und Bayley-Skala-Bewertungen. • Wenn zwei aufeinanderfolgende CHOP-INTEND-Werte ≥ 62 sind, dann wird der Status der/des Patientin/Patienten überprüft, um festzustellen, ob weitere CHOP-INTEND-Beurteilungen notwendig sind. • Feinmotorik-, Grobmotorik- und Kognitionstests nach Bayley-Skala für Patientinnen und Patienten mit einer CHOP-INTEND-Punktzahl ≥ 60. • Während des ersten Jahres wurden die Subskalen der Bayley-Skala (Grob- und Feinmotorik) monatlich und die Subskala der Bayley-Skala (Kognition) alle 3 Monate ausgewertet. Im zweiten Jahr wurden alle Bayley-Skalen-Subtests alle 3 Monate ausgewertet, mit Ausnahme der Patientinnen und Patienten, die monatlich eine CHOP-INTEND-Auswertung hatten. Bei diesen Personen wurde das Bewertungsschema des ersten Jahres befolgt. • Beurteilung des Alters, in dem signifikante motorische Meilensteine erreicht wurden, unter Verwendung des Motor Milestone Development Survey und der Gross Motor Skills Checkliste.

¹⁾ Hinweis: Dieses Einschlusskriterium wurde überarbeitet, um die Aufnahme von Patientinnen und Patienten im Alter von 6 Monaten oder jünger zu ermöglichen (Protokolladdendum, 24.06.2015). Die ersten 9 Patientinnen und Patienten wurden unter der/den früheren Version/en des Protokolls rekrutiert, die eine Altersspanne von 9 Monaten oder jünger erlaubte.

Abkürzungen: CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CTCAE: Common Terminology Criteria for Adverse Events; k. A.: keine Angabe; NIH: National Institutes of Health; UE: Unerwünschte/s Ereignis/se; vg: Vektorgenom/e.

Tabelle 7: Charakterisierung der Intervention in der Studie CL-303

Intervention	Kontrolle
Onasemnogen-Abeparvovec ¹⁾ , <ul style="list-style-type: none"> • Einmalige intravenöse Applikation von 1,1 x 10¹⁴ vg/kg Onasemnogen-Abeparvovec über 30–60 Minuten (abhängig vom Gesamtvolumen). • Patientinnen/Patienten standen für 48 Stunden nach Verabreichen der Studienmedikation unter Beobachtung. 	Keine ²⁾
<p>Nicht erlaubte Begleitmedikation</p> <p>Die gleichzeitige Einnahme eines der folgenden Medikamente war verboten:</p> <ul style="list-style-type: none"> • Medikamente zur Behandlung von Diabetes mellitus, Myopathie oder Neuropathie. • Eine Therapie mit der Absicht SMA zu behandeln (z. B. Nusinersen, Valproinsäure). • Orale β-Agonisten wurden mindestens 30 Tage vor Dosierung der Gentherapie abgesetzt. <ul style="list-style-type: none"> ○ Inhalative β-Agonisten wurden möglicherweise zur Behandlung respiratorischer Komplikationen von SMA unter der Voraussetzung eingesetzt, dass solche Medikamente in klinisch angemessenen Mengen dosiert wurden. • Alle anderen Prüfpräparate während der Studie außer Onasemnogen-Abeparvovec. • Laufende immunsuppressive Therapie, Plasmapherese, Immunmodulatoren, wie Adalimumab, oder eine immunsuppressive Therapie innerhalb von 3 Monaten nach Beginn der Studie (z. B. Kortikosteroide, Cyclosporin, Tacrolimus, Methotrexat, Cyclophosphamid, IV Immunglobulin, Rituximab). <p>Die Verwendung von Kortikosteroiden nach Abschluss der Prednisolon-Anpassung (s. oben) war zulässig als Teil der routinemäßigen klinischen Behandlung. Der Einsatz von Kortikosteroiden unter solchen Umständen sollte in geeigneter Weise als Begleitmedikation dokumentiert und die beschleunigte Einnahme als UE dokumentiert werden.</p> <p>Wenn die Anwendung von Kortikosteroiden (abgesehen von inhalativen Kortikosteroiden bei Bronchospasmen) als Teil der Versorgung im Verlauf der Prednisolon-Anpassung in Betracht gezogen wurde, war diese medizinische Behandlung mit dem medizinischen Monitoring zu besprechen.</p> <p>Erlaubte Begleitmedikation</p> <p>Um die Immunreaktion der AAV-basierten Therapie zu verringern, erhielten alle Patientinnen und Patienten 24 Stunden vor Gentransfer und bis mindestens 30 Tage nach Infusion eine prophylaktische Gabe von Prednisolon ca. 1 mg/kg/Tag. Nach 30 Tagen konnte die Dosis Prednisolon angepasst werden für Personen, deren ALT-, AST-Werte und die T-Zell-Antwort $\leq 2 \times$ ULN für ALT und AST und < 100 SFC/106 PBMC war in Übereinstimmung mit folgendem Behandlungsschema:</p> <ul style="list-style-type: none"> • Bis mindestens 30 Tage nach der Infusion: 1 mg/kg/Tag. • Woche 5 und 6: 0,5 mg/kg/Tag. • Woche 7 und 8: 0,25 mg/kg/Tag. <p>Bei Werten $> 2 \times$ ULN für AST oder ALT oder einer T-Zell-Antwort von ≥ 100 SFC/106 PBMC nach 30 Tagen, wurde die Prednisolon-Dosis beibehalten bis die AST- und ALT-Werte unter den Schwellenwert gesunken waren. Sollte die T-Zell-Antwort nach 60 Tagen unverändert hoch sein, lag es im Ermessen des ärztlichen Prüfpersonals nach Einschätzung des Nutzen-Risiko-Verhältnisses die Prednisolon-Behandlung noch fortzuführen.</p> <p>Impfungen</p> <ul style="list-style-type: none"> • Patientinnen und Patienten wurden ermutigt während der gesamten Studie alle routinemäßig geplanten Impfungen zu befolgen, wie durch das CDC empfohlen. • Saisonale Impfungen, inklusive der Palivizumab-Prophylaxe (Synagis) zur Verhinderung des respiratorischen Synzytial-Virus wurden auch in Übereinstimmung mit der American Academy of Pädiatrie empfohlen. 	

¹⁾ Vor Verabreichen von Onasemnogen-Abeparvovec sind Ausgangs-Laboruntersuchungen erforderlich, einschließlich: AAV9-Antikörperbestimmung mit einem entsprechend validierten Test; Leberfunktion: ALT, AST und Gesamtbilirubin; Thrombozytenzahl; Troponin-I. Bei der Festlegung des Zeitpunktes der Therapie mit Onasemnogen-Abeparvovec sind die Notwendigkeit einer engmaschigen Überwachung der Leberfunktion, der Thrombozytenzahl und des Troponin-I-Spiegels nach der Verabreichung sowie die erforderliche begleitende Kortikosteroid-Behandlung zu berücksichtigen. [9]

²⁾ Einarmige Studie, keine Intervention.

Abkürzungen: AAV: Adeno-assoziiertes Virus; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; CDC: Center for Disease Control; PMBC: Peripheral Blood Mononuclear Cell; SFC: Spot Forming Cell; SMA: Spinale Muskelatrophie; UE: Unerwünschte/s Ereignis/se; ULN: Obere Grenze des Normwerts; vg: Vektorgenom/e.

Tabelle 8: Charakterisierung der Intervention in der Studie CL-101

Intervention	Kontrolle
Onasemnogen-Abeparvovec ¹⁾ , einmalige intravenöse Applikation von a) Kohorte 1: $6,7 \times 10^{13}$ vg/kg b) Kohorte 2 (2A und 2B): $2,0 \times 10^{14}$ vg/kg ²⁾ über ca. 60 Minuten (abhängig vom Gesamtvolumen).	Keine ³⁾
<p>Nicht erlaubte Begleitmedikation Die gleichzeitige Einnahme eines der folgenden Medikamente war verboten:</p> <ul style="list-style-type: none"> • Medikamente zur Behandlung von Diabetes mellitus, Myopathie oder Neuropathie. • Laufende immunsuppressive Therapie oder eine immunsuppressive Therapie innerhalb von 3 Monaten nach Beginn der Studie (z. B. Kortikosteroide, Cyclosporin, Tacrolimus, Methotrexat, Cyclophosphamid, IV Immunglobulin, Rituximab). <p>Erlaubte Begleitmedikation Um die Immunreaktion der AAV-basierten Therapie zu verringern, erhielten alle Patientinnen und Patienten 24 Stunden vor Gentransfer und bis mindestens 30 Tage nach Infusion eine prophylaktische Gabe von Prednisolon ca. 1 mg/kg/Tag. Nach 30 Tagen konnte die Dosis Prednisolon angepasst werden. Bei Werten $> 2 \times$ ULN für ALT und AST, wurde die Prednisolon-Dosis beibehalten, bis die Enzyme unter diesen Wert fielen, bei gleichzeitiger Überwachung der T-Zell-Antwort auf unter 100 SFC pro PBMC.</p> <p>Impfungen</p> <ul style="list-style-type: none"> • Patientinnen und Patienten wurden ermutigt während der gesamten Studie alle routinemäßig geplanten Impfungen zu befolgen, wie durch das CDC empfohlen. • Saisonale Impfungen, inklusive der Palivizumab-Prophylaxe (Synagis) zur Verhinderung des respiratorischen Synzytial-Virus wurden empfohlen. 	

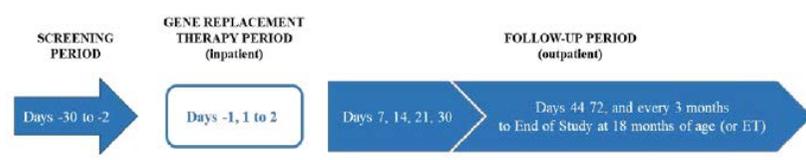
¹⁾ Vor Verabreichen von Onasemnogen-Abeparvovec sind Ausgangs-Laboruntersuchungen erforderlich, einschließlich: AAV9-Antikörpertestung mit einem entsprechend validierten Test; Leberfunktion: ALT, AST und Gesamtbilirubin; Thrombozytenzahl; Troponin-I. Bei der Festlegung des Zeitpunktes der Therapie mit Onasemnogen-Abeparvovec sind die Notwendigkeit einer engmaschigen Überwachung der Leberfunktion, der Thrombozytenzahl und des Troponin-I-Spiegels nach der Verabreichung sowie die erforderliche begleitende Kortikosteroid-Behandlung zu berücksichtigen. [9]

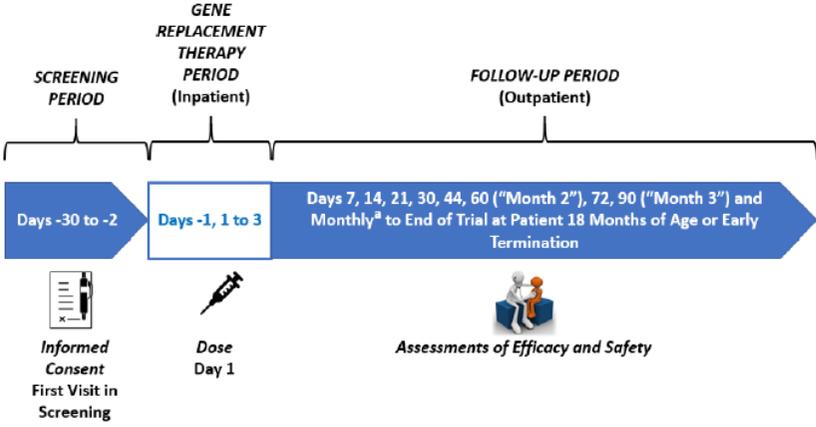
²⁾ Die beiden verschiedenen Konzentrationsangaben (2A und 2B) beziehen sich ausschließlich auf eine unterschiedliche Messmethode (qPCR vs. ddPCR), entsprechen jedoch der fachinformationskonformen Dosierung von Onasemnogen-Abeparvovec ($1,1 \times 10^{14}$ vg/kg).

³⁾ Einarmige Studie, keine Intervention.

Abkürzungen: AAV: Adeno-assoziiertes Virus; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; CDC: Center for Disease Control; ddPCR: droplet digital Polymerase Chain Reaction; PMBC: Peripheral Blood Mononuclear Cell; qPCR: quantitative Polymerase Chain Reaction; SFC: Spot Forming Cell; vg: Vektorgenom/e.

Tabelle 9: Charakterisierung der Studien CL-302 (STR1VE-EU) und CL-304

Charakteristikum	Beschreibung CL-302	Beschreibung CL-304
Design	<p>Studie CL-302 (STR1VE-EU) ist eine einarmige, offene, multizentrische Phase-III-Studie. Das Ziel der Studie besteht darin, die Wirksamkeit und Sicherheit von Onasemnogen-Abeparvovec als einmalige intravenöse Injektion bei symptomatischen und präsymptomatischen Kindern mit SMA Typ 1 mit ein oder zwei SMN2-Genkopien zu untersuchen.</p> <p>Die Studie gliedert sich in 3 Studienperioden: Screening, Gensersatztherapie und Nachbeobachtung. Während der Screeningperiode (Tage -30 bis -2) erfolgte die Überprüfung der Eignung für den Einschluss in die Studie. Nach Bestätigung der Eignung für die Studie schloss sich die Periode der stationären Gensersatztherapie (Tag -1 bis Tag 3) an. An Tag -1 wurden die Patientinnen und Patienten zur Vorbehandlung in ein Krankenhaus eingewiesen und erhielten an Tag 1 eine einmalige intravenöse Infusion Onasemnogen-Abeparvovec über etwa 60 Minuten. Daran anschließend verblieben die Patientinnen und Patienten 48 Stunden in stationärer Sicherheitsüberwachung. Die Entlassung erfolgte danach nach Einschätzung des ärztlichen Prüfpersonals. Während der ambulanten Nachbeobachtungszeit (Tag 4 bis zum Ende der Studie (Alter 18 Monate)) erfolgten regelmäßige Studienvisiten zur Wirksamkeits- und Sicherheitsbeurteilung.</p>	<p>Studie CL-304 ist eine einarmige, offene, multizentrische Phase-III-Studie. Das Ziel der Studie besteht darin, die Wirksamkeit und Sicherheit von Onasemnogen-Abeparvovec als einmalige intravenöse Injektion bei genetisch diagnostizierten und präsymptomatischen Kindern mit SMA Typ 1 mit biallelischer Mutation des SMN1-Gens mit zwei oder drei SMN2-Genkopien zu untersuchen.</p> <p>Die Studie gliedert sich in 3 Studienperioden: Screening, Gensersatztherapie und Nachbeobachtung. Während der Screeningperiode (Tage -30 bis -2) erfolgte die Überprüfung der Eignung für den Einschluss in die Studie. Nach Bestätigung der Eignung für die Studie schloss sich die Periode der stationären Gensersatztherapie (Tag -1 bis Tag 2) an. An Tag -1 wurden die Patientinnen und Patienten zur Vorbehandlung in ein Krankenhaus eingewiesen und erhielten an Tag 1 eine einmalige intravenöse Infusion Onasemnogen-Abeparvovec über etwa 60 Minuten. Daran anschließend verblieben die Patientinnen und Patienten 24 Stunden in stationärer Sicherheitsüberwachung. Die Entlassung erfolgte danach nach Einschätzung des ärztlichen Prüfpersonals. Während der ambulanten Nachbeobachtungszeit (Tag 3 bis zum Ende der Studie (Alter 18 (2 SMN2-Genkopien) oder 24 Monate (3 SMN2-Genkopien))) erfolgten regelmäßige Studienvisiten zur Wirksamkeits- und Sicherheitsbeurteilung.</p> <div data-bbox="1254 1021 2060 1181" style="text-align: center;">  </div> <p>Abbildung 4: Studienverlauf CL-304 – Kohorte mit 2 SMN2-Genkopien nach Studienperioden (Quelle: Studienbericht)</p>

Charakteristikum	Beschreibung CL-302	Beschreibung CL-304
	 <p>Abbildung 3: Studienverlauf CL-302 (STR1VE-EU) nach Studienperioden (Quelle: Studienbericht)</p> <p>Die Studie ist laufend (Datenschnitt letzte Interimsanalyse: Dezember 2019). Im Anschluss an die Studie haben die Patientinnen und Patienten die Möglichkeit an einer Langzeitbeobachtungsstudie (LT-002) nach separatem Protokoll teilzunehmen; dasselbe gilt für Patientinnen und Patienten, die die Studienteilnahme vorzeitig beenden.</p>	 <p>Abbildung 5: Studienverlauf CL-304 – Kohorte mit 3 SMN2-Genkopien nach Studienperioden (Quelle: Studienbericht)</p> <p>Die Studie ist laufend (Datenschnitt letzte Interimsanalyse: Dezember 2019). Im Anschluss an die Studie haben die Patientinnen und Patienten die Möglichkeit an einer Langzeitbeobachtungsstudie (LT-002) nach separatem Protokoll teilzunehmen; dasselbe gilt für Patientinnen und Patienten, die die Studienteilnahme vorzeitig beenden.</p>
<p>Population</p>	<p>Wesentliche Einschlusskriterien Wie Studie CL-303 (s. Tabelle 3)</p> <p>Wesentliche Ausschlusskriterien Wie Studie CL-303 (s. Tabelle 3)</p> <p><u>Abweichend:</u></p> <ul style="list-style-type: none"> Einsatz invasiver Beatmungsunterstützung (Tracheostomie mit positivem Druck) oder Pulsoximetrie < 95 % Sättigung beim Screening: 	<p>Wesentliche Einschlusskriterien – alle Patientinnen und Patienten</p> <ul style="list-style-type: none"> Alter ≤ 6 Wochen (≤ 42 Tage) zum Zeitpunkt des Gentransfers. Fähigkeit, dünne Flüssigkeiten zu vertragen, nachgewiesen durch einen formellen Schlucktest. CMAP ≥ 2 mV zu Baseline. Gestationsalter von 35 bis 42 Wochen. Impfstatus entsprechend den aktuellen Empfehlungen (umfasst auch Palivizumab-Prophylaxe (auch bekannt als Synagis) zur Verhinderung von RSV-Infektionen).

Charakteristikum	Beschreibung CL-302	Beschreibung CL-304
	<p>a) Sättigung der Pulsoximetrie kann nach dem Screening auf < 96 % sinken, vorausgesetzt, dass die Sättigung nicht um ≥ 4 Prozentpunkte abnimmt.</p> <p>b) Patientinnen und Patienten können nach Ermessen des ärztlichen Prüf- oder Studienpersonals für <12 Stunden pro Tag auf nicht-invasive Beatmungsunterstützung gesetzt werden.</p> <ul style="list-style-type: none"> • Verwendung oder Notwendigkeit einer nicht-invasiven Beatmungsunterstützung für ≥ 12 Stunden täglich in den 2 Wochen vor der Infusion. 	<ul style="list-style-type: none"> • Genetische Diagnose, durch ein akzeptables Neugeborenen- oder pränatales Screening-Testverfahren. <p><u>Einschlusskriterien Kohorte mit 2 SMN2-Genkopien</u> Personen mit präsymptomatischer SMA Typ 1, wie durch die folgenden Merkmale bestimmt: 2 SMN2-Genkopien.</p> <p><u>Einschlusskriterien Kohorte mit 3 SMN2-Genkopien</u> Personen mit präsymptomatischer SMA Typ 2, wie durch die folgenden Merkmale bestimmt: 3 SMN2-Genkopien.</p> <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Gewicht zur Screening-Visite < 2 kg. • Die Sättigung der Pulsoximetrie < 96 % beim Screening, während die/der Patient/in wach ist oder schläft, ohne zusätzlichen Sauerstoff oder Atemunterstützung, oder für Höhen > 1.000 m, Sauerstoffsättigung < 92 % wach oder schläft, ohne zusätzlichen Sauerstoff oder Atemunterstützung bei Screening. • Alle klinischen Anzeichen oder Symptome zur Screening-Visite oder unmittelbar vor der Infusion, die nach Meinung der Untersucherin / des Untersuchers, stark bezeichnend sind für SMA (z. B. Zungenfaszikulation, Hypotonie, Areflexie). • Tracheostomie oder gegenwärtige prophylaktische Anwendung oder Notwendigkeit einer nicht-invasiven Beatmungsunterstützung zu jeder Zeit und für jede Dauer vor dem Screening oder während der Screeningperiode. • Patientinnen/Patienten mit Anzeichen von Aspiration / Unvermögen nicht-verdickte Flüssigkeiten zu tolerieren, basierend auf einem formalen Schlucktest, der als Teil des Screenings durchgeführt wurde, oder Personen, die eine nicht-orale Ernährungsmethode erhalten. • Klinisch pathologische Laborwerte: GGT, ALT, AST oder Gesamt-Bilirubin > 2 x ULN, Kreatinin $\geq 1,0$ mg/dl, Hgb < 8 oder > 18 g/dl; WBC > 20.000 pro ccm); Patientinnen/Patienten mit einem erhöhten

Charakteris- tikum	Beschreibung CL-302	Beschreibung CL-304
		<p>Bilirubinspiegel aufgrund einer Neugeborenenengelbsucht sind nicht ausgeschlossen.</p> <ul style="list-style-type: none"> • Patientinnen/Patienten mit anderen klinisch signifikanten Anomalien in der Hämatologie oder klinisch-chemischen Parametern, die vom Prüfpersonal oder medizinischen Monitor bestimmt wurden. • Behandlung mit einem Prüf- oder Handelsprodukt, einschließlich Nusinersen, das zur Behandlung von SMA verabreicht wird. Dies schließt jegliche Vorgeschichte der Gentherapie, vorherige Antisense-Oligonukleotid-Behandlung oder Zelltransplantation ein. • Patientinnen/Patienten, deren Gewicht im Alter unter dem 3. Perzentil liegt, basierend auf den Kinderwachstumsstandards der WHO. • Biologische Mutter hat eine aktive virale Infektion (inklusive HIV oder Seropositivität auf Hepatitis B oder C oder bekannte Infektion mit Zika-Virus). • Schwerwiegende Erkrankung (nicht der Atemwege) mit Bedarf für systemische Behandlung und/oder Hospitalisierung innerhalb der 2 Wochen vor Screening. • Infektion der oberen oder unteren Atemwege, die medizinische Intervention, Aufmerksamkeit oder eine Zunahme der unterstützenden Pflegemaßnahmen erfordern, innerhalb der 4 Wochen vor Screening. • Schwere nicht-pulmonale/respiratorische Infektion (z. B. Pyelonephritis oder Meningitis) innerhalb von 4 Wochen vor Verabreichung der Gensatztherapie oder Begleiterkrankungen, die nach Meinung des ärztlichen Prüfpersonals unnötige Risiken für die Gensatztherapie mit sich bringen, wie erhebliche Nieren- oder Leberfunktionsstörungen, bekannte Anfallsleiden, Diabetes mellitus, Idiopathische Hypokalziurie, Symptomatische Kardiomyopathie. • Bekannte Allergien oder Überempfindlichkeit gegenüber Prednisolon oder anderen Glukokortikosteroiden oder deren weiteren Inhaltsstoffen. • Geplante oder erwartete größere chirurgische Eingriffe einschließlich Skoliose-Korrektur während der Studienbeurteilung.

Charakteristikum	Beschreibung CL-302	Beschreibung CL-304
		<ul style="list-style-type: none"> • Gleichzeitige Anwendung eines oder mehrerer der folgenden Arzneimittel: Medikamente zur Behandlung von Myopathie oder Neuropathie, Diabetes mellitus, andauernde immunsuppressive Therapie, Plasmapherese, Immunmodulatoren wie Adalimumab, oder immunsuppressive Therapie (z. B. Kortikosteroide, Cyclosporin, Tacrolimus, Methotrexat, Cyclophosphamid, IV-Immunglobulin, Rituximab) innerhalb der letzten 4 Wochen vor Gentherapie. • Anti-AAV9-Antikörpertiter > 1:50 bestimmt über ELISA. Bei Anti-AAV9-Antikörpertiter > 1:50 kann innerhalb von 30 Tagen während der Screeningphase erneut getestet werden (wenn Titer ≤ 1:50 kann Person in die Studie eingeschlossen werden, vorausgesetzt, sie ist zum Zeitpunkt der Verabreichung noch < 6 Wochen alt).
Intervention und Zahl der Patienten	Screening geplant: N = 30 Studieneinschluss und -behandlung: N = 33 (Rekrutierung abgeschlossen). Studie ist laufend, Datenschnitt der letzten Interimsanalyse: 31.12.2019.	Screening geplant: N = 26 <ul style="list-style-type: none"> • Kohorte mit zwei SMN2-Genkopien mindestens N = 14 • Kohorte mit drei SMN2-Genkopien mindestens N = 12 Studie ist laufend, Datenschnitt der letzten Interimsanalyse 31.12.2019.
Ort und Zeitraum der Durchführung	Die Studie wird an 12 bis 16 europäischen Zentren in folgenden Ländern durchgeführt: Belgien, Deutschland, Frankreich, Italien, Niederlande, Spanien, Schweden, Vereinigtes Königreich. Einschluss erste/r Patient/in: 2. Quartal 2018 Letzte/r Patient/in abgeschlossen: 3. Quartal 2020	Die Studie wird an 32 Zentren weltweit durchgeführt. Japan nimmt an dieser globalen Studie unter separatem Protokoll teil. Einschluss erste/r Patient/in: 1. Quartal 2018 Letzte/r Patient/in abgeschlossen: 2 SMN2-Genkopien: 4. Quartal 2020 3 SMN2-Genkopien: 3. Quartal 2021
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß	Primärer Endpunkt Erreichen des Entwicklungs-Meilensteins „Sitzen ohne Unterstützung für mindestens 10 Sekunden“ (definiert entsprechend der „WHO Developmental Milestones“) im Alter von 18 Monaten. Sekundäre Endpunkte Überleben im Alter von 14 Monaten. Überleben ist definiert als das Nicht-Eintreten von (a) Tod oder (b) dauerhafter Beatmung, definiert als Tracheostomie oder der Notwendigkeit respiratorischer Unterstützung an	Primäre Endpunkte <ul style="list-style-type: none"> • Bewertung der Sicherheit der Behandlung mit Onasemnogen-Abeparvovec anhand der Inzidenz von UE und/oder SUE. • Bewertung der Sicherheit der Behandlung mit Onasemnogen-Abeparvovec auf der Grundlage der Veränderung der klinischen Laborparameter gegenüber Baseline.

Charakteristikum	Beschreibung CL-302	Beschreibung CL-304
Studienprotokoll	<p>≥ 16 Stunden pro Tag (durch nicht-invasive ventilatorische Unterstützung) für einen Zeitraum von ≥ 14 aufeinanderfolgenden Tagen in der Abwesenheit einer akuten, reversiblen Erkrankung, ausgenommen perioperative Beatmung.</p> <p>Explorative Endpunkte</p> <ul style="list-style-type: none"> • Erreichen des Entwicklungs-Meilensteins „Kopfkontrolle“ (ohne Unterstützung). • Erreichen des Entwicklungs-Meilensteins „Rollen“. • Erreichen des Entwicklungs-Meilensteins „Sitzen ohne Unterstützung“ (≥ 30 Sekunden). • Erreichen des Entwicklungs-Meilensteins „Krabbeln“. • Erreichen des Entwicklungs-Meilensteins „Hochziehen in eine stehende Position“. • Erreichen des Entwicklungs-Meilensteins „Stehen mit Unterstützung“. • Erreichen des Entwicklungs-Meilensteins „Freies Stehen“. • Erreichen des Entwicklungs-Meilensteins „Gehen mit Unterstützung“. • Erreichen des Entwicklungs-Meilensteins „Freies Gehen“. • Motorische Funktion mittels fein- und grobmotorischen Subskala der Bayley-Skala (Version 3.0). • Grobmotorik bestimmt über eine Verbesserung des CHOP-INTEND-Wertes. • Frei von ventilatorischer Unterstützung (täglich) in einem Alter von 18 Monaten. • Gedeihfähigkeit definiert als das Erreichen folgender Kriterien: <ul style="list-style-type: none"> ○ Keine mechanische Unterstützung bei der Nahrungsaufnahme, z. B. durch eine Sonde. ○ Fähigkeit dünnflüssige Nahrung aufzunehmen, demonstriert über einen formalen Schlucktest. • Erhalt des Gewichts (> 3. Perzentil für Alter und Geschlecht). 	<p><u>Kohorte mit 2 SMN2-Genkopien</u> Erreichen des Entwicklungs-Meilensteins „Sitzen mit Unterstützung“ (≥ 30 Sekunden) im Alter bis zu 18 Monaten.</p> <p><u>Kohorte mit 3 SMN2-Genkopien</u> Erreichen des Entwicklungs-Meilensteins „Stehen ohne Unterstützung“ (≥ 3 Sekunden) im Alter bis zu 24 Monaten.</p> <p>Sekundäre Endpunkte</p> <p><u>Kohorte mit 2 SMN2-Genkopien</u></p> <ul style="list-style-type: none"> • Überleben im Alter von 14 Monaten. Überleben ist definiert als das Nicht-Eintreten von (a) Tod oder (b) dauerhafter Beatmung in der Abwesenheit einer akuten Erkrankung oder perioperativ. • Nachweis der Fähigkeit das Gewicht auf oder über dem 3. Perzentil zu halten, ohne Erfordernis einer nicht-oralen oder mechanischen Ernährungsunterstützung bis zum Alter von 18 Monaten. <p><u>Kohorte mit 3 SMN2-Genkopien</u> Erreichen des Entwicklungs-Meilensteins „Gehen ohne Unterstützung“ (mindestens fünf Schritte gehen und Koordination und Gleichgewicht zeigen) bei jeder Studiervisite bis zum Alter von 24 Monaten.</p> <p>Explorative Endpunkte</p> <p><u>Kohorte mit 2 SMN2-Genkopien</u></p> <ul style="list-style-type: none"> • Erreichen motorischer Meilensteine nach den Kriterien der WHO MGRS bis zum Alter von 18 Monaten: <ul style="list-style-type: none"> ○ Sitzen ohne Unterstützung. ○ Krabbeln (Hände und Knie). ○ Stehen mit Unterstützung. ○ Gehen mit Unterstützung. ○ Stehen ohne Unterstützung. ○ Gehen ohne Unterstützung. • Zeit bis zur ventilatorischen Unterstützung. • Notwendigkeit ventilatorischer Unterstützung im Alter von 18 Monaten.

Charakteristikum	Beschreibung CL-302	Beschreibung CL-304
	<p>Sicherheit</p> <ul style="list-style-type: none"> • Bewertung der Sicherheit der Behandlung mit Onasemnogen-Abeparvovec bei Patienten mit SMA Typ 1 • Bestimmung der Sicherheit von Onasemnogen-Abeparvovec basierend auf der Entwicklung von inakzeptabler Toxizität definiert als das Auftreten von jedweden unerwarteten behandlungsassoziierten UE von CTCAE-Grad 3 oder höher. 	<ul style="list-style-type: none"> • Überleben im Alter von 18 Monaten. Überleben ist definiert als das Nicht-Eintreten von (a) Tod oder (b) dauerhafter Beatmung in der Abwesenheit einer akuten Erkrankung oder perioperativ. • Anteil lebender Patientinnen/Patienten und ohne Tracheostomie im Alter von 18 Monaten. • Verbesserung um ≥ 15 Punkte im Vergleich zu Baseline der fein- und grobmotorischen Subskala (Rohwerte) der Bayley-Skala (Version 3.0) zu jeder Studienvsiste bis zum Alter von 18 Monaten. • Erreichen eines (skalierten) Wertes der grob- und feinmotorischen Subskalen der Bayley-Skala (Version 3.0) innerhalb von 1,5 SD eines Referenzstandards für die chronologische Entwicklung, bei jeder Studienvsiste bis zum Alter von 18 Monaten. • Erreichen eines CHOP-INTEND-Wertes ≥ 40 zu jeder Studienvsiste bis zum Alter von 18 Monaten. • Erreichen eines CHOP INTEND-Wertes > 50 zu jeder Studienvsiste bis zum Alter von 18 Monaten. • Erreichen eines CHOP INTEND-Wertes ≥ 58 zu jeder Studienvsiste bis zum Alter von 18 Monaten. • Erhalt erreichter Meilensteine bei Studienvsiten bis zum Alter von 18 Monaten, in Abwesenheit einer akuten Erkrankung oder perioperativ. <p><u>Kohorte mit 3 SMN2-Genkopien</u></p> <ul style="list-style-type: none"> • Erreichen motorischer Meilensteine nach den Kriterien der WHO MGRS bis zum Alter von 24 Monaten: <ul style="list-style-type: none"> ○ Stehen mit Unterstützung. ○ Gehen mit Unterstützung. • Zeit bis zur ventilatorischen Unterstützung. • Notwendigkeit ventilatorischer Unterstützung im Alter von 24 Monaten. • Überleben im Alter von 24 Monaten. Überleben ist definiert als das Nicht-Eintreten von (a) Tod oder (b) dauerhafter Beatmung in der Abwesenheit einer akuten Erkrankung oder perioperativ.

Charakteristikum	Beschreibung CL-302	Beschreibung CL-304
		<ul style="list-style-type: none"> • Verbesserung um ≥ 15 Punkte im Vergleich zu Baseline der fein- und grobmotorischen Subskala (Rohwerte) der Bayley-Skala (Version 3.0) zu jeder Studienvisite bis zum Alter von 24 Monaten. • Erreichen eines (skalierten) Wertes der grob- und feinmotorischen Subskalen der Bayley-Skala (Version 3.0) innerhalb von 1,5 SD eines Referenzstandards für die chronologische Entwicklung, bei jeder Studienvisite bis zum Alter von 24 Monaten. • Nachweis der Fähigkeit das Gewicht auf oder über dem 3. Perzentil zu halten, ohne Erfordernis einer nicht-oralen oder mechanischen Ernährungsunterstützung bis zum Alter von 24 Monaten. • Erhalt erreichter Meilensteine bei Studienvisiten bis zum Alter von 24 Monaten, in Abwesenheit einer akuten Erkrankung oder perioperativ.
Subgruppenanalysen	A priori geplante Subgruppenanalysen wie Studie CL-303 (s. Tabelle 3).	Separate Auswertung für Subgruppen nach SMN1-Punktmutationen oder der SMN2-Genmodifikator-Mutation (c.859G>C); zudem berücksichtigt: Anzahl SMN2-Genkopien, Dauer der Beobachtung je Kohorte sowie die relevanten primären, sekundären und explorativen Endpunkte. Entsprechend werden alle Wirksamkeits- und Sicherheitsendpunkte für jede der Subgruppen zusammengefasst.
Intervention	<ul style="list-style-type: none"> • Onasemnogen-Abeparvovec¹⁾, einmalige intravenöse Applikation von $1,1 \times 10^{14}$ vg/kg über 60 Minuten (abhängig vom Gesamtvolumen). • Patientinnen/Patienten standen für 48 Stunden nach Verabreichen der Studienmedikation unter Beobachtung. 	<ul style="list-style-type: none"> • Onasemnogen-Abeparvovec¹⁾, einmalige intravenöse Applikation von $1,1 \times 10^{14}$ vg/kg über 60 Minuten (abhängig vom Gesamtvolumen). • Patientinnen/Patienten standen für 24 Stunden nach Verabreichen der Studienmedikation unter Beobachtung.
Begleitmedikation	<p>Nicht erlaubte Begleitmedikation Wie Studie CL-303 (s. Tabelle 7).</p> <p>Erlaubte Begleitmedikation Wie Studie CL-303.</p> <p><u>Abweichend:</u> Um die Immunreaktion der AAV-basierten Therapie zu verringern, erhielten alle Patientinnen und Patienten prophylaktisch Prednisolon</p>	<p>Nicht erlaubte Begleitmedikation Wie Studie CL-303 (s. Tabelle 7).</p> <p>Erlaubte Begleitmedikation Wie Studie CL-303.</p> <p><u>Abweichend:</u> Um die Immunreaktion der AAV-basierten Therapie zu verringern, erhielten alle Patientinnen und Patienten prophylaktisch Prednisolon</p>

Charakteristikum	Beschreibung CL-302	Beschreibung CL-304
	ca. 2 mg/kg/Tag an den Tagen -1, 1, 2 und dann beginnend mit ca. 1 mg/kg/Tag an Tag 3 bis mindestens Tag 30 nach Infusion. Impfungen Wie Studie CL-303, wenn möglich Impfplan angepasst auf Prednisolon-Einnahme.	ca. 2 mg/kg/Tag an den Tagen -1, 1, 2, und dann beginnend mit ca. 1 mg/kg/Tag an Tag 3 bis mindestens Tag 30 nach Infusion. Impfungen Wie Studie CL-303, wenn möglich Impfplan angepasst auf Prednisolon-Einnahme.
Protokoll-änderungen	<u>Es liegen 6 Versionen des Studienprotokolls vor:</u> <ul style="list-style-type: none"> • Protokollversion 1.0 (17.02.2017 und 29.03.2017) • Protokollversion 2.0 (21.12.2017); Protokollversion 2.1 / Amendment 1.1 (09.01.2018); Protokollversion 2.2 / Amendment 1.2 (13.06.2018) • Protokollversion Version 3.0 / Amendment 2.0 (04.10.2018) • Protokollversion 4.0 (17.05.2019) • Protokollversion 5.0 / Amendment 4 (31.07.2019) • Protokollversion 6.0 / Amendment 5 (26.11.2019) 	<u>Es liegen 5 Versionen des Studienprotokolls vor:</u> <ul style="list-style-type: none"> • Protokollversion 1.0 (18.12.2017) • Protokollversion 2.0 (27.09.2018) • Protokollversion 3.0 (19.10.2018) • Protokollversion 4.0 / Amendment 3.0 (30.05.2019) • Protokollversion 5.0 / Amendment 4.0 (26.11.2019)

¹⁾ Vor Verabreichen von Onasemnogen-Abeparvovec sind Ausgangs-Laboruntersuchungen erforderlich, einschließlich AAV9-Antikörpertestung mit einem entsprechend validierten Test; Leberfunktion: ALT, AST und Gesamtbilirubin; Thrombozytenzahl; Troponin-I. Bei der Festlegung des Zeitpunktes der Therapie mit Onasemnogen-Abeparvovec sind die Notwendigkeit einer engmaschigen Überwachung der Leberfunktion, der Thrombozytenzahl und des Troponin-I-Spiegels nach der Verabreichung sowie die erforderliche begleitende Kortikosteroidbehandlung zu berücksichtigen. [9]

Abkürzungen: AAV: Adeno-assoziiertes Virus; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CMAP: Compound Muscle Action Potential; CTCAE: Common Terminology Criteria for Adverse Events; ELISA: Enzyme-linked Immunosorbent Assay; GGT: Gamma-Glutamyl-Transferase; Hgb: Hämoglobin; RSV: Respiratory Syncytial Virus; SD: Standardabweichung; SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron; (S)UE: (Schwerwiegende/s) Unerwünschte/s Ereignis/se; ULN: Obere Grenze des Normwerts; vg: Vektorgenom/e; WBC: Weiße Blutkörperchen; WHO: World Health Organization; WHO MGRS: World Health Organization Multicenter Growth Reference Study.

2.2.2 Externe Kontrollen

Mit dem Nutzendossier zu Onasemnogen-Abeprarvovec (Modul 4) legt der pU externe Kontrollstudien zum natürlichen Verlauf der SMA vor (PNCR, NeuroNEXT, ENDEAR) und stellt einen Vergleich mit den Zulassungsstudien CL-101 und CL-303 dar (siehe auch Anhang 1 bis 3).

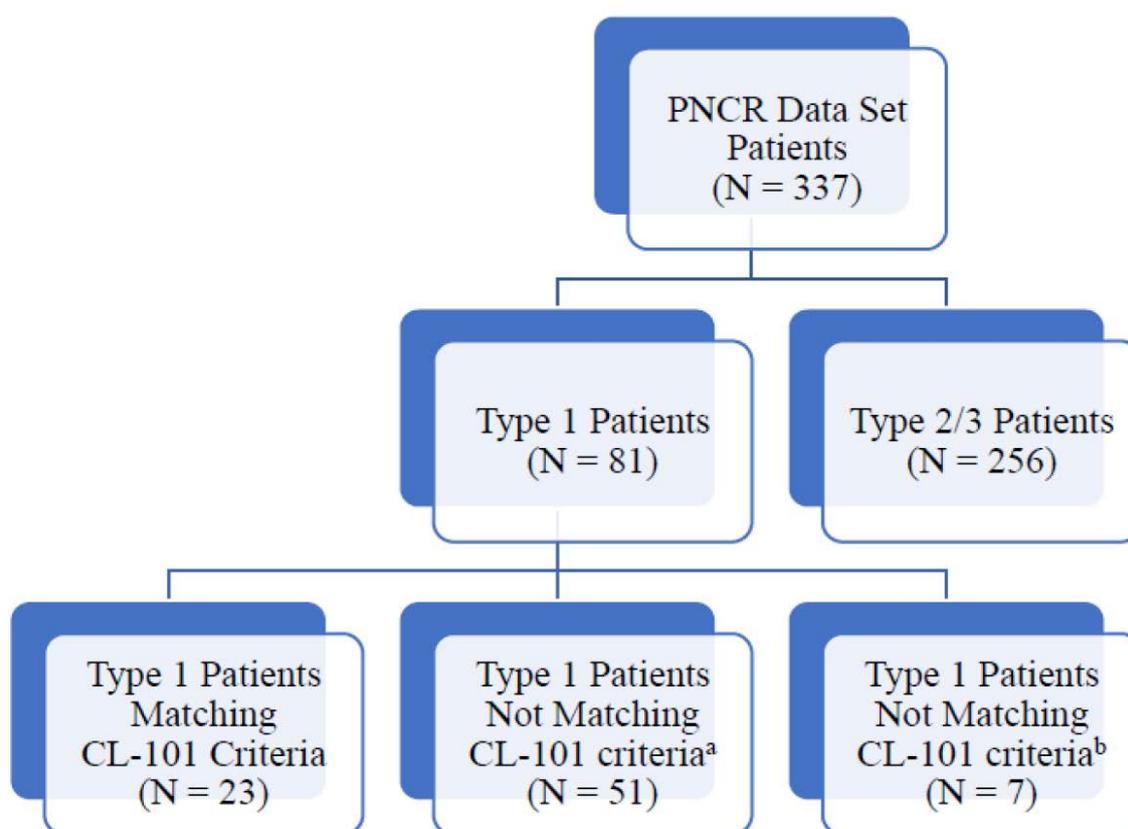
Insgesamt fehlen Angaben einer systematischen Recherche nach vergleichbaren Studien und der daraus resultierenden Identifikation der beiden Studien PNCR und NeuroNEXT. Im Rahmen des Nutzendossiers von Onasemnogen-Abeprarvovec (Modul 4) führte der pU zusätzlich eine systematische Literaturrecherche durch, um weitere relevante Studien für einen Vergleich zu identifizieren. Relevante Studien sind in diesem Kontext solche, die eine Patientenpopulation untersuchen, die mit den Studienpopulationen von CL-303 und CL-101 vergleichbar sind (2 SMN2-Genkopien, Erkrankungsbeginn in einem Alter ≤ 6 Monate). Im Ergebnis wurden neben den bereits genannten Studien (PNCR, NeuroNEXT) zusätzlich die Studien ENDEAR und Finkel et al. (2014 b) [20] identifiziert. Letztere wird, wie in Tabelle 2 bereits adressiert, nicht für die vorliegende Nutzenbewertung herangezogen. Die Studie ENDEAR ist eine randomisierte kontrollierte Studie zum Vergleich von Nusinersen mit einer Scheinintervention und gleichzeitiger Behandlung mit „Best Supportive Care“ (BSC) bei Patientinnen und Patienten mit infantiler SMA. Die Vergleichsgruppe (BSC) wurde als Kontrollgruppe für Onasemnogen-Abeprarvovec herangezogen.

Von Relevanz für die vorliegende Nutzenbewertung sind die Studien PNCR, NeuroNEXT und ENDEAR (BSC-Arm). Eine Darstellung der Studiencharakteristik findet sich in Tabelle 10. Angaben zur Charakterisierung der Patientenpopulation finden sich in Anhang 1 bis 3.

PNCR

Bei der Studie PNCR [18] handelt es sich um eine prospektive Beobachtungsstudie mit dem Ziel, den natürlichen Verlauf und die klinische Charakteristik von Patientinnen und Patienten mit SMA Typ 1 und 2 zu untersuchen. Die Studie wurde an drei medizinischen Zentren in den USA durchgeführt: Columbia University Medical Center (New York), University of Pennsylvania (The Children's Hospital of Philadelphia) und Harvard University (Boston Children's Hospital). Angaben, ob eine formale Ermittlung einer Studiengröße durchgeführt wurde, wird in der Publikation nicht adressiert. Der Einschluss des Patientenkollektivs in die Studie erfolgte zwischen Mai 2005 und April 2009. Die Beobachtungsdauer war bis zu 36 Monate geplant. Allen Patientinnen und Patienten wurde entsprechend ihrer Diagnose eine Teilnahme in die Studie angeboten. Dabei wurden sowohl Personen, die bereits in den Zentren in Behandlung waren, als auch neu diagnostizierte Personen mit SMA aufgenommen. Aus der Studie geht nicht hervor, welche und wie viele Patientendaten prospektiv und/oder retrospektiv erfasst wurden bzw. welche Daten retrospektiv in die Analyse gingen. Die Diagnose SMA und Klassifizierung des vorliegenden Typs wurden durch die Studienleitungen der Zentren vorgenommen. Zusätzlich erfolgte eine Unterscheidung der SMA-Typen hinsichtlich der Fähigkeit ohne Unterstützung (> 10 Sekunden) Sitzen zu können (Kriterien der World Health Organization Multicenter Growth Reference Study, WHO MGRS). Kinder, die zu keinem Zeitpunkt in der Lage waren diesen Meilenstein zu erreichen, wurden als SMA Typ 1 klassifiziert. Nur Personen mit homozygoter und biallelischer Deletion des SMN1-Gens und Alter bei Symptombeginn ≤ 6 Monate wurden für die Vergleichskohorte berücksichtigt. Ein genetischer Test sowie eine Untersuchung nach der genetischen Modifikation c.859G>C war in der PNCR-Studie nicht vorgesehen. Andere Ein- und Ausschlusskriterien wurden nicht definiert.

Die Datenbasis für die PNCR-Studie stellten 337 Personen mit SMA (Typen 1–3) [7]. Das Patientenkollektiv in der Studie von Finkel et al. (2014 a) [18] umfasste jedoch 79 Patientinnen und Patienten mit SMA Typ 1 und 2. Es bestehen Unklarheiten zum Patienteneinschluss der Studie. 81 Patientinnen und Patienten wiesen einen SMA Typ 1 auf (siehe Abbildung 1). In der Publikation von Finkel et al. (2014 a) wird jedoch adressiert, dass 54 Patientinnen und Patienten die o. g. Einschlusskriterien einer Diagnose SMA Typ 1 erfüllten. Aufgrund fehlender Angaben in den Studienunterlagen ist nicht ersichtlich, welche möglichen zusätzlichen Kriterien zu einer Abnahme der Population von 79 bzw. 81 auf 54 Personen führte. Auch liegen für diese Patientenkollektive keine soziodemographischen und krankheitsbezogenen Charakteristika in den vom pU referenzierten Publikationen vor.



^{a)} Patientinnen/Patienten mit SMN2-Genkopienzahl ungleich 2.

^{b)} Alter bei Symptombeginn > 6 Monate oder Alter bei Diagnose > 24 Monate.

Abbildung 6: Patientenfluss im PNCR-Datensatz [8]

In die Studie von Finkel et al. (2014 a) [18] wurden letztendlich 34 Patientinnen und Patienten mit SMA Typ 1 (18 mit Typ 1B und 16 mit Typ 1C) eingeschlossen. Für die Analyse wurden die Personen mit SMA Typ 1 in die Gruppen „recent“ vs „chronic“ (1B vs. 1C) unterteilt und als „recent“ klassifiziert, wenn die Aufnahme in die Studie innerhalb von 3 Monaten nach Diagnose erfolgte, und als „chronic“, wenn die Diagnose länger zurücklag. In der Publikation findet sich keine Darstellung der soziodemographischen und krankheitsbezogenen Charakteristika für die gesamte Population (N = 34), sondern nur differenziert nach den Gruppen 1B und 1 C (siehe Anhang 1).

Zum Vergleich des natürlichen Verlaufs der Erkrankung mit den Studienpopulationen von CL-101 und CL-303 wurde aus dem Datenset der PNCr-Studie eine Patientenpopulation mit den folgenden übereinstimmenden Charakteristika definiert:

- Alter zu Beginn der Symptome: ≤ 6 Monate,
- biallelische Deletion von SMN1 und
- zwei SMN2-Genkopien.

Insgesamt erfüllten 23 der 34 eingeschlossenen Patientinnen und Patienten die Kriterien (siehe Abbildung 6) und waren für einen Vergleich mit den Studien CL-101 und CL-303 vorgesehen. Von diesen Patientinnen und Patienten erreichten 50 % eine Nachbeobachtung von mindestens 12 Monaten. 16 Personen sind frühzeitig aus der Studie ausgeschieden: Tod ($n = 9$), Erkrankung ($n = 1$), fehlende Compliance ($n = 2$), Aufnahme in eine interventionelle Studie ($n = 2$), zeitliche Probleme der Betreuungsperson ($n = 1$) und Rückzug der Einverständniserklärung ($n = 1$).

NeuroNEXT (National Network for Excellence in Neuroscience Clinical Trials)

Die Studie NeuroNEXT ist eine prospektive, longitudinale Studie zum natürlichen Verlauf von Personen mit genetisch bestätigter SMA im Vergleich zu gesunden Kontrollen über einen Beobachtungszeitraum von 24 Monaten [25,26]. Der Kontrollarm der Studie (gesunde Patientinnen und Patienten) ist für die Nutzenbewertung nicht relevant.

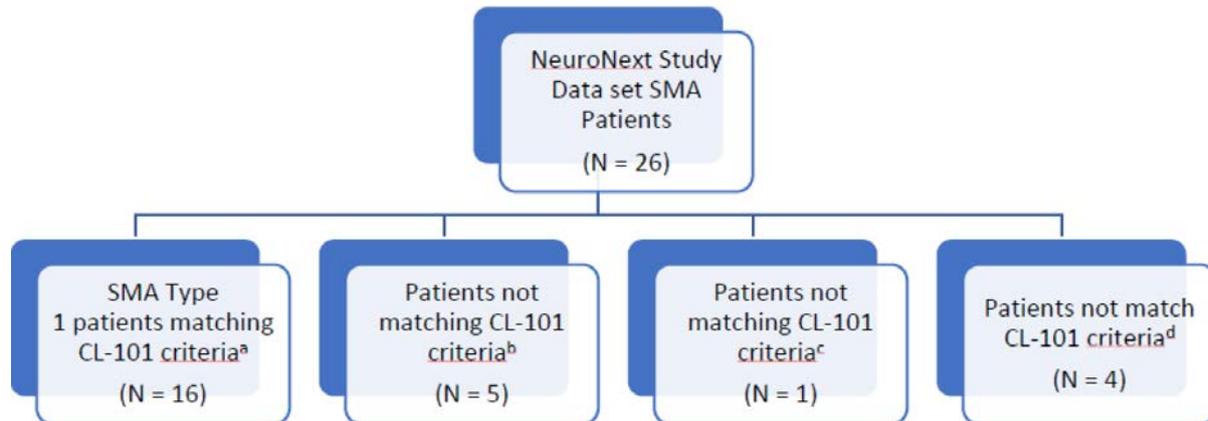
Eingeschlossen wurden Patientinnen und Patienten in einem Alter ≤ 6 Monate und die zwischen der 36. und 42. Schwangerschaftswoche geboren wurden. Die Diagnostik der SMA wurde entweder durch Studienärztinnen/Studienärzte oder der/dem behandelnden Neurologin/Neurologen vorgenommen. Vor Aufnahme in die Studie musste die SMA-Erkrankung durch einen Gentest bestätigt werden. Auch asymptomatische Personen, wenn sie vor der Einschreibung genetisch getestet wurden, waren für den Einschluss in die Studie vorgesehen. Von der Studie ausgeschlossen waren Personen, die mehr als 12 Stunden/Tag nicht-invasive ventilatorische Unterstützung benötigten (z. B. BiPAP), eine Komorbidität aufwiesen oder in eine andere therapeutische klinische Studie eingeschlossen waren. Personen wurden auch ausgeschlossen, wenn sie irgendeine Intervention erhielten, die das Potential hat, die SMN-Expression zu erhöhen (z. B. Valproinsäure).

Eine formale Ermittlung einer Studiengröße wurde nicht durchgeführt. Die Rekrutierung des Patientenkollektivs erfolgte an 14 Studienzentren in den USA zwischen November/Dezember 2012 bis September 2014. Studienvisiten erfolgten in einem Alter von 3 (wenn möglich), 6, 9, 12, 18 und 24 Monaten. In die Studie wurden 26 Personen mit SMA Typ 1 eingeschlossen (siehe Anhang 2).

Zum Vergleich des natürlichen Verlaufs der Erkrankung mit der Studienpopulation von CL-101 und CL-303 wurde dementsprechend aus dem Datenset der NeuroNEXT-Studie eine Patientenpopulation mit den folgenden Charakteristika definiert:

- Alter ≤ 6 Monate,
- biallelische Deletion von SMN1,
- zwei SMN2-Genkopien und
- Abwesenheit von c.859G>C.

16 Personen wiesen 2 SMN2-Genkopien auf (siehe Abbildung 7).



^{a)} 2 SMN2-Genkopien.

^{b)} 3 SMN2-Genkopien.

^{c)} 4 SMN2-Genkopien

^{d)} Keine Information zur SMN2-Genkopienzahl.

Abbildung 7: Patientenfluss im NeuroNEXT-Datensatz [8]

ENDEAR

Bei der Studie ENDEAR handelt es sich um eine randomisiert kontrollierte Phase-III-Zulassungsstudie von Nusinersen. Die Studie gliederte sich in eine Screeningphase von 21 Tagen, eine anschließende Behandlungsphase von 10 Monaten und eine Nachbeobachtung von 3 Monaten. Die Studie wurde in einem Zeitraum von August 2014 bis November 2016 in 31 Studienzentren in 13 Ländern durchgeführt. Nusinersen wurde gegenüber einer Scheinintervention mit gleichzeitiger Behandlung mit BSC verglichen. Die Vergleichsgruppe (BSC) wird als zusätzliche Kontrollgruppe für Onasemnogen-Abepravovec herangezogen. Die BSC-Gruppe in der Studie ENDEAR erhielt an denselben Studientagen wie die Nusinersen-Gruppe Scheininjektionen in Form eines Nadelstichs am unteren Rücken (ohne Lumbalpunktion). Die Injektionsstelle wurde analog zur Injektionsstelle der Nusinersen-Gruppe versorgt. Ausgewählte wesentliche Einschlusskriterien waren:

- Homozygote Deletion oder homozygote Mutation des SMN1-Gens,
- erste Symptome in einem Alter ≤ 6 Monate,
- zwei SMN2-Kopien und
- adäquate Nahrungs- und Flüssigkeitszufuhr (mit oder ohne Gastrostomie) nach Einschätzung des ärztlichen Prüfpersonals.

Die Intention-to-Treat (ITT)-Population, gemäß o. g. Einschlusskriterien, der Studie ENDEAR umfasste 41 Personen, die für den Vergleich mit den Zulassungsstudien CL-101 und CL-303 post hoc im Nutzendossier (Modul 4) verglichen wurden.

Vergleichbarkeit der Einschlusskriterien zwischen den externen Kontrollen und den Studien CL-101 und CL-303

Anhand a priori definierter Einschlusskriterien wurde in den externen Kontrollen jeweils eine Population identifiziert, die homogen zum Patientenkollektiv der Studien CL-101 und CL-303 ist. In den Studien CL-101 und CL-303 wurden Patientinnen und Patienten mit SMA Typ 1 und biallelischer Mutation des SMN1-Gens (Deletion oder Punktmutation) sowie bis zu 2 SMN2-Genkopien eingeschlossen. In die Studien PNCR, NeuroNEXT und ENDEAR wurden Patientinnen und

Patienten mit biallelischer oder homozygoter Deletion bzw. Mutation des SMN1-Gens und 2 SMN2-Genkopien eingeschlossen.

In Studie CL-101 wurde im Vergleich zu den Studien CL-303, PNCR und NeuroNEXT die genetische Modifikation c.859G>C im SMN2-Gen bei Patientinnen und Patienten berücksichtigt. In der ENDEAR-Studie finden sich keine entsprechenden Informationen zum Modifikator. Weitere Unterschiede zwischen den Studien zeigen sich für das Einschlusskriterium „Alter“:

- Alter bei Symptombeginn \leq 6 Monate: Studien CL-101, PNCR, ENDEAR
- Alter $<$ 6 Monate: Studie CL-303 (Zeitpunkt d. Gentransfer); \leq 6 Monate: Studie NeuroNEXT

Demnach waren in den Studien CL-101, PNCR und ENDEAR ausschließlich symptomatische Patientinnen und Patienten eingeschlossen, während an den Studien CL-303 und NeuroNEXT auch asymptomatische Patientinnen und Patienten teilnehmen konnten. In der Studie ENDEAR galt zudem im Vergleich zu den anderen externen Kontrollstudien eine adäquate Nahrungs- und Flüssigkeitszufuhr nach Einschätzung des ärztlichen Prüfpersonals als Einschlusskriterium.

In der Studie NeuroNEXT werden im Gegensatz zu den anderen externen Kontrollen Patientinnen und Patienten ausgeschlossen, die mehr als 12 Stunden/Tag nicht-invasive ventilatorische Unterstützung benötigten (z. B. BiPAP) oder eine bedeutsame Komorbidität aufwiesen. In den externen Kontrollen finden sich keine Angaben zur erlaubten und nicht-erlaubten Begleitmedikation. Lediglich in der Studie PNCR finden sich Angaben, dass 4 Personen begleitend Valporinsäure einnahmen.

Nähere Angaben zu den externen Kontrollen bezogen auf die Vergleichbarkeit von Endpunktoperationalisierung, Statistik und der Studienpopulationen finden sich in den jeweiligen Kapiteln.

Tabelle 10: Charakterisierung der externen Kontrollen PNCR, NeuroNEXT und ENDEAR (BSC)

Kriterium	PNCR	NeuroNEXT	ENDEAR (BSC)
Studien-design / Datenquelle	Prospektive Beobachtungsstudie zum natürlichen Verlauf von SMA Typ 1 und 2.	Prospektive, multizentrische, longitudinale Beobachtungsstudie zum natürlichen Verlauf der SMA und gesunden Kleinkindern.	Prospektive, multizentrische, randomisierte, doppelblinde Phase-III-Studie zur Wirksamkeit und Sicherheit von Nusinersen im Vergleich zu einer Scheinintervention bei Patientinnen und Patienten mit infantiler SMA.
Population	<p>Quellpopulation: unklar</p> <p><u>Wesentliche Einschlusskriterien</u></p> <ul style="list-style-type: none"> • Bestätigung Deletion SMN1-Gen, homozygote Deletion. • Bestimmung der SMN2-Genkopienzahl. • Allen Personen mit bestätigter SMA-Diagnose wurde eine Teilnahme angeboten. • Sowohl Personen, die bereits in den Zentren in Behandlung waren, als auch neu diagnostizierte Personen wurden aufgenommen. <p><u>Wesentliche Ausschlusskriterien</u></p> <p>Keine definiert</p>	<p>Quellpopulation N = 26</p> <p><u>Wesentliche Einschlusskriterien</u></p> <ul style="list-style-type: none"> • Alter ≤ 6 Monate und Geburt in der 36. bis 42. Schwangerschaftswoche • Genetische Diagnostik. • Bestimmung der SMN2-Genkopienzahl. <p><u>Wesentliche Ausschlusskriterien</u></p> <ul style="list-style-type: none"> • Personen wurden aus der Studie ausgeschlossen, wenn sie mehr als 12 Stunden/Tag nicht-invasive ventilatorische Unterstützung benötigten (z. B. BiPAP), eine bedeutsame Komorbidität aufwiesen oder in eine andere therapeutische klinische Studie eingeschlossen waren. • Personen wurden auch ausgeschlossen, wenn sie irgendeine Intervention erhielten, die das Potential hat, die SMN-Expression zu erhöhen (z. B. Valproinsäure). • Abwesenheit von c.859G>C 	<p>Quellpopulation N = 41 (ITT BSC-Arm)</p> <p><u>Wesentliche Einschlusskriterien</u></p> <ul style="list-style-type: none"> • Alter bei Symptombeginn ≤ 6 Monate Infantile SMA • Homozygote Deletion oder homozygote Mutation des SMN1-Gens. • 2 SMN2-Kopien. <p><u>Wesentliche Ausschlusskriterien</u></p> <ul style="list-style-type: none"> • Hypoxämie (O₂-Sättigung wach < 96 % oder schlafend < 96 % ohne unterstützende Beatmung) während des Screenings. • Vorliegen einer unbehandelten oder unzureichend behandelten aktiven Infektion, die eine systemische antivirale oder antimikrobielle Therapie zu jeglichem Zeitpunkt während des Screenings erfordert. • Behandlung von SMA mit einem nicht zugelassenen Wirkstoff (z. B. Salbutamol/ Salmeterol oral, Riluzol, Carnitin, Natriumphénylbutyrat, Valproat, Hydroxyharnstoff etc.), einem biologischen Wirkstoff oder einem Medizinprodukt innerhalb von 30 Tagen vor dem Screening oder zu jedem Zeitpunkt während der Studie; jegliche

Kriterium	PNCR	NeuroNEXT	ENDEAR (BSC)
	<u>Zielpopulation für den Vergleich: N = 23</u> <ul style="list-style-type: none"> • Alter bei Symptombeginn ≤ 6 Monate. • Biallelische Deletion von SMN1. • Zwei SMN2-Genkopien. • Daten zum Studieneinschluss liegen vor. 	<u>Zielpopulation für den Vergleich: N = 16</u> <ul style="list-style-type: none"> • Alter ≤ 6 Monate. • Biallelische Deletion von SMN1. • Zwei SMN2-Genkopien. • Abwesenheit von c.859G>C. 	Vorbehandlung mit Gentherapie, Antisense-Oligonukleotiden oder Zelltransplantation. <u>Zielpopulation für den Vergleich: N = 41 (ITT)</u> <ul style="list-style-type: none"> • Siehe Einschlusskriterien in die Studie. • Behandlung mit Scheinintervention unter BSC.
Diagnose	SMA-Diagnose erfolgte im jeweiligen Studienzentrum durch die Studienleitung. Die Bestätigung der gemeinsamen Deletion des SMN1-Exons 7/8 erfolgte durch PCR-Amplifikation und Restriktionsverdau der DNA unter Verwendung von Primern. Keine Untersuchung auf die SMN2-Gen-positive Modifikator-Mutation c.859G>C.22	SMA-Diagnose wurde durch ärztliches Studienpersonal oder örtliche Neurologinnen/Neurologen vorgenommen und vor Aufnahme in die Studie durch einen genetischen Test bestätigt. Das Fehlen einer Deletion/Mutation des SMN1-Gens wurde für jedes gesunde Kontrollkind bestätigt. Zusätzlich wurde die DNA von SMA-Patientinnen und -Patienten auf die SMN2-Gen-positive Modifikator-Mutation c.859G>C.22 untersucht.	Genetische Dokumentation der homozygoten Deletion, homozygoten Mutation oder Compound-Heterozygotie von 5q SMA.
Intervention und Zahl der Patientinnen und Patienten	Im Zeitraum zwischen Mai 2005 und April 2009 wurden Patientinnen und Patienten in die Studie aufgenommen. <u>Intervention</u> Um eine einheitliche Behandlung der Patientinnen und Patienten zu gewährleisten, wurden der Behandlung in allen teilnehmenden Studienzentren die Behandlungsstandards von Wang et al. (2007) zugrunde gelegt [29].	Die Rekrutierung der Studie startete im November 2012 und endete im September 2014. <u>Intervention</u> Es sind in den Publikationen keine detaillierten Angaben zur unterstützenden Behandlung enthalten.	Die Studie begann im Juli 2014 und dauerte bis November 2016 an. Für die Studie ENDEAR war eine Beobachtungszeit von 13 Monaten geplant, aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht in etwa 6,2 Monaten. <u>Intervention</u> Der BSC-Arm (Scheinintervention) erhielt an denselben Studientagen wie der Nusinersen-Arm Scheininjektionen in Form eines Nadelstichs am unteren Rücken (ohne Lumbalpunktion).

Kriterium	PNCR	NeuroNEXT	ENDEAR (BSC)
Ort und Zeitraum der Durchführung	<p>Die Studie wurde an 3 Studienzentren in den USA durchgeführt:</p> <ul style="list-style-type: none"> • Columbia University Medical Center (New York) • University of Pennsylvania (The Children's Hospital of Philadelphia) • Harvard University (Boston Children's Hospital) <p>Geplante Beobachtungsdauer: 36 Monate</p>	<p>Die Studie wurde an 14 Studienzentren in den USA durchgeführt.</p> <p>Die NeuroNEXT-Infrastruktur besteht aus 25 klinischen Zentren, die geographisch über die USA verteilt sind, einem zentralen Koordinierungszentrum am Massachusetts General Hospital und einem zentralen Datenkoordinierungszentrum an der University of Iowa.</p> <p>Geplante Beobachtungsdauer: 24 Monate</p>	<p>31 Zentren in 13 Ländern (Australien, Belgien, Deutschland, Frankreich, Italien, Japan, Kanada, Korea, Schweden, Spanien, Türkei, Vereinigtes Königreich, USA)</p> <p>August 2014 bis November 2016</p> <p>Geplante Beobachtungsdauer: 13 Monate</p>
Studienvisiten	<p>Studienvisiten wurden zu Baseline, nach Monat 2, 4, 6, 9 und 12, danach alle 6 Monate durchgeführt.</p>	<p>Studienvisiten erfolgten in einem Alter von 3 (falls zutreffend) 6, 9, 12, 18 und 24 Monaten.</p>	<p>Studienvisiten entsprechend der Endpunkterfassung [21].</p>
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte	<p><u>Primärer Endpunkt</u> Alter zum Zeitpunkt des Todes oder das Alter, in dem der kombinierte Endpunkt Tod oder dauerhafte Beatmung definiert als ventilatorische Unterstützung für mindestens 16 Stunden/Tag für mindestens 14 aufeinanderfolgende Tage in Abwesenheit einer akuten Erkrankung oder Operation erreicht wurde.</p> <p><u>Weitere Endpunkte</u></p> <ul style="list-style-type: none"> • Körperliche Untersuchungen (Gewicht, Größe, Kopf- und Brustumfang, Vitalparameter, motorische Funktion, Skoliose und Gelenkkontraktionen). • Zu jeder Studienvisite wurde eine Serumuntersuchung (metabolisches Panel) und eine Zählung der Blutkörperchen vorgenommen. Abweichungen der Laborparameter wurden über einen Vergleich mit Normwer- 	<p><u>Endpunkte</u></p> <ul style="list-style-type: none"> • Tod oder Intubation (dauerhafte Beatmung) • EIM-Test • CMAP • Blutabnahme • Motorische Funktionstests • TIMPSI • CHOP INTEND • AIMS • Biomarker: CMAP, EIM, SMN mRNA und SMA MAP 	<p><u>Primärer Endpunkt</u> Anteil der HINE-Responder (Subskala 2), Zeit bis zum Tod oder bis zur dauerhaften Beatmung (≥ 16 Stunden/Tag durchgehend für > 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses ODER Tracheostomie).</p> <p><u>Weitere Endpunkte</u></p> <ul style="list-style-type: none"> • CHOP-INTEND. • Prozentualer Anteil der Patientinnen und Patienten, die keine dauerhafte Beatmung benötigen. • Anteil der CMAP-Responder. • Zeit bis zum Tod oder bis zur dauerhaften Beatmung in den Subgruppen „Patienten unterhalb der mittleren Erkrankungsdauer“ bzw. „Patienten oberhalb der mittleren Erkrankungsdauer“.

Kriterium	PNCR	NeuroNEXT	ENDEAR (BSC)
	ten in der entsprechenden Altersgruppe bestimmt. <ul style="list-style-type: none"> • Motorische Funktion mittels CHOP INTEND. • Elektrophysiologische Untersuchungen umfassten CMAP und MUNE. 		

Abkürzungen: AIMS: Alberta Infant Motor Scale; BiPAP: Bi-level Positive Airway Pressure; BSC: Best Supportive Care; CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CMAP: Compound Muscle Action Potential; EIM: Electrical Impedance Myography; HINE: Hammersmith Infant Neurological Examination; ITT: Intention-to-Treat; MUNE: Motor Unit Number Estimate; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; PCR: Polymerase Chain Reaction; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron; TIMPSI: Test of Infant Motor Performance Screening Items.

2.3 Endpunkte

In diesem Kapitel werden die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Nutzendossier, die relevanten Studienberichte, einschließlich Studienprotokoll und SAP sowie die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt.

Aktuell liegen für die Studien CL-302, CL-304 und LT-001 (Langzeit-Sicherheitsstudie der Studie CL-101) ausschließlich Interimsanalysen vor, weshalb für die vorliegende Nutzenbewertung lediglich die Endpunkte „Mortalität“, „Tod oder dauerhafte Beatmung“ sowie die aggregierten Sicherheitsendpunkte dargestellt werden (siehe Tabelle 2).

Tabelle 11: Zusammenfassung der Endpunktbewertung

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU				Berücksichtigung in der Nutzenbewertung
		CL-303	CL-101 ¹⁾	CL-302 ²⁾	CL-304 ²⁾	
Mortalität ³⁾	Mortalität	ja	ja	ja	ja	ja
Tod oder dauerhafte Beatmung (Überleben) ⁴⁾		ja	ja	ja	ja	ja
Dauerhafte Beatmung	Morbidität	ja	ja	ja	ja	ja
Ventilatorische Unterstützung (Beatmungsfreiheit)	Morbidität	ja	ja	nein	nein	ja
Invasive Beatmung	Morbidität	ja	ja	ja	ja	ja
<u>Erreichen motorischer Meilensteine</u>	Morbidität	ja	ja	ja	ja	ja
Sitzen ohne Unterstützung (≥ 30 Sekunden)		ja	ja	ja	ja	ja
Sitzen ohne Unterstützung (> 10 Sekunden) ⁵⁾		ja	ja	ja	ja	ja
Sitzen ohne Unterstützung (≥ 5 Sekunden)		-	ja	-	-	ja
Kopfkontrolle (ohne Unterstützung)		ja	ja	ja	ja	ja
Rollen		ja	ja	ja	ja	ja
Krabbeln		ja	-	ja	ja	ja
Hochziehen in eine stehende Position		nein	-	nein	nein	ja
Stehen mit Unterstützung		ja	-	ja	ja	ja
Freies Stehen		ja	-	ja	ja	ja
Gehen mit Unterstützung		ja	-	ja	ja	ja
Freies Gehen		ja	-	ja	ja	ja
Bayley-Skala (Grob- und Feinmotorik)		Morbidität	ja	ja	ja	ja
CHOP INTEND (Grobmotorik)	Morbidität	ja	ja	ja	ja	ja

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU				Berücksichtigung in der Nutzenbewertung
		CL-303	CL-101 ¹⁾	CL-302 ²⁾	CL-304 ²⁾	
Bewertung d. Gedeihfähigkeit	Morbidität	ja	ja	ja	ja	nein
Nahrungsaufnahme	Morbidität	ja	ja	ja	ja	nein
Unerwünschte Ereignisse	Sicherheit	ja	ja	ja	ja	ja
Hospitalisierungen		ja	ja	nein	nein	ja

¹⁾ In der Langzeitbeobachtungsstudie der Studie CL-101, Studie LT-001, werden primär Sicherheitsendpunkte erfasst.

²⁾ In Tabelle 11 aufgeführt sind die in den Studien CL-302 und CL-304 erhobenen Endpunkte. Da aktuell ausschließlich Interimsanalysen vorliegen, werden lediglich die Endpunkte „Mortalität“, „Tod oder dauerhafte Beatmung“ sowie die aggregierten Sicherheitsendpunkte dargestellt.

³⁾ Todesfälle wurden im Rahmen der Sicherheitsbewertung erfasst.

⁴⁾ Kombiniertes Endpunkt.

⁵⁾ Meilenstein entsprechend der Definition der WHO Developmental Milestones.

Abkürzungen: CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; pU: pharmazeutischer Unternehmer; WHO: World Health Organization.

2.3.1 Mortalität

Mortalität (Todesfälle)

Der Endpunkt „Mortalität“ wird für die Nutzenbewertung berücksichtigt.

Operationalisierung

CL-303: Todesfälle wurden im Rahmen der Sicherheitsbewertung erfasst; darüber hinaus gibt der pU das Alter zum Todeszeitpunkt an.

CL-101: Todesfälle wurden im Rahmen der Sicherheitsbewertung erfasst; Zeit bis zum Eintreten eines Todes wird als Teilendpunkt des kombinierten primären Endpunktes „Überleben“ erhoben.

CL-302: Todesfälle wurden im Rahmen der Sicherheitsbewertung erfasst.

CL-304: Todesfälle wurden im Rahmen der Sicherheitsbewertung erfasst.

- Erhebungszeitpunkte:
CL-303, CL-101, CL-302 und CL-304: Über den gesamten Studienzeitraum.
- Auswertung:
CL-303, CL-101, CL-302 und CL-304: Auf Basis der Sicherheitspopulation.

Bewertung

Die Operationalisierung des Endpunktes ist nachvollziehbar.

Patientenrelevanz und Validität

Die Validität des Endpunktes wird als gegeben angesehen. 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“.

Tod oder dauerhafte Beatmung (Überleben)

Der Endpunkt „Tod oder dauerhafte Beatmung“ wird für die Nutzenbewertung berücksichtigt.

Operationalisierung

CL-303: Das Überleben ist definiert als das Vermeiden von entweder a) Tod oder b) dauerhafter Beatmung im Alter von 14 Monaten. Dauerhafte Beatmung ist definiert als die Notwendigkeit entweder einer Tracheostomie (invasiv) oder ventilatorischen Unterstützung von ≥ 16 Stunden pro Tag (nicht-invasiv) über einen Zeitraum von ≥ 14 aufeinanderfolgenden Tagen in Abwesenheit einer akuten reversiblen Erkrankung, ausgenommen perioperative Beatmung.

Die Daten zur Verwendung der nicht-invasiven Beatmungsunterstützung werden direkt aus dem Gerät jeder Patientin / jedes Patienten extrahiert; die Daten werden über eine vom Hersteller bereitgestellte Software von der Speicherkarte des Geräts heruntergeladen und als externer Datensatz übertragen. Berechnet wird die durchschnittliche tägliche Nutzung der nicht-invasiven Beatmungsunterstützung zu Studienbeginn, an den Tagen 7, 14, 21, 30 und danach monatlich. Wenn eine Person zu Beginn des Screenings kein für die Studie geeignetes Beatmungsgerät besaß und vor dem Screening eine nicht-invasive Beatmungsunterstützung verwendete, wurde die durchschnittliche tägliche Nutzung auf Grundlage des Berichts der Eltern berechnet und in der Ausgangsdokumentation festgehalten.

CL-101: Das Überleben ist definiert als die Zeit von der Geburt bis a) Tod oder b) dauerhafte Beatmung bis zu einem Alter von 13,6 und 20 Monaten. Dauerhafte Beatmung war in diesem Kontext definiert als die Notwendigkeit entweder einer invasiven Beatmung (Endotrachealtubus über Mund/Nase) oder respiratorischer Unterstützung ≥ 16 Stunden pro Tag (inklusive nicht-invasiver Beatmung) über einen Zeitraum von ≥ 14 aufeinanderfolgenden Tagen in Abwesenheit einer akuten reversiblen Erkrankung, ausgenommen perioperative Beatmung.

CL-302: Das Überleben wird definiert durch das Vermeiden des kombinierten Endpunktes im Alter von 14 Monaten; entweder a) Tod oder b) dauerhafte Beatmung, definiert durch Tracheostomie (invasiv) oder durch die Anforderung einer Beatmungsunterstützung von ≥ 16 Stunden pro Tag (nicht-invasiv) an ≥ 14 aufeinanderfolgenden Tagen in Abwesenheit einer akuten reversiblen Erkrankung, ausgenommen perioperative Beatmung.

CL-304: Das Überleben ist definiert als das Vermeiden von entweder a) Tod oder b) dauerhafter Beatmung über einen Zeitraum von ≥ 14 aufeinanderfolgenden Tagen in Abwesenheit einer akuten reversiblen Erkrankung bis zu einem Alter von 14, 18 und 24 Monaten. Als dauerhafte Beatmung gilt eine Tracheostomie (invasiv) oder die Anforderung einer Beatmungsunterstützung von ≥ 16 Stunden pro Tag (nicht-invasiv) für ≥ 14 aufeinanderfolgende Tage in Abwesenheit einer akuten reversiblen Erkrankung, mit Ausnahme der perioperativen Beatmung.

CL-303, CL-101, CL-302: Eine akute reversible Erkrankung ist definiert als jede andere Erkrankung mit Ausnahme der SMA, die zu einem erhöhten Bedarf an medizinischer Intervention führt (d. h. erhöhter Bedarf an externer Beatmungsunterstützung; Einsatz anderer begleitender Medikamente als Rettungsmaßnahme) und von der erwartet wird, dass sie nach einer konkreten Intervention (d. h. Operation, Antibiotika) oder der Einführung einer eskalierten unterstützenden Versorgung,

wie z. B. Krankenhausaufenthalt (d. h. bei Infektionen der oberen Atemwege, Spontanfraktur), reversibel ist oder sich verbessert.¹

Die perioperative Beatmung beinhaltet jede Beatmungsgeräteverwendung im Zusammenhang mit einem chirurgischen oder jeglichem anderen medizinischen Eingriff, für den die Patientin / der Patient Medikamente erhielt, die die Atemfunktion beeinträchtigen oder stören könnten.

- Erhebungszeitpunkte:
 - CL-303 und CL-101: Über den gesamten Studienzeitraum.
 - CL-302: Keine Angabe.
 - CL-304: Keine Angabe.
- Auswertung:
 - CL-303: Auf Basis der ITT-Population, finaler Datenschnitt 12. Dezember 2019.
 - CL-101: Auf Basis der ITT-Population, finaler Datenschnitt 14. Dezember 2017.
 - CL-302: Auf Basis der ITT-Population; aktuell liegen lediglich Interimsanalysen vor, letzter Datenschnitt Dezember 2019.
 - CL-304: Auf Basis der ITT-Population; aktuell liegen lediglich Interimsanalysen vor, letzter Datenschnitt Dezember 2019.

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar beschrieben. Es handelt sich um einen kombinierten Endpunkt.

Patientenrelevanz und Validität

Die Beatmungspflichtigkeit wird als patientenrelevant angesehen und gilt entsprechend als valider Endpunkt. Mortalität gilt ebenfalls als valider Endpunkt.

Gemäß § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustands, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Aufgrund von qualitativer und quantitativer Heterogenität sowie möglicher Verzerrungseffekte des kombinierten Endpunktes werden die Teilkomponenten des Endpunktes zusätzlich einzeln betrachtet.

¹ Die spezifische Dauer der dem Zustand vorausgehenden Intervention wird bei der Definition von "akut" nicht berücksichtigt. Das Datum der "konkreten Intervention" wird definiert als das Datum der Bereitstellung eines Verfahrens (d. h. einer Operation usw.) oder einer Medikation (d. h. Antibiotika), mit dem der Zustand geheilt oder wesentlich verbessert werden soll. Bei Zuständen, wie z. B. viralen Atemwegsinfektionen, für die eine unterstützende Behandlung vorgesehen ist, gilt das Datum des "endgültigen Eingriffs" als Datum des Krankenhausaufenthalts oder der wesentlichen Eskalation der Behandlung.

2.3.2 Morbidität

Ventilatorische Unterstützung (Beatmungsfreiheit)

Der Endpunkt „Ventilatorische Unterstützung (Beatmungsfreiheit)“ wird in der Nutzenbewertung ergänzend dargestellt.

Operationalisierung

CL-303: Unabhängigkeit von ventilatorischer Unterstützung ist definiert als der Anteil Patientinnen/Patienten, die von der unterstützenden Beatmung unabhängig waren, im Alter von 18 Monaten keine tägliche Beatmungsunterstützung benötigten, in Abwesenheit einer akuten reversiblen Erkrankung sowie perioperativer Beatmung, wie oben definiert durch Bewertung der tatsächlichen Nutzungsdaten des Beatmungsgerätes (BiPAP).

CL-101: Unabhängigkeit von ventilatorischer Unterstützung ist definiert als der Anteil Patientinnen/Patienten, die von einer unterstützenden Beatmung unabhängig waren und die an keinem der drei Zeitpunkte zur Wirksamkeitsanalyse eine tägliche Beatmungsunterstützung² benötigten, bei Abwesenheit einer akuten reversiblen Erkrankung und Ausschluss einer perioperativen Beatmung³.

- Erhebungszeitpunkte:
CL-303 und CL-101: Über den gesamten Studienzeitraum.
- Auswertung:
 - CL-303: Auf Basis der ITT-Population, finaler Datenschnitt: 12. Dezember 2019.
 - CL-101: Auf Basis des FAS (Full Analysis Set), finaler Datenschnitt: 14. Dezember 2017.

Bewertung

Die Operationalisierung des Endpunktes ist nachvollziehbar.

Patientenrelevanz und Validität

Der Endpunkt „Ventilatorische Unterstützung (Beatmungsfreiheit)“ wird als patientenrelevant (gemäß § 2 Satz 3 AM-NutzenV) und validiert angesehen. Da die Beatmungssituation der Patientinnen und Patienten bereits über den Teilendpunkt „Dauerhafte Beatmung“ des kombinierten Endpunktes „Tod oder dauerhafte Beatmung“ erfasst ist, wird der Endpunkt „Ventilatorische Unterstützung“ nur ergänzend berichtet.

² Zusatzgeräte, die in der Lungenheilkunde eingesetzt werden, wie z. B. ein Husten-Hilfsgerät, wurden nicht als Beatmungsunterstützung betrachtet.

³ Bei Personen, bei denen zum Zeitpunkt des Wirksamkeitsdatenschnitts eine akute reversible Erkrankung vorlag oder eine perioperative Beatmung erforderlich war, wurde die Unabhängigkeit der Beatmung definiert als die Tatsache, dass während des 2-Wochen-Intervalls vor dem Auftreten der akuten reversiblen Erkrankung oder der perioperativen Anwendung keine Beatmungsunterstützung erforderlich war.

Invasive Beatmung

Der Endpunkt „Invasive Beatmung“ wird in der Nutzenbewertung ergänzend dargestellt.

Operationalisierung

CL-303 und CL-101: Der Endpunkt wird in den Studien narrativ im Rahmen der dauerhaften Beatmung erfasst.

- Erhebungszeitpunkte: Über den gesamten Studienzeitraum.
- Auswertung:
 - CL-303: Auf Basis der ITT-Population, finaler Datenschnitt: 12. Dezember 2019.
 - CL-101: Auf Basis des FAS, finaler Datenschnitt: 14. Dezember 2017.

Validität und Patientenrelevanz

Der Endpunkt „Invasive Beatmung“ wird als patientenrelevant erachtet. Die Beatmung einer Patientin / eines Patienten stellt im Kontext der Erkrankung einen relevanten Faktor für die Schwere und den Verlauf der Erkrankung dar. Der Endpunkt wird als Teilkomponente des kombinierten Endpunktes „Tod oder dauerhafte Beatmung“ sowie des Endpunktes „Ventilatorische Unterstützung (Beatmungsfreiheit)“ erfasst. Die Ergebnisse werden ergänzend berichtet.

Erreichen motorischer Meilensteine

Der Endpunkt „Erreichen motorischer Meilensteine“ wird für die Nutzenbewertung berücksichtigt.

Operationalisierung

CL-303: Das Erreichen motorischer Meilensteine umfasst verschiedene motorische Entwicklungsstufen im Säuglings- und Kleinkindalter, die zur Beurteilung der Wirksamkeit der Studienmedikation untersucht wurden. Das unabhängige Sitzen von mindestens 30 Sekunden in einem Alter von 18 Monaten war einer der zwei primären Wirksamkeitsendpunkte der Studie CL-303. Der Endpunkt wurde erfasst mittels Bayley-Skala für die Entwicklung von Säuglingen und Kleinkindern (Version 3) [11], bestätigt durch Videoaufnahmen. Eine Patientin / ein Patient musste mindestens 30 Sekunden lang aufrecht sitzen (Bayley-Skala – Grobmotorik, Item #26).

Die Erhebung weiterer Meilensteine erfolgt für Responder, definiert als Patientin/Patient, die/der das Erreichen eines Meilensteins bei jeder Studienvisite nach Gentransfer bis zu und einschließlich der letzten Studienvisite im Alter von 18 Monaten vorweist. Für den Meilenstein „Sitzen mit Unterstützung“ ist keine Videobestätigung durchgeführt worden.

Weitere motorische Meilensteine:

- Sitzen ohne Unterstützung für > 10 Sekunden (gemäß WHO-Definition): Das Kind sitzt aufrecht mit erhobenem Kopf für mindestens 10 Sekunden. Das Kind benutzt weder Arme noch Hände, um sich zu stabilisieren oder die Position zu stützen.
- Sitzen mit Unterstützung (Bayley-Skala – Grobmotorik, Item #19, ohne Videoaufzeichnung)
- Kopfkontrolle ohne Unterstützung (Bayley-Skala – Grobmotorik, Item #4): Das Kind hält den Kopf aufrecht ohne Unterstützung für mindestens 3 Sekunden.
- Rollen (Bayley-Skala – Grobmotorik, Item #20): Das Kind rollt sich aus der Rückenlage sowohl auf die rechte als auch auf die linke Seite.

- Krabbeln (Bayley-Skala – Grobmotorik, Item #34): Das Kind bewegt sich vorwärts für mindestens 5 Fuß (= 152,4 cm) durch Krabbeln auf Händen und Knien.
- Hochziehen in eine stehende Position (Bayley-Skala – Grobmotorik, Item #35): Das Kind zieht sich selbst in eine stehende Position unter Verwendung eines Stuhls oder eines anderen geeigneten Gegenstands zur Unterstützung.
- Stehen mit Unterstützung / ohne Unterstützung (Bayley-Skala – Grobmotorik, Items #33 / #40): Das Kind trägt das eigene Gewicht für mindestens 2 Sekunden. / Das Kind steht alleine für mindestens 3 Sekunden, nachdem dessen Hand losgelassen wurde.
- Gehen mit Unterstützung / ohne Unterstützung (Bayley-Skala – Grobmotorik, Items #37 / #43): Das Kind geht (mit Unterstützung), indem es koordinierte, alternierende Schrittbewegungen macht. / Das Kind geht mindestens 5 Schritte ohne Unterstützung, zeigt Koordination und Gleichgewicht.

Die Entwicklungsmeilensteine wurden unter Verwendung relevanter Item-Definitionen der Bayley-Skala (Version 3.0) beurteilt. Das Erreichen jedes Entwicklungsmeilensteins wurde durch das ärztliche Prüfpersonal bestimmt und durch die weitere Bewertung eines Videos bestätigt. Bei jeder monatlichen Studienvisite wurden die zu erreichenden Entwicklungsmeilensteine bestimmt. Zu Baseline wurde vom ärztlichen Prüfpersonal die Ausgangssituation bewertet. Die Bewertung bezog sich auf alle Meilensteine entsprechend der erwarteten altersbedingten Funktion des Kleinkindes. Items der Bayley-Skala, die unter der erwarteten Funktion lagen und zu Baseline nicht erreicht wurden, sind bei den nachfolgenden Studienvisiten bis zur erfolgreichen Durchführung überprüft worden. Die Meilensteine „Unabhängiges Sitzen“, „Unabhängiges Sitzen für mindestens 5 Sekunden“ (Bayley-Skala – Grobfunktion, Item #22) und „Unabhängiges Sitzen für mindestens 30 Sekunden“ (Bayley-Skala – Grobfunktion, Item #26) wurden bei jeder Studienvisite bis zum Erreichen des Meilensteins und unabhängig von der Bayley-Skala bewertet.

Die Bayley-Skala ist eine standardisierte, normenbezogene Säuglingsbeurteilung von Entwicklungsfunktionen in 5 Bereichen: kognitives, sprachliches, motorisches, sozial-emotionales und adaptives Verhalten. Die Erhebung erfolgte durch eine qualifizierte Physiotherapeutin / einen qualifizierten Physiotherapeuten. Die grob- und feinmotorischen Tests der Bayley-Skala wurden mit der Videokamera aufgenommen. Die Videos wurden hinsichtlich des Nachweises für das Erreichen des Meilensteins von einer unabhängigen Person beurteilt. Dabei wurde der früheste Zeitpunkt für das Ereignis auf den Videos ermittelt. Darüber hinaus wurde den Eltern/Erziehungsberechtigten gestattet, zusätzliche Heimvideos einzureichen, die den Nachweis des Erreichens von Entwicklungsmeilensteinen zu einem beliebigen Zeitpunkt während der Studie belegen. Die Bewertung des Meilensteins „Unabhängiges Sitzen > 10 Sekunden“ erfolgt nach WHO-Definition unter Verwendung folgender Kriterien [30]:

- a) Der Kopf des Kindes ist aufgerichtet.
- b) Das Kind benutzt die Arme oder Hände nicht, um den Körper auszubalancieren oder eine Stützposition einzunehmen.
- c) Das Kind sitzt mindestens 10 Sekunden lang aufrecht.

CL-101: Die Verbesserung der motorischen Funktion und der Muskelkraft, die durch das Erreichen wichtiger Entwicklungsmeilensteine bestimmt wird – einschließlich, aber nicht beschränkt auf die Fähigkeit, allein zu sitzen und sich ohne Hilfe umzudrehen – war einer der sekundären Endpunkte der Studie. Das funktionell unabhängige Sitzen wurde definiert als die Fähigkeit, eine Sitzposition unabhängig für mindestens 30 Sekunden aufrechtzuerhalten, was per Videoevaluierung durch eine zentrale Begutachtung bestätigt wurde.

Das Erreichen motorischer Meilensteine umfasst verschiedene motorische Entwicklungsstufen im Säuglings- und Kleinkindalter, die zur Beurteilung der Wirksamkeit der Studienmedikation untersucht wurden. In Studie CL-101 wurden motorische Meilensteine mit Hilfe einer Standardübersicht zur Entwicklung motorischer Meilensteine (Motor Milestone Development Survey) (Tabelle 12) und der Checkliste für grobe motorische Fertigkeiten (GMS-Checkliste; Gross Motor Skills) bewertet. Die motorischen Meilensteine sowie die Entwicklung wurden vom ärztlichen Prüfpersonal ermittelt und dokumentiert; die Daten für die GMS-Checkliste wurden von der Physiotherapeutin / dem Physiotherapeuten durch eine körperliche Untersuchung erhoben und dokumentiert. Darüber hinaus wurden Videos von Entwicklungsmeilensteinen verwendet, die entweder bei geplanten Besuchen aufgenommen oder von den Eltern/Erziehungsberechtigten zur Verfügung gestellt wurden. Die Videos bestätigen das Erreichen der Meilensteine; die Bewertung der Videos erfolgte durch eine zentrale unabhängige Begutachtung. Neben den oben beschriebenen waren vier weitere motorische Meilensteine von besonderem Interesse der externen Begutachtung.

- **Kopfkontrolle:** Kind hält den Kopf mindestens 3 Sekunden lang ohne Unterstützung aufrecht (Bayley-Skala⁴ – Grobmotorik, Item #4).
- **Rollen:** Kind dreht sich aus der Rückenlage nach rechts und links (Bayley-Skala – Grobmotorik, Item #20).
- **Sitzen mit Unterstützung:** Sitzt mit leichter Unterstützung mindestens 30 Sekunden lang (Bayley-Skala – Grobmotorik, Item #19).
- **Sitzen ohne Unterstützung:**
 - Sitzt ohne Unterstützung für ≥ 30 Sekunden (Bayley-Skala – Grobmotorik, Item #26).
 - Sitzt mit aufrechtem Kopf mindestens 10 Sekunden lang aufrecht; benutzt weder Arme noch Hände, um den Körper auszubalancieren oder zu stützen (Definiert nach WHO [32]).
 - Sitzen ohne Unterstützung für ≥ 15 Sekunden (Definiert nach Nationwide Children's Hospital⁵).
 - Sitzen ohne Unterstützung für ≥ 5 Sekunden (Bayley-Skala – Grobmotorik, Item #22; dieser Meilenstein wird in den Quelldaten als "Sitzt allein <10 Sekunden" dargestellt).

Tabelle 12: Standardübersicht Entwicklung motorischer Meilensteine bis zum 60. Lebensmonat [6]

Alter (gesunde Kinder)	Meilenstein
4 Monate	Bauchlage: Das Kind hält den Kopf hoch, Beine ausgestreckt.
6 Monate	Rückenlage: Das Kind rollt von der Rücken- in die Bauchlage.
9 Monate	Freies Sitzen, gerader Rücken.
12 Monate	Das Kind bewegt sich im Raum, hält sich fest und steht ggf. ohne Hilfe.
15 Monate	Freies Gehen.
18 Monate	Das Kind rennt, geht an der Hand gehalten Treppen hinunter.

⁴ Die Bayley-Skala ist eine standardisierte, normenbezogene Säuglingsbeurteilung von Entwicklungsfunktionen in 5 Bereichen: kognitives, sprachliches, motorisches, sozial-emotionales und adaptives Verhalten [11]: Die Bayley-Skala wurde von einer Physiotherapeutin / einem Physiotherapeuten bewertet.

⁵ Lehrkrankenhaus für die pädiatrische Akutversorgung mit Sitz in Columbus, Ohio.

Alter (gesunde Kinder)	Meilenstein
24 Monate	Das Kind geht stufenweise Treppen hoch und runter, springt vom Boden ab.
30 Monate	Das Kind geht Treppen hoch (ein Fuß pro Stufe), steht auf einem Bein.
36 Monate	Das Kind geht Treppen runter (ein Fuß pro Stufe), fährt Dreirad.
48 Monate	Das Kind hüpfte auf einem Bein; wirft Bälle mit dem Handrücken nach oben.
60 Monate	Das Kind ist in der Lage zu springen.

Die Entwicklung signifikanter motorischer Meilensteine nach GMS-Checkliste wird als vorläufige Beurteilung des Erreichens von motorischen Meilensteinen verwendet. Sofern das Prüfpersonal feststellte, dass ein Meilenstein erreicht wurde, wurden die entsprechenden Videos zur Bestätigung an eine zentrale Begutachtung gesandt. In anderen Fällen wurden potentielle Meilensteine möglicherweise nicht mit der GMS-Checkliste dokumentiert, die das Prüfpersonal auf den zu Hause angefertigten Videos identifizierte und anschließend der zentralen Videoevaluation übergab, wo das Erreichen des Meilensteins festgestellt wurde. Daher sind Diskrepanzen zwischen der GMS-Checkliste und der zentralen Videoüberprüfung möglich. In allen Fällen wird die unabhängige zentrale Videoevaluation als maßgebliche Dokumentation des Erreichens der Meilensteine in dieser Studie angesehen.

- Erhebungszeitpunkte:
 - CL-303: Zu Baseline, nach Gentransfer monatlich bis zum Ende der Studie (oder bis zum Alter von 18 Monaten).
 - CL-101: Zu Baseline, nach Infusion der Studienmedikation monatlich bis Studienende.
- Auswertung:
 - CL-303: Auf Basis der ITT-Population, finaler Datenschnitt: 12. Dezember 2019.
 - CL-101: Auf Basis des FAS nach Kohorte, finaler Datenschnitt: 14. Dezember 2017.

Bewertung

Die Operationalisierung des Endpunktes ist nachvollziehbar.

Validität und Patientenrelevanz

Die Bayley-Skala ist eine standardisierte, normenbezogene Säuglingsbeurteilung von Entwicklungsfunktionen in 5 Bereichen: kognitives, sprachliches, motorisches, sozial-emotionales und adaptives Verhalten [1,11]. In den Studien CL-303 und CL-101 wurde insbesondere die motorische Entwicklung der Studienteilnehmenden beurteilt, weshalb für den vorliegenden Endpunkt die grobmotorische Skala der Bayley-Skala eingesetzt wurde. Die Bewertung darüber, ob ein altersgerechter Meilenstein erreicht wurde, erfolgte auf Basis einzelner Item-Definitionen dieser grobmotorischen Skala. Die Items der Skala sind vergleichbar mit einer standardisierten, altersabhängigen Entwicklung von Säuglingen, wie z. B. die altersabhängigen Entwicklungsstufen eines Säuglings nach WHO (siehe Abbildung 8). Leichte Unterschiede zwischen der grobmotorischen Bayley-Subskala und den WHO-Entwicklungsstufen zeigen sich beispielsweise für die Definition des unabhängigen Sitzens, wobei in beiden Fällen ein freies Sitzen ohne Unterstützung erreicht werden muss [31,32].

Im Gegensatz zu Studie CL-303 erfolgte die Datenerhebung zum Erreichen motorischer Meilensteine in Studie CL-101 bei geplanten Studienvisiten zunächst auf Grundlage einer Standardüber-

sicht zur Entwicklung motorischer Meilensteine (Motor Milestone Development Survey) und der GMS-Checkliste. Daraus ergeben sich mögliche Diskrepanzen zwischen der Bewertung durch das Prüfpersonal und der Bewertung durch die zentrale Begutachtung. Zum einen, da die Meilensteine ähnlich, aber teilweise anders definiert sind (z. B. Rollbewegung, Kopfsteuerung), andererseits bewertete die unabhängige, zentrale Begutachtung Videoaufnahmen der Studienvisiten sowie Heimvideos der Erziehungsberechtigten. Da in beiden Studien die unabhängige, externe Begutachtung auf Basis spezieller Items der Bayley-Skala erfolgte und diese Bewertung die maßgebliche für die Auswertung des Endpunktes darstellt, ist von einer validen Endpunkterhebung auszugehen. Ein motorischer Meilenstein gilt in Studie CL-101 als erreicht, sofern die Patientin / der Patient zum Zeitpunkt der festgelegten Datenschnitte nach Gentransfer mittels Videoevidenz die Leistung der einzelnen motorischen Fähigkeiten zeigte. In Studie CL-303 war ein erreichter, altersbedingter Meilenstein nach Gentransfer bei jeder entsprechenden Studienvisite bis zu und einschließlich der letzten Studienvisite im Alter von 18 Monaten vorzuzwies (Responderdefinition).

Die Veränderung der motorischen Funktionsfähigkeit in Form der Entwicklung motorischer Meilensteine wird als patientenrelevant angesehen. Kinder mit der infantilen Form der SMA werden per Krankheitsdefinition niemals das Sitzen erlernen, die Lebenserwartung wird in der Literatur mit < 2 Jahren angegeben. Das Erreichen motorischer Meilensteine im Sinne einer normalen Kindesentwicklung wird entsprechend als Unterbrechen der Krankheitsprogression gewertet [10,13].

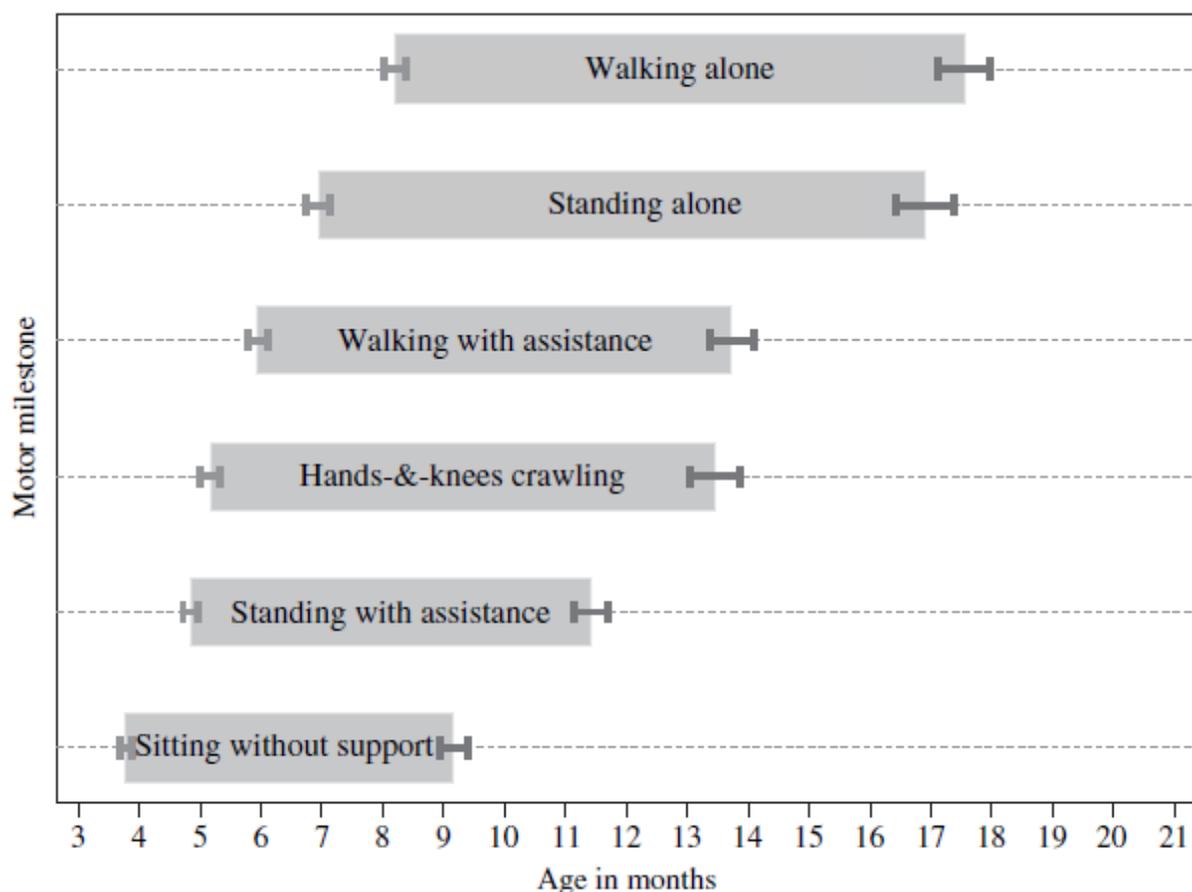


Abbildung 8: Altersabhängige Entwicklungsstufen Säuglinge, gemäß WHO [32]

Bayley-Skala (Grob- und Feinmotorik)

Der Endpunkt „Bayley-Skala (Grob- und Feinmotorik)“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Die Bayley-Skala ist eine standardisierte, normenbezogene Säuglingsbeurteilung von Entwicklungsfunktionen in 5 Bereichen: kognitives, sprachliches, motorisches, sozial-emotionales und adaptives Verhalten.

CL-303: In Studie CL-303 wurden die grobmotorische und feinmotorische Subskala der Bayley-Skala durch qualifiziertes klinisches Studienpersonal erhoben und per Videoevaluation durch eine zentrale, unabhängige Begutachtung bewertet.

CL-101: Personen, die im CHOP-INTEND (The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders) einen Wert von ≥ 60 erreicht haben, wurden mit der Bayley-Skala bewertet. Bei Patientinnen und Patienten, die an mehr als zwei aufeinanderfolgenden Studienvisiten einen CHOP-INTEND-Wert von 62 überschritten, wurde ausschließlich die Bayley-Skala erhoben. Die Erhebung des Endpunktes erfolgte durch qualifiziertes klinisches Studienpersonal, die Bewertung per Videoevaluation durch eine zentrale, unabhängige Begutachtung.

- Erhebungszeitpunkte:
 - CL-303: Die Subskalen zur Grob- und Feinmotorik der Bayley-Skala wurden zur Baseline erhoben sowie, nach Gentransfer, monatlich, beginnend ab Visite 6.
 - CL-101: Die Skalen zur Grob- und Feinmotorik wurden nach Gentransfer monatlich bis Monat 12 oder bis zu einem Patientenalter von 15 Monaten erhoben, je nachdem was später eintrat, sofern eine Patientin / ein Patient einen CHOP-INTEND-Wert von 60 der 64 Punkte erreichte oder überschritt. Die Bewertung des CHOP INTEND wurde abgebrochen und nur die Bayley-Skala erhoben, wenn 2 aufeinanderfolgende CHOP-INTEND-Werte von ≥ 62 erreicht wurden. Im 2. Jahr der Nachbeobachtung erfolgte die Erhebung vierteljährlich.
- Auswertung:
 - CL-303: Auf Basis der ITT-Population, finaler Datenschnitt: 12. Dezember 2019.
 - CL-101: Auf Basis des FAS, finaler Datenschnitt: 26. Dezember 2017.

Bewertung

Die Operationalisierung des Endpunkts ist nachvollziehbar.

Mit der feinmotorischen Subskala der Bayley-Skala [1,11] werden insbesondere die Hand-Finger-Koordination sowie die Mimik (Augen-, Mund- und Gesichtsbewegung) der Patientinnen und Patienten erfasst. Mit der grobmotorischen Subskala der Bayley-Skala werden die Bewegungsfunktionen des kindlichen Körpers, wie Kopfheben, Rollen, Krabbeln, Sitzen, Laufen, Gehen, Springen untersucht. Die feinmotorische Skala besteht insgesamt aus 66 Items, die grobmotorische Skala aus 72 Items. Die Bewertung der Bewegungen erfolgte altersabhängig. In Studie CL-101 wurde die grob- und feinmotorische Subskala der Bayley-Skala nur für Patientinnen und Patienten mit einem CHOP-INTEND-Wert ≥ 60 erhoben. Die Erhebung der Bayley-Skala erfolgte in beiden Studien durch qualifiziertes Prüfpersonal, die Auswertung per Videoevaluation durch eine unabhängige zentrale Begutachtung.

Patientenrelevanz und Validität

Die gestörte Entwicklung der kindlichen Grob- und Feinmotorik ist Teil der Krankheitssymptomatik der infantilen SMA. Eine Veränderung dieser Symptomatik im Sinne einer verbesserten motorischen Entwicklung des Säuglings oder Kleinkindes wird als patientenrelevant erachtet.

Die Bayley-Skala ist eine standardisierte, etablierte, fremdbewertete Skala zur Beurteilung verschiedener Entwicklungsfunktionen bei Säuglingen und Kleinkinder von 1 bis 42 Monaten: kognitives, sprachliches, motorisches, sozial-emotionales und adaptives Verhalten. In den vorliegenden Studien wurden die grob- und feinmotorische Subskala bewertet. Die grobmotorische Skala umfasst 72 Items, mit denen folgende Funktionen beurteilt werden: statische Haltung (z. B. Kopfsteuerung, Sitzen, Stehen), dynamische Bewegung einschließlich Fortbewegung (Krabbeln, Gehen, Laufen, Springen, Treppen steigen und Fallen), Qualität der Bewegung (Koordination beim Aufstehen, Gehen, Treten), Gleichgewicht, Motorik (motor planning), Wahrnehmungsmotorische Koordination (z. B. Nachahmung von Körperhaltungen). Die feinmotorische Subskala umfasst 66 Items und beurteilt die Funktionen Auffassungsgabe, Wahrnehmungsmotorische Koordination, Motorik (motor planning), motorische Geschwindigkeit, wobei die Items altersgerechte Fähigkeiten messen, einschließlich der visuellen Erfassung, Reichweite, Handhabung von Objekten, Greifen, Bewegungsqualität von Kindern, funktionelle Fertigkeiten der Hand, Reaktionen auf tastbare Informationen (sensorische Integration). Beurteilt wird die gegenwärtige motorische Funktion der Kinder mit einer binären Antwort-Skala (ja/nein).

Ein Kind muss zu Beginn bei den ersten drei aufeinanderfolgenden Items des entsprechenden Alters eine Punktzahl von 1 erreichen, um auf der Skala voranzukommen. Wenn das Kind bei einem der ersten drei Items eine Nullbewertung erhält, muss es zum Startpunkt für das vorangegangene Alter zurückgehen und diese Items erneut erfüllen. Sofern das Kind bei 5 aufeinanderfolgenden Items eine Punktzahl von Null erreicht, sollte der Test beendet werden. Punkte werden pro Subskala berechnet, eine Gewichtung wird nicht vorgenommen. Ein höherer Punktwert weist auf eine bessere Entwicklung hin. Für die motorische Skala wird ein Komposit-Wert generiert. Die motorische Domäne der Bayley-Skala wird als reliabel angesehen. Die interne Konsistenz ist mit 0,57–0,97 (Kuder-Richardson-Koeffizient) bei Säuglingen im Alter zwischen 1 und 15 Monaten angegeben. Die Test-Retest-Reliabilität erreichte für die motorischen Items in einer Population von 8 Monate alten Säuglingen und einem Zeitintervall von einer Woche für die Wiederholung des Tests eine Übereinstimmung von 75,5 (SD: 14,5). Bei der Inter-Interviewer-Reliabilität wurde eine Übereinstimmung von 93,4 (SD: 3,2) erreicht. In klinischen Studien zeigte sich, dass die Bayley-Skala (Version 3.0) eine hohe Diskriminanz sowie eine gute Sensitivität und Spezifität bei der Unterscheidung klinisch auffälliger von normal entwickelten Kindern aufweist. Zur Änderungssensitivität des Instrumentes liegen keine Untersuchungen vor.

CHOP INTEND

Der Endpunkt „CHOP INTEND“ wird für die Nutzenbewertung berücksichtigt.

Operationalisierung

Der CHOP INTEND ist eine Funktionsskala für die Bewertung der Entwicklung der motorischen Funktion, speziell für neuromuskulär erkrankte Säuglinge oder Kleinkinder. Die Skala besteht aus 16 Items, die mit 0 (ohne Funktion), 1, 2, 3 oder 4 (volle Funktion) bewertet werden, wobei maximal ein Wert von 64 erreicht werden kann. Höhere Werte weisen auf eine bessere motorische Funktion hin. Beurteilt werden entweder spontane Bewegungen (durch Beobachtung) oder aktive Bewe-

gungen: (1) spontane Bewegung (obere Extremität); (2) spontane Bewegung (untere Extremität); (3) Handgriff; (4) Kopf in der Mittellinie; (5) Hüftadduktoren; (6) Drehen von den Beinen ausgelöst; (7) Drehen von den Armen ausgelöst; (8) Schulter- und Ellbogenflexion und horizontale Abduktion; (9) Flexion von Schulter und Ellbogen; (10) Knieextension; (11) Flexion der Hüfte und Dorsalflexion des Fußes, (12) Kopfkontrolle, (13) Ellbogenflexion, (14) Halsflexion, (15) Kopf-/Nackenextension, (16) Lateralflexion der Wirbelsäule.

Die bei jeder Studienvsiste erforderlichen physiotherapeutischen Beurteilungen mittels CHOP INTEND wurden auf Video aufgezeichnet, um Veränderungen der funktionellen Fähigkeiten zu bestimmen. Die Videos hinsichtlich des Nachweises für das Erreichen des Meilensteins wurden durch eine zentrale, unabhängige Begutachtung beurteilt. Dabei wurde der früheste Zeitpunkt für das Erreichen des Meilensteins auf den Videos ermittelt. Zusätzlich konnten die Eltern/Erziehungsberechtigten während der gesamten Studie Heimvideos einreichen, die das Erreichen der Entwicklungsmeilensteine belegen. Videoaufzeichnungen wurden in allen berücksichtigten Studien durchgeführt.

CL-303: Patientinnen und Patienten, die drei aufeinanderfolgende CHOP-INTEND-Werte ≥ 58 erreichen, werden keiner zusätzlichen CHOP-INTEND-Untersuchungen unterzogen.

CL-101: Bei Patientinnen und Patienten, die zwei aufeinanderfolgende CHOP-INTEND-Werte ≥ 62 erreichten, sollte überprüft werden, ob weitere CHOP-INTEND-Beurteilungen erforderlich waren.

- Erhebungszeitpunkte:
 - CL-303: Zu Baseline, nach Gentransfer an den Tagen⁶ 7, 14, 21, 30, dann monatlich bis zum Alter von 14 Monaten sowie zu Studienende (Alter 18 Monate).
 - CL-101: Zu Baseline, nach Gentransfer monatlich bis Monat 12 im ersten Jahr der Nachbeobachtung, im zweiten Jahr der Nachbeobachtung vierteljährlich.
- Auswertung:
 - CL-303: Auf Basis der ITT-Population, finaler Datenschnitt 12. Dezember 2019.
 - CL-101: Auf Basis der ITT-Population, finaler Datenschnitt 14. Dezember 2017.

Bewertung

Die Operationalisierung des Endpunkts in den Studien ist weitgehend nachvollziehbar. Jede Erhebung des CHOP INTEND erfolgte durch qualifiziertes Prüfpersonal (Physiotherapeut/in), wurde per Videoaufzeichnung dokumentiert und durch eine zentrale, unabhängige Begutachtung bewertet.

Validität und Patientenrelevanz

Der CHOP INTEND ist ein von Expertengruppen entwickeltes, etabliertes, krankheitsspezifisches Instrument. In mehreren Studien konnten gute Testgütekriterien nachgewiesen werden, wobei zur Validität weitere Studien notwendig sind. Untersuchungen zur klinischen Relevanzschwelle des Instrumentes existieren nicht. Unter Berücksichtigung des natürlichen Krankheitsverlaufs bei Personen mit bestätigter SMA Typ 1, bei denen eine Verbesserung des Gesundheitszustandes nicht zu erwarten ist, erscheint jeglicher Fortschritt motorischer Fähigkeiten als erstrebenswert und

⁶ Wenn eine CHOP-INTEND-Beurteilung an Tag -1 nicht durchgeführt wird, sollte eine CHOP-INTEND-Beurteilung an Tag 1 vor der Verabreichung der Dosis abgeschlossen werden.

patientenrelevant. Der CHOP INTEND gilt als etabliertes Instrument in der Patientenpopulation, da bereits in verschiedenen klinischen Studien eingesetzt, ohne jedoch abschließend validiert zu sein. [17,22,23]

Die Erhebung der motorischen Fähigkeiten mittels CHOP INTEND wird als patientenrelevant angesehen. Gemäß § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustands, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“.

Gedeihfähigkeit

Der Endpunkt „Gedeihfähigkeit“ wird für die Nutzenbewertung nicht berücksichtigt.

Bei der Gedeihfähigkeit handelt es sich um einen kombinierten Endpunkt, der sich in den Studien CL-101 und CL-303 jeweils aus drei Teilkomponenten zusammensetzt:

- Toleranz von Flüssigkeiten nachgewiesen durch formalen Schlucktest.
- Keine Ernährung über unterstützende mechanische Maßnahmen.
- Gewichtserhalt innerhalb einer altersgerechten Spanne (in Studie CL-101: > 3. Perzentil für Alter und Geschlecht in Bezug auf die Werte der WHO-Leitlinien).

Die Auswahl dieser Kriterien für den Endpunkt „Gedeihfähigkeit“ wird im vorliegenden Nutzen-dossier nicht begründet.

Als Gedeihstörung gilt das Ausbleiben der erwarteten physiologischen Entwicklung eines Kindes (ICD 62.8 [14]) mit Verzögerung der körperlichen Entwicklung oder mangelnder Gewichtszunahme, die häufig mit motorischen und psychosozialen Auffälligkeiten der Entwicklung verbunden sind. Das altersabhängige Gewicht hat somit einen unmittelbaren Einfluss auf die Entwicklung und Gedeihfähigkeit eines Säuglings bzw. Kleinkindes. Bei Kindern mit infantiler SMA ist in einem bestimmten Krankheitsstadium das Schlucken und die Aufnahme von Nahrung nahezu immer erschwert, was zu einer Ernährungsunterstützung führen kann, z. B. durch eine Sonde. Durch eine entsprechende unterstützende Ernährung ist es möglich, eine ausreichende Nahrungsaufnahme sowie einen Gewichtserhalt zu erreichen. Die Einzelkriterien des Endpunktes stehen miteinander in Zusammenhang, jedoch sind die Kriterien von unterschiedlicher Relevanz für die Gedeihfähigkeit des Kindes, da lediglich ein unzureichendes altersbedingtes Gewicht auf eine Gedeihstörung hinweist. Aufgrund qualitativer und quantitativer Heterogenität des kombinierten Endpunktes werden die Ergebnisse des Endpunktes „Gedeihfähigkeit“ nachfolgend nicht dargestellt. Ergebnisse zum altersbedingten Gewicht der Patientinnen und Patienten werden in den Studien CL-101 und CL-303 nicht berichtet.

Non-orale Ernährung

Der Endpunkt „Non-orale Ernährung“ wird für die Nutzenbewertung nicht berücksichtigt (siehe Bewertung der Gedeihfähigkeit).

2.3.3 Lebensqualität

Die Lebensqualität wurde in den eingeschlossenen Studien CL-101 und CL-303 nicht erhoben.

2.3.4 Sicherheit

Unerwünschte Ereignisse

Unerwünschte Ereignisse (UE) werden für die Nutzenbewertung berücksichtigt.

Operationalisierung

CL-101, CL-303, CL-302 und CL-304: UE sind definiert als die Entwicklung eines unerwünschten medizinischen Zustandes oder die Verschlechterung eines bereits bestehenden medizinischen Zustandes nach oder während der Exposition gegenüber einem pharmazeutischen Produkt, unabhängig davon, ob es als zufällig mit dem Produkt verbunden angesehen wird oder nicht. Die Erhebung des Endpunktes erfolgte im Rahmen der Sicherheitsbewertung durch das Prüfpersonal und ein unabhängiges Komitee zur Sicherheitsüberwachung (Data Safety Monitoring Board, DSMB).

Alle UE werden in Übereinstimmung mit den CTCAE (Common Terminology Criteria for Adverse Events) Version 4.03 eingestuft:

- Grad 1: mild (UE ist asymptomatisch oder mit milden Symptomen verbunden; lediglich klinische oder diagnostische Beobachtung, keine Intervention erforderlich).
- Grad 2: moderat (minimale, lokale oder nicht-invasive Intervention nötig; UE unterbricht die gewöhnlichen Aktivitäten des täglichen Lebens).
- Grad 3: schwer (schwer oder medizinisch bedeutsam, aber nicht unmittelbar lebensbedrohlich, Hospitalisierung oder Verlängerung der Hospitalisierung notwendig, zur Behinderung/Invalidität führend, Einschränkung der Aktivitäten des täglichen Lebens und der Selbstversorgung).
- Grad 4: lebensbedrohlich (lebensbedrohliche Konsequenzen, dringende Intervention erforderlich).
- Grad 5: Tod (UE steht im Zusammenhang mit dem Tod).

Der Einschluss in die Studie wird unterbrochen, wenn bei einer Person ein unerwarteter CTCAE-Grad 3 oder eine höhere Toxizität durch UE auftritt, die möglicherweise, wahrscheinlich oder definitiv mit der Genersatztherapie in Zusammenhang stehen. Das Ereignis wird dann vom DSMB überprüft, und es wird beurteilt, ob die Studie nach den Abbruchregeln vorzeitig beendet werden sollte.

Unerwartete UE CTCAE-Grad 3 oder höher, die möglicherweise, wahrscheinlich oder definitiv mit der Genersatztherapie in Zusammenhang stehen, müssen gemäß Sicherheitsmanagementplan der Studie innerhalb von 24 Stunden gemeldet werden, um eine rechtzeitige Eskalation an das DSMB zu gewährleisten.

Schwerwiegende unerwünschte Ereignisse

Ein schwerwiegendes unerwünschtes Ereignis (SUE) ist definiert als UE, das

- den Tod einer Patientin / eines Patienten zur Folge hatte,
- unmittelbar lebensbedrohlich war,
- eine Einweisung ins Krankenhaus oder einen verlängerten Krankenhausaufenthalt erforderlich machte,
- eine anhaltende oder schwerwiegende Behinderung oder Einschränkung oder
- eine kongenitale Anomalie oder einen Geburtsfehler nach sich zog.

Unerwünschte Ereignisse von besonderem Interesse

Als unerwünschte Ereignisse von besonderem Interesse (AESI) wurden erhöhte Leberenzymwerte im Studienprotokoll präspezifiziert. Darüber hinaus finden sich im Studienbericht noch weitere Kategorien mit dazugehörigen Preferred Terms: „Hepatotoxizität“, „Thrombozytopenie“, „Kardiale Ereignisse“, „Auf Ganglionopathie hindeutende sensorische Anomalien“. In Studie CL-101 wurden diese AESI nicht präspezifiziert.

Signifikante unerwünschte Ereignisse von besonderer klinischer Bedeutung, bei denen es sich nicht um ein SUE und UE handelt, die zum Abbruch der Studie durch die Patientin / den Patienten führen, werden als AESI klassifiziert.

- Erhebungszeitpunkte:
 - CL-303: Zu jeder geplanten Studienvsiste bis Studienende (Alter 18 Monate) + 0–14 Tage.
 - CL-101: Zu jeder geplanten Studienvsiste bis 30 Tage nach letzter Studienvsiste.
 - CL-302+CL-304: Zu jeder geplanten Studienvsiste bis zur letzten Studienvsiste.
- Auswertung:
CL-101, CL-303, CL-302, CL-304: Auf Basis der Sicherheitspopulation.

Bewertung

Die Operationalisierung des Endpunktes ist nachvollziehbar.

Validität und Patientenrelevanz

Das Auftreten von UE ist patientenrelevant. Die Patientenrelevanz von Laborparametern ist jedoch unklar. Der Endpunkt wird als valide bewertet. 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“.

Hospitalisierung

Der Endpunkt „Hospitalisierung“ wird in der Nutzenbewertung berücksichtigt.

CL-303: Als Hospitalisierung gilt jede Aufnahme einer Patientin / eines Patienten ins Krankenhaus. Der Endpunkt wurde im Rahmen der Sicherheitsbewertung unter SUE erfasst, für Patientinnen und Patienten im Alter von 13,6 und 24 Monaten. Die Erhebung erfolgte durch das Prüfpersonal und das DSMB.

CL-101: Erfasst wurde jeder Krankenhausaufhalt einer Patientin / eines Patienten während der Studie mit Angabe des Grundes für die Einweisung sowie der Dauer des Aufenthaltes (Tage). Die Erhebung erfolgte im Rahmen der Sicherheitsbewertung (SUE) durch das Prüfpersonal und das DSMB.

- Erhebungszeitpunkte:
 - CL-303: Über den gesamten Zeitraum der Studie.
 - CL-101: Über den gesamten Zeitraum der Studie.
- Auswertung:
 - CL-303: Auf Basis der Sicherheitspopulation, finaler Datenschnitt: 12. Dezember 2019
 - CL-101: Auf Basis der Sicherheitspopulation, finaler Datenschnitt: 14. Dezember 2017.

Bewertung

Die Operationalisierung des Endpunktes ist nur teilweise nachvollziehbar. Es fehlen Angaben zum Mindestaufenthalt in der Einrichtung und darüber, ob der Aufenthalt im Zusammenhang mit der Erkrankung oder einem UE steht. Darüber hinaus bleibt offen, ob auch geplante Krankenhausaufenthalte mit in die Bewertung eingingen. Von regionalen Unterschieden bei der Erhebung ist nicht auszugehen, da beide Studien in den USA durchgeführt wurden.

Validität und Patientenrelevanz

Der Endpunkt wurde im Rahmen der Sicherheitsüberwachung durch ein verblindetes Komitee beurteilt.

Gemäß § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustands, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität“. Entsprechend ist der Endpunkt „Hospitalisierungen“ als patientenrelevant einzustufen und wird für die Bewertung des Zusatznutzens berücksichtigt.

2.3.5 Übersicht der Vergleichbarkeit der Endpunkte in den Studien CL-101 und CL-303 mit den externen Kontrollen (PNCR, NeuroNEXT, ENDEAR)

In Tabelle 13 findet sich eine Auflistung aller in der Nutzenbewertung berücksichtigten patientenrelevanten Endpunkte der Studien CL-303 und CL-101 (siehe Tabelle 11), für die indirekte Vergleiche mit den externen Kontrollen (PNCR, NeuroNEXT, ENDEAR) vorgelegt wurden.

Tabelle 13: Vergleichbarkeit der Endpunkte der Studien CL-303 und CL-101 mit den externen Vergleichskohorten (PNCR, NeuroNEXT, ENDEAR (BSC))

Endpunkt	Kategorie	PNCR	NeuroNEXT	ENDEAR (BSC)
Mortalität	Mortalität	ja	ja	ja
Tod oder dauerhafte Beatmung (Überleben)		ja	nein ¹⁾	nein ¹⁾
Dauerhafte Beatmung	Morbidität	ja	nein ¹⁾	nein ¹⁾
Ventilatorische Unterstützung (Beatmungsfreiheit)		– ²⁾	– ²⁾	ja
Erreichen motorischer Meilensteine		nein ³⁾	nein ³⁾	nein ⁴⁾
CHOP INTEND		nein ⁵⁾	nein ⁶⁾	nein ⁷⁾
Bayley-Skala		– ²⁾	– ²⁾	– ²⁾
Unerwünschte Ereignisse	Sicherheit	– ²⁾	– ²⁾	ja
Hospitalisierungen		– ²⁾	– ²⁾	nein ⁸⁾

¹⁾ Die Operationalisierung der dauerhaften Beatmung ist nicht vergleichbar mit den Studien CL-303 und CL-101.

²⁾ Dieser Endpunkt wurde nicht erfasst bzw. es liegen keine Angaben zu dem Endpunkt vor.

³⁾ Die Operationalisierung und Erfassung der motorischen Meilensteine in den Studien PNCR und NeuroNEXT sind nicht vergleichbar mit den Studien CL-303 und CL-101.

⁴⁾ In der Studie ENDEAR wurde die motorische Entwicklung des Kleinkindes mittels HINE-Subskala 2 erfasst. Diese umfasst jedoch keine in den Studien CL-101 und CL-303 untersuchten vergleichbaren motorischen Meilensteine.

⁵⁾ Für die Studie PNCR liegen nur Daten für das Erreichen eines CHOP-INTEND-Wertes von ≥ 40 Punkten in einem Alter von ≥ 6 Monaten vor. In den Studien CL-101 und CL-303 waren a priori Responderauswertungen für den Anteil an Personen mit

einem CHOP-INTEND-Schwellenwert von ≥ 40 , ≥ 50 und ≥ 60 (CL-101) bzw. ≥ 58 (CL-303) geplant. Es liegen keine Angaben zur Endpunkterfassung/-systematik (u. a. Erhebungszeitpunkte, Umgang mit fehlenden Werten) des CHOP INTEND in der Studie PNCR vor. Darüber hinaus fehlen Angaben über die tatsächliche Beobachtungszeit.

- ⁶⁾ Für die Studie NeuroNEXT liegen CHOP-INTEND-Daten für die Veränderung pro Monat, die Veränderung über den gesamten Studienverlauf sowie für das Erreichen eines CHOP-INTEND-Wertes von ≥ 40 Punkten in einem Alter von ≥ 6 Monaten vor. Die Erfassung des CHOP INTEND in der Studie NeuroNEXT unterscheidet sich bezogen auf die Anzahl an Messhäufigkeiten zu den Zulassungsstudien CL-303 und CL-101 (nähere Informationen siehe Abschnitt 2.4.4). Zudem wurden für die Auswertung nur Patientinnen und Patienten in der Studie NeuroNEXT berücksichtigt, für die Daten in einem Alter ≤ 12 Monate vorliegen (10/16 Personen). Analog zu Studie PNCR liegen keine Angaben zur Endpunkterfassung/-systematik (u. a. Erhebungszeit-punkte, Umgang mit fehlenden Werten) des CHOP INTEND in der Studie NeuroNEXT vor. Darüber hinaus fehlen Angaben über die tatsächliche Beobachtungszeit.
- ⁷⁾ In der Studie ENDEAR liegen Auswertungen für die Veränderung des CHOP-INTEND-Wertes vor, operationalisiert als Anteil an Personen mit einer Verbesserung, wobei ein Responder definiert war als eine Person mit einer Verbesserung von mindestens 4 Punkten im Verlauf der Studie.
- ⁸⁾ Die Operationalisierung der Hospitalisierung und Beobachtungszeiten in der Studie ENDEAR sind nicht vergleichbar mit den Studien CL-303 und CL-101.

Abkürzungen: BSC: Best Supportive Care; CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; HINE: Hammersmith Infant Neurological Examination; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA.

Dauerhafte Beatmung

Der Endpunkt „Tod oder dauerhafte Beatmung (Überleben)“ war in den Studien CL-101 und CL-303 folgendermaßen operationalisiert:

- **CL-303:** Überleben ist definiert als das Vermeiden von entweder a) Tod oder b) dauerhafter Beatmung. Dauerhafte Beatmung ist definiert als die Notwendigkeit entweder einer Tracheostomie (invasiv) oder ventilatorischer Unterstützung von ≥ 16 Stunden pro Tag (nicht-invasiv) über einen Zeitraum von ≥ 14 aufeinanderfolgenden Tagen in Abwesenheit einer akuten reversiblen Erkrankung, ausgenommen perioperative Beatmung.
- **CL-101:** Überleben ist definiert als die Zeit von der Geburt bis a) Tod oder b) dauerhafte Beatmung bis zum Alter von 13,6 und 20 Monaten. Dauerhafte Beatmung war in diesem Kontext definiert als die Notwendigkeit entweder einer invasiven Beatmung oder respiratorischer Unterstützung ≥ 16 Stunden pro Tag (inklusive nicht-invasiver Beatmung) über einen Zeitraum von ≥ 14 aufeinanderfolgenden Tagen in Abwesenheit einer akuten reversiblen Erkrankung, ausgenommen perioperative Beatmung.

Entsprechend Definition wurde in beiden Zulassungsstudien sowohl die invasive als auch die nicht-invasive Beatmung in dem Teilendpunkt „Dauerhafte Beatmung“ berücksichtigt. Eine vergleichbare Operationalisierung für die dauerhafte Beatmung ist ausschließlich für Studie PNCR gegeben. Hier ist das Überleben definiert als das Alter beim Erreichen von (a) Tod oder bei (b) dauerhafter Beatmung. Dauerhafte Beatmung war in diesem Kontext definiert als die Notwendigkeit entweder einer Tracheostomie oder ≥ 16 Stunden ventilatorischer Unterstützung pro Tag (nicht-invasiv) über einen Zeitraum von mindestens 14 aufeinanderfolgenden Tagen in Abwesenheit einer akuten reversiblen Erkrankung und perioperativer Beatmung.

In der Studie NeuroNEXT wurde das Überleben definiert als Zeit bis (a) zum Tod oder (b) zur dauerhaften invasiven Beatmung (Intubation). Entgegen der Definition in den Zulassungsstudien CL-303 und CL-101 ist die nicht-invasive Beatmung in der Studie NeuroNEXT nicht erfasst. Darüber hinaus fehlen Angaben über die Mindestdauer eines Ereignisses (invasive Beatmung), wie sie für die Studien CL-303, CL-101 und PNCR vorliegen (≥ 16 Stunden ventilatorischer Unterstützung pro Tag von ≥ 14 aufeinanderfolgenden Tagen). Die Tracheostomie (CL-303, CL-101) und die Intubation (NeuroNEXT) werden für die vorliegende Nutzenbewertung jedoch als hinreichend vergleichbar im Sinne einer invasiven Beatmungsunterstützung angesehen. Die Unterschiede der Studien in der Erfassung der dauerhaften Beatmung können zu einer Überschätzung

des Effektes zuungunsten von Onasemnogen-Abeparvovec führen, da in den Zulassungsstudien CL-303 und CL-101 zusätzlich die nicht-invasiven Fälle erfasst werden. Diese stellen im Vergleich zu einer invasiven Beatmung eher mildere Formen der Beatmungssituation dar und werden womöglich häufiger erfasst, was zu höheren Ereignishäufigkeiten in den entsprechenden Studienarmen führen kann.

In der Studie ENDEAR wurde sowohl die invasive als auch nicht-invasive Beatmung für mindestens 16 Stunden am Tag als dauerhaft gewertet. Die Operationalisierungen der Studien unterscheiden sich jedoch hinsichtlich des Kriteriums „Mindestdauer für das Ereignis“. In den Studien CL-303, CL-101 und PNCR ist der Zeitraum mit 14 Tagen kürzer als in der Studie ENDEAR mit 21 Tagen. Auch hier besteht die Möglichkeit einer Verzerrung zuungunsten von Onasemnogen-Abeparvovec, da der Endpunkt in den Studien CL-101 und CL-303 früher erreicht wird als in der Studie ENDEAR. Für die Studie ENDEAR war eine Beobachtungszeit von 13 Monaten geplant, aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht in etwa 6,2 Monaten. Durch die längere Beobachtungsdauer in den Studien CL-101 und CL-303 ist das Ergebnis unter Berücksichtigung der ITT-Population potentiell zuungunsten von Onasemnogen-Abeparvovec verzerrt.

Der Zeitpunkt der Endpunktauswertung wurde für die Vergleichbarkeit mit den Zulassungsstudien CL-101 und CL-303 auf das Alter 13–14 Monate festgelegt. In der Studie CL-303 sind die Personen bis zu einem Alter von mindestens 18 Monaten, in der Studie CL-101 bis zu 24 Monate nach Gentransfer beobachtet worden. Für die Studien PNCR und NeuroNEXT finden sich nur Angaben über die geplante Beobachtungsdauer (PNCR: bis zu 36 Monate; NeuroNEXT: bis zu 24 Monate). Vom pU wurde der frühere Vergleichszeitpunkt 13–14 Monate gewählt, um einen adäquateren Vergleich mit den Studien zum natürlichen Verlauf zu ermöglichen. Beobachtungszeit-adjustierte Effektschätzer werden vom pU nicht berichtet.

Ventilatorische Unterstützung (Beatmungsfreiheit)

Der Endpunkt „Ventilatorische Unterstützung (Beatmungsfreiheit)“ war in den Studien CL-303 und CL-101 folgendermaßen operationalisiert:

- CL-303: Unabhängigkeit von ventilatorischer Unterstützung ist definiert als der Anteil der Personen, die von der unterstützenden Beatmung unabhängig waren, im Alter von 18 Monaten keine tägliche Beatmungsunterstützung benötigten, in Abwesenheit einer akuten reversiblen Erkrankung sowie perioperativer Beatmung, durch Bewertung der tatsächlichen Nutzungsdaten des Beatmungsgerätes (siehe Abschnitt 2.3.1).
- CL-101: Unabhängigkeit von ventilatorischer Unterstützung ist definiert als der Anteil Personen, die von einer unterstützenden Beatmung unabhängig waren und die an keinem der 3 Zeitpunkte zur Wirksamkeitsanalyse eine tägliche Beatmungsunterstützung⁷ benötigten, bei Abwesenheit einer akuten reversiblen Erkrankung und Ausschluss einer perioperativen Beatmung⁸.

⁷ Zusatzgeräte, die in der Lungenheilkunde eingesetzt werden, wie z. B. ein Husten-Hilfsgerät, wurden nicht als Beatmungsunterstützung betrachtet.

⁸ Bei Personen, bei denen zum Zeitpunkt des Wirksamkeitsdatenschnitts eine akute reversible Erkrankung vorlag oder eine perioperative Beatmung erforderlich war, wurde die Unabhängigkeit der Beatmung definiert als die

Für die Studie ENDEAR liegen Daten für die Anzahl der Personen ohne Beatmung (unabhängig von der Beatmungsdauer) vor, die für einen Vergleich mit den Studien CL-303 und CL-101 herangezogen wurden. Es gehen in der Studie ENDEAR alle Personen in die Bewertung ein, unabhängig von der Dauer der Beatmung. Daraus resultierend werden die Personen als „Responder“ gewertet, die nie beatmet werden mussten. Die Endpunktauswertung erfolgte in der Studie CL-303 bis zu einem Alter von mindestens 18 Monaten und in der Studie CL-101 bis 24 Monate nach Gentransfer. Für die Studie ENDEAR war eine Beobachtungszeit von 13 Monaten geplant, aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht in etwa 6,2 Monaten. Durch die längere Beobachtungsdauer in den Studien CL-101 und CL-303 ist das Ergebnis unter Berücksichtigung der ITT-Population potentiell zuungunsten von Onasemnogen-Abeparvovec verzerrt. Darüber hinaus ist die Beatmungssituation der Patientinnen und Patienten bereits über den Teilendpunkt des kombinierten Endpunktes „Tod oder dauerhafte Beatmung“ erfasst und wird lediglich für die Studien CL-101 und CL-303 ergänzend dargestellt.

Erreichen motorischer Meilensteine

Für den Vergleich der Studien CL-303 und CL-101 mit den Studien PNCR und NeuroNEXT legt der pU Daten zu folgenden vier motorischen Meilensteinen vor:

- „Freies Sitzen \geq 10 Sekunden“ (CL-101, CL-303 (gemäß WHO-Definition)),
- „Freies Sitzen \geq 30 Sekunden“ (CL-101, CL-303: Bayley-Skala – Grobmotorik, Item #26),
- „Freies Stehen“ (CL-101, CL-303: Bayley-Skala – Grobmotorik, Item #40) und
- „Freies Gehen“ (CL-101: nicht erfasst; CL-303: Bayley-Skala – Grobmotorik, Item #43).

Für die Studien PNCR und NeuroNEXT liegen keine Angaben zu motorischen Meilensteinen in einer vergleichbaren objektiven Endpunkterfassung wie in den Zulassungsstudien CL-303 und CL-101 vor. So wurden in der Studie PNCR die motorischen Fähigkeiten ausschließlich mittels CHOP INTEND erfasst. In der Studie NeuroNEXT wurden die Instrumente Test of Infant Motor Performance Screening Items (TIMPSI), CHOP INTEND und der Alberta Infant Motor Scale (AIMS [27,28])⁹ eingesetzt. Alle drei Skalen dienen der Erfassung motorischer Fähigkeiten, jedoch fehlen Angaben über die Vergleichbarkeit mit den in den Zulassungsstudien CL-101 und CL-303 erfassten motorischen Meilensteinen. In der Studie ENDEAR erfolgte die Beurteilung der Entwicklung motorischer Meilensteine anhand der Subskala 2 des Hammersmith Infant Neurological Examination (HINE), einem Instrument zur Untersuchung der motorischen Funktionsfähigkeit bei Säuglingen und Kleinkindern zwischen 2 und 24 Monaten. Die Erfassung sowie die Bewertung der motorischen

Tatsache, dass während des 2-Wochen-Intervalls vor dem Auftreten der akuten reversiblen Erkrankung oder der perioperativen Anwendung keine Beatmungsunterstützung erforderlich war.

⁹ AIMS ist ein Screening-Instrument zur Identifizierung einer verzögerten motorischen Entwicklung. Dieses kann von Geburt bis zum Alter von 18 Monaten (bis hin zum selbstständigen Gehen) angewendet werden. Die AIMS besteht insgesamt aus 58 Items und unterteilt sich in 4 Kategorien: Bauchlage (21 Items), Rückenlage (9 Items), Sitzen (12 Items) und Stehen (16 Items). Jedes Item kann als „beobachtet“ oder „nicht beobachtet“ bzw. mit ja oder nein bewertet werden. Die Gesamtauswertung aller Items beschreibt dabei die Aspekte Gewichtsbelastung, Körperhaltung und Bewegungen gegen die Schwerkraft. Es erfolgt die Bildung eines Gesamtwertes (0–58 Punkte) ohne unterschiedliche Gewichtung von Items. Ein höherer Wert steht für eine reifere grobmotorische Entwicklung.

Funktionsfähigkeit im Sinne der Entwicklung motorischer Meilensteine mittels HINE ist dabei nicht vergleichbar mit der Erfassung in den Studien CL-101 und CL-303.

Zusammenfassend liegt für die Studien PNCR, NeuroNEXT und ENDEAR keine mit den Studien CL-303 und CL-101 vergleichbare Endpunkterfassung für die motorischen Meilensteine vor.

CHOP INTEND

Der CHOP INTEND wurde in allen drei externen Kontrollstudien (PNCR, NeuroNEXT, ENDEAR) erfasst. In Studie PNCR wurde die zeitliche Veränderung (d. h. Steigungen) des CHOP-INTEND-Wertes mit Hilfe eines linearen Regressionsmodells mit gemischten Effekten mit zufälligen Slopes und Intercepts geschätzt [18]. Nach Angaben des pU im Nutzendossier ist die PNCR-Studie für einen validen Vergleich der Veränderung der CHOP-INTEND-Werte nicht geeignet, da in die Studie sowohl neu rekrutierte als auch bereits behandelte Patientinnen / Patienten eingeschlossen wurden. Weitere Informationen liegen nicht vor.

In der Studie NeuroNEXT wurde der CHOP INTEND zu jeder Visite (alle 3 Monate) erfasst, jedoch nur bei den Personen, die einen TIMPSI-Wert von weniger als 41 Punkten erreichten [25,26]. Die Auswertung der CHOP-INTEND-Werte basierte auf zeitlichen Veränderungen (d. h. Steigungen) und ist beschränkt auf Personen, für die Daten in einem Alter ≤ 12 Monate vorliegen, da im natürlichen Verlauf der Erkrankung in diesem Zeitintervall eine wesentliche Verschlechterung der motorischen Fähigkeiten auftritt. Dies umfasst in der Studie NeuroNEXT 10 von 16 Personen. Die Analyse basiert auf der Veränderung in den CHOP-INTEND-Werten von der ersten bis zur letzten vorhandenen Messung, wobei ausschließlich der letzte und der erste Datenpunkt für die Berechnung verwendet wurden.

In der Studie ENDEAR liegen Auswertungen für die Veränderung des CHOP-INTEND-Wertes vor, operationalisiert als Anteil an Personen mit einer Verbesserung. Ein Responder war definiert als eine Person mit einer Verbesserung von mindestens 4 Punkten im Verlauf der Studie. Darüber hinaus liegen Angaben über die Veränderung des CHOP INTEND über den gesamten Studienverlauf vor. Angaben zur monatlichen Änderung des CHOP-INTEND-Wertes und zur Anzahl an Patientinnen und Patienten, die einen Wert von über 40 Punkten in einem Alter von ≥ 6 Monaten erreicht haben, liegen nicht vor. Für die Studie ENDEAR war eine Beobachtungszeit von 13 Monaten geplant, aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht in etwa 6,2 Monaten.

Eine Vergleichbarkeit des Endpunktes zwischen den Studien PNCR, NeuroNEXT, ENDEAR und den Studien CL-101 und CL-303 ist nicht gegeben (u. a. Operationalisierung, Erhebungszeitpunkte, Umgang mit fehlenden Werten, Analysepopulation).

Hospitalisierung

In der Studie ENDEAR war die Hospitalisierung definiert als die Aufnahme in eine medizinische Einrichtung von mehr als 24 Stunden. Dabei wurde die Häufigkeit von Hospitalisierungen aufgrund von Symptomen nach Dosierung/Scheinintervention unter BSC, von SUE oder von zusätzlichen Untersuchungen (z. B. geplanter Eingriff) erfasst. „Überwachung zur allgemeinen Beobachtung“ schloss Behandlungen ein, die nicht im Zusammenhang mit der Studienmedikation/Scheinintervention unter BSC stehen oder aufgrund von UE/SUE erfolgten. „Dosierung/Scheinintervention unter BSC“ schloss Behandlungen aus Sicherheitsgründen ein (nicht aus bestimmten Gründen wie

z. B. UE oder SUE). „SUE“ schloss Behandlungen aufgrund eines SUE ein. „Zusätzliche Untersuchungen“ schloss geplante Behandlungen ein, wie z. B. das Legen einer Magensonde aus präventiven Gründen und nicht aufgrund eines UE [19,21]. Demgegenüber erfolgte die Erfassung der Hospitalisierung in den zulassungsbegründenden Studien ausschließlich unter SUE, jedoch fehlen Angaben zum Mindestaufenthalt in der Einrichtung und darüber, ob der Aufenthalt im Zusammenhang mit der Erkrankung oder einem UE steht. Darüber hinaus bleibt offen, ob auch geplante Krankenhausaufenthalte mit in die Bewertung eingingen, wie in der Studie ENDEAR. Es kann abschließend nicht beurteilt werden, ob Gründe, Anzahl und Dauer von Krankenhausaufenthalten zwischen den Zulassungsstudien und der Studie ENDEAR vergleichbar sind. Darüber hinaus war in der Studie ENDEAR eine Beobachtungszeit von 13 Monaten geplant, aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht in etwa 6,2 Monaten.

Eine Vergleichbarkeit des Endpunktes zwischen der Studie ENDEAR und den Studien CL-303 und CL-101 ist nicht gegeben (u. a. Operationalisierung, Analysepopulation).

2.4 Statistische Methoden

2.4.1 Studie CL-303

Die nachfolgenden Ausführungen basieren auf SAP-Version 3.0 (Stand: 21. Dezember 2017). Diese wurde nach Einschluss der ersten Person in die Studie am 24. Oktober 2017 veröffentlicht.

Analysepopulation

Angaben zu den Analysepopulationen dieser Studie sind Tabelle 14 zu entnehmen.

Fallzahlplanung

Es war geplant bis zu 20 Personen in die Studie aufzunehmen. Auf Basis der Phase-I-Studie (CL-101), wonach mindestens 40 % der behandelten Personen den ersten ko-primären Wirksamkeitsendpunkt (Sitzen ohne Unterstützung für mindestens 30 Sekunden in einem Alter von 18 Monaten) erreichten sowie unter Annahme einer wahren Ansprechrate von 30–40 %, wäre eine Fallzahl von 15 Personen (Annahme: 30 % Drop-out-Rate) ausreichend, um mit einer Power von > 90 % einen signifikanten Unterschied zu 0,1 % zu zeigen ($\alpha = 0,025$, einseitiger exakter Test mittels Binomialverteilung).

Für den zweiten ko-primären Endpunkt (Überleben bis zu einem Alter von 14 Monaten) wird angenommen, dass 75 % der Patientinnen und Patienten in der PNCR-Kohorte ein Alter von 13,6 Monaten nicht erreichen und somit eine 25%ige Überlebensrate aufweisen. Vorläufige Ergebnisse aus der Phase-I-Studie CL-101 lassen annehmen, dass nach der Behandlung 80 % der Patientinnen und Patienten in der ITT-Population für eine entsprechende Zeit überleben. Mit diesem Wirksamkeitsunterschied würde eine Fallzahl von 15 Personen, die den ITT-Kriterien entsprechen (Annahme: 30 % Drop-out-Rate) ausreichen, um mit einer Power von > 80 % einen signifikanten Unterschied im Vergleich zur PNCR-Kohorte zu detektieren ($\alpha = 0,05$ zweiseitiger exakter Test nach Fisher).

Statistische Analysen

Alle Hypothesen wurden einseitig zum Signifikanzniveau von $\alpha = 0,025$ oder zweiseitig zum Signifikanzniveau $\alpha = 0,05$ getestet.

Kategoriale Variablen, so wie der Anteil der Personen ohne Ereignis, wurden mit Anzahl und Prozentwert zusammenfassend dargestellt. Für kontinuierliche Daten (z. B. CHOP INTEND) wurden Mittelwert, Median, Standardabweichung (SD) sowie Minimum und Maximum angegeben.

Wirksamkeitsanalysen werden in der ITT- und der Efficacy-Completer-Population durchgeführt. Die auf der ITT-Population basierenden Analysen werden als Primäranalyse betrachtet. Sicherheitsanalysen wurden in der Sicherheitspopulation durchgeführt.

Die primären und sekundären Endpunkte wurden hierarchisch getestet: 1. funktionelles unabhängiges Sitzen (erster ko-primärer Endpunkt); 2. ereignisfreies Überleben (zweiter ko-primärer Endpunkt); 3. Erhaltung der Gedeihfähigkeit (erster ko-sekundärer Endpunkt); 4. frei von ventilatorischer Unterstützung (zweiter ko-sekundärer Endpunkt). Nur wenn die primären Endpunkte die Signifikanz erreichen, werden die sekundären Endpunkte getestet.

Endpunkt: Sitzen ohne Unterstützung in einem Alter von 18 Monaten

Für den ersten ko-primären Endpunkt (Sitzen ohne Unterstützung in einem Alter von 18 Monaten) ergeben sich folgende statistische Hypothesen:

- $H_0: p_{\text{Zolgensma}} = 0,1\%$
- $H_1: p_{\text{Zolgensma}} > 0,1\%$

Dabei ist p der Anteil der Personen, die den motorischen Meilenstein des funktionell unabhängigen Sitzens ≥ 30 Sekunden bis zu einem Alter von 18 Monaten erreicht haben ($\alpha = 0,025$; einseitiger exakter Test mittels Binomialverteilung).

Die Analyse dieses Endpunkts erfolgt in der ITT-Population. Zusätzlich werden für Patientinnen und Patienten, die das funktionelle unabhängige Sitzen im Alter von 18 Monaten erreicht haben, das Alter, in dem die Probanden das erste Mal 30 Sekunden lang unabhängig sitzen konnten, sowie die Zeit bis zum Erreichen des Meilensteins zusammengefasst.

Endpunkt: Überleben bis zu einem Alter von 14 Monaten

Für den zweiten ko-primären Endpunkt (Überleben in einem Alter von 14 Monaten) ergeben sich folgende statistische Hypothesen

- $H_0: p_{\text{Zolgensma}} = \text{PNCr}$
- $H_1: p_{\text{Zolgensma}} \neq \text{PNCr}$

Dabei ist p der Anteil der Personen, die im Alter von 14 Monaten überlebt haben ($\alpha = 0,05$; zweiseitiger exakter Test nach Fisher).

Die Analyse dieses Endpunkts erfolgt in der ITT-Population. Als Vergleichsgruppe schätzten Finkel et al. (2014 a) [18] in einer natürlichen Verlaufskohorte (PNCr) an Patientinnen und Patienten mit infantiler SMA (Typ 1), dass nur ein Anteil von 25 % mit zwei SMN2-Genkopien ereignisfrei bis zum Alter von 13,6 Monaten überleben würde.

Zusätzlich wird der Anteil der Personen dargestellt, die jedes der folgenden Ereignisse bis zum Alter von 14 Monaten ereignisfrei erleben:

- Tod
- Dauerhafte Beatmung

Explorative Endpunkte: Motorische Meilensteine, Bayley-Skala, CHOP INTEND

Dargestellt werden Anzahl und Anteil der Personen, die zum Alter von 18 Monate die motorischen Meilensteine erreichen. Zusätzlich wurde das Alter beim erstmaligen Erreichen der motorischen Meilensteine angegeben.

Dargestellt werden für Bayley-Skala und CHOP INTEND Anzahl und Anteil der Patientinnen und Patienten zu jeder Visite und die Veränderung zum Alter von 18 Monate im Vergleich zu Baseline. Für CHOP INTEND werden zusätzlich Anzahl und Anteil der Patientinnen und Patienten dargestellt, die einen CHOP-INTEND-Schwellenwert von ≥ 40 , ≥ 50 und ≥ 58 zu irgendeinem Zeitpunkt innerhalb der Beobachtungsdauer erreichten. Personen, die bereits zu Baseline jeweils einen Wert ≥ 40 , ≥ 50 und ≥ 58 Punkte aufwiesen, werden in dieser Analyse nicht berücksichtigt.

Endpunkt: Hospitalisierung

Keine Angaben im SAP.

Umgang mit fehlenden Werten

Für die ko-primären und ko-sekundären Endpunkte wurden fehlende Werte als Non-Responder imputiert.

Für kombinierte Endpunkte (z. B. Gedeihfähigkeit) wurde bei Fehlen von Teilkomponenten des Endpunkts zur jeweiligen Visite die Methode „Last Observation Carried Forward“ (LOCF) mit gemischten Komponenten verwendet – einer Kombination aus Endpunktwerten zu Baseline und aus der letzten Visite.

Für explorative Endpunkte, kontinuierliche und geordnet-kategoriale Endpunkte, die erst im Alter von 18 Monaten (und zu Baseline) erhoben werden sollen, werden Werte genutzt, die bei vorzeitigen Studienabbruchvisiten erfasst wurden, um fehlende Werte im Alter von 18 Monaten zu imputieren.

Für CHOP INTEND und Bayley-Skala erfolgen die Berechnungen der Werte nach Manual, wenn einzelne Fragen/Items fehlen. Wenn diese Regeln für die Berechnung eines Wertes nicht ausreichen, wird der Endpunkt als fehlend betrachtet und nicht imputiert.

Darüber hinaus werden fehlende Werte für Sicherheitsendpunkte nicht imputiert.

Subgruppenanalysen

Siehe Tabelle 3.

2.4.2 Studie CL-101

Für die Phase-I-Studie wurde kein SAP erstellt, da auf Empfehlung der amerikanischen Zulassungsbehörde (FDA) deskriptive Analysen für die Wirksamkeits- und Sicherheitsendpunkte empfohlen wurden. Protokollversion 14.0 (21. April 2016) stellt die Basis für die definierten Ziele und Endpunkte der Studie dar. Die letzte Protokollversion 15.0 (21. November 2016) wurde auf das von der FDA erhaltene Feedback zurückgenommen.

Datenschnitt

Der Datenschnitt für die primäre Wirksamkeitsanalyse erfolgte am 20. Januar 2017. Zu diesem Zeitpunkt hatten alle Patientinnen und Patienten die Studiervisite nach 13,6 Monaten erreicht. Zum Datenschnitt 7. August 2017 hatte die zuletzt eingeschlossene Person in die Studie das Alter von 20 Monaten erreicht. Die 24-monatige Nachbeobachtungsphase war zum Datenschnitt 14. Dezember 2017 abgeschlossen. Die ersten zwei Datenschnitte dienen dem Vergleich mit der PNCR-Kohorte (historische Kontrolle). Der Vergleich erfolgte unter der Annahme, dass 25 % der Personen mit infantiler SMA (Typ 1) mit zwei SMN2-Genkopien bis Lebensmonat 13,6 und 8 % bis Lebensmonat 20 ereignisfrei überleben. Zusätzlich war für die Personen eine Langzeitbeobachtung von bis zu 15 Jahren geplant (Studie LT-001).

Analysepopulation

Angaben zu den Analysepopulationen dieser Studie sind Tabelle 14 zu entnehmen.

Fallzahlplanung

Eine formale Fallzahlplanung wurde nicht durchgeführt.

Statistische Analysen

Es handelt sich um eine einarmige Studie, Daten wurden deskriptiv ausgewertet. Alle getesteten Hypothesen wurden zum Signifikanzlevel von $\alpha = 0,05$ getestet, mit Ausnahme der Analyse des Überlebens ($\alpha = 0,025$).

Kategoriale Variablen, so wie der Anteil der Personen ohne Ereignis, wurden mit Anzahl und Prozentwert zusammenfassend dargestellt. Für kontinuierliche Daten wurden Mittelwert, Median, SD sowie Minimum und Maximum angegeben.

Wirksamkeitsanalysen wurden für die FAS- und ITT-Analysepopulationen durchgeführt. Sicherheitsanalysen wurden in der Sicherheitspopulation durchgeführt.

Endpunkt: Überleben

Die primäre Wirksamkeitsanalyse erfolgte nach Erreichen eines Patientenalters von 13,6 Monaten zum Datenschnitt 20. Januar 2017 und nach Erreichen eines Alters von 20 Monaten für die zuletzt eingeschlossene Person zum Datenschnitt 7. August 2017. Patientinnen und Patienten, die die Studie vorzeitig aufgrund jeglicher Ursache abbrachen, wurden als Ereignis gewertet.

Für den Endpunkt Überleben wurde angenommen, dass 25 % der Patientinnen und Patienten mit infantiler SMA (Typ 1) und zwei SMN2-Genkopien bis zum Alter von 13,6 Monaten und 8 % bis zum Alter von 20 Monaten überleben [18]. Die Analysen erfolgten in der ITT-Population.

Endpunkt: Erreichen motorischer Meilensteine

Dargestellt werden Anzahl und Anteil der Patientinnen und Patienten, die zum Zeitpunkt der Analysen die motorischen Meilensteine erreichen.

Endpunkt: Bayley-Skala

Die Bayley-Skala (Grob- und Feinmotorik) wurde als explorativer Endpunkt für Personen festgelegt, die einen CHOP-INTEND-Wert von ≥ 60 erreichen. Die Veränderungen gegenüber Baseline für die fein- und grobmotorischen Subskalen wurden im FAS ausgewertet.

Endpunkt: CHOP INTEND

Die Veränderung des CHOP-INTEND-Wertes gegenüber dem Baseline-Wert wurde mit Hilfe eines gemischten Modells für wiederholte Messwerte (Mixed Model of Repeated Measures, MMRM) analysiert. Das Modell für das FAS beinhaltete die Veränderung gegenüber dem Baseline-Wert als abhängige Variable sowie fixe Effekte (Kohorte, Visite), Kovariate (Baseline-Wert) und Interaktion (Kohorten-Visiten, Baseline-Visiten). Zunächst wurde eine unstrukturierte (allgemeine) Kovarianzstruktur angenommen, um die Fehler innerhalb einer Person zu modellieren; wenn jedoch die unstrukturierte Kovarianz zu Nichtkonvergenz führte, wurde die Varianzkomponente verwendet. Die kleinsten Quadrate (LS-Mittelwerte), die Differenzen zwischen den LS-Mittelwerten, zweiseitige 95 % Konfidenzintervalle (KI) für jede Differenz und die p-Werte aus den Modelleffekten wurden für jede geplante Visite berichtet. Darüber hinaus wurde der Anteil an Personen, die einen CHOP-INTEND-Schwellenwert von ≥ 40 , ≥ 50 und ≥ 60 zum Zeitpunkt der primären Wirksamkeitsanalyse sowie nach 24 Monaten erreichten, für das FAS ausgewertet (explorativer Endpunkt).

Endpunkt: Hospitalisierung

Annualisierte Hospitalisierungsrate / Anzahl der Hospitalisierungen pro Jahr: Für jede Person wurde die Anzahl der Hospitalisierungen durch die Anzahl der Tage in der Studie geteilt, um die Hospitalisierungen pro Tag zu berechnen, dann auf 365 Tage standardisiert.

Umgang mit fehlenden Werten

Tod oder dauerhafte Beatmung: Personen, die vor dem Datenschnitt aus der Studie ausschieden ohne dass dauerhafte Beatmung notwendig war, wurden als Ereignis gewertet.

CHOP INTEND: In Fällen, in denen eine erhebliche, systemische und äußerliche Einschränkung vorhanden war, welche die akkurate Bewertung mehrerer Items verhinderte, wurde der Gesamtwert als fehlend eingestuft. Ein vollständiger Wert wurde nicht berechnet, wenn ein oder mehrere Items als „nicht abgefragt“ bewertet werden konnten.

Für andere Endpunkte wurden keine Zensierungen aufgrund fehlender Werte vorgenommen.

Subgruppenanalysen

Es waren keine Subgruppen präspezifiziert.

2.4.3 Studien CL-302 und CL-304

Bei den Studien CL-302 und CL-304 handelt es sich um noch laufende, einarmige, multizentrische Phase-III-Studien. Die finalen Ergebnisse für Studie CL-302 sind für das 4. Quartal 2020 (2 SMN2-Genkopien) bzw. das 2. Quartal 2021 (3 SMN2-Genkopien) geplant und für Studie CL-304 für das 3. Quartal 2020.

Vor dem Hintergrund, dass für die vorliegende Nutzenbewertung ausschließlich nicht a priori geplante Datenschnitte vorliegen, werden die Ergebnisse dieser Studien für die Nutzenbewertung ergänzend für die Endpunkte „Mortalität“, „Tod oder dauerhafte Beatmung“ und die aggregierten Sicherheitsendpunkte herangezogen.

Für Studie CL-302 liegt die SAP-Version 2.1 vom 9. Januar 2018 vor. Vorherige Versionen des SAP wurden nicht eingereicht; auch finden sich in den Unterlagen keine Angaben über relevante Änderungen zu den vorherigen SAP-Versionen. Für Studie CL-304 liegt die SAP-Version 1.0 vom 18. Dezember 2017 vor.

Analysepopulation

Angaben zu den Analysepopulationen dieser Studie sind Tabelle 14 zu entnehmen.

2.4.4 Indirekter Vergleich der Studien CL-303 und CL-101 vs. externe Kontrolle (PNCR, NeuroNEXT, ENDEAR)

Aufgrund des einarmigen Designs der zulassungsbegründenden Studien CL-101 und CL-303 wurden vom pU nicht-adjustierte indirekte Vergleiche (ohne Brückenkomparator) in den Endpunktkategorien Mortalität, Morbidität und Sicherheit vorgelegt. Hierfür wurden folgende Studien zum natürlichen Verlauf der SMA-Erkrankung eingereicht: PNCR [18], NeuroNEXT [25,26] und der BSC-Studienarm der ENDEAR-Studie zu Nusinersen [19,21].

Eine differenzierte Betrachtung der externen Kontrollen erfolgte bereits in Abschnitt 2.2.2, bezogen auf die Studiencharakteristik (u. a. Ein- und Ausschlusskriterien), und in Abschnitt 2.3.5, bezogen auf die Vergleichbarkeit der Endpunkte.

Die nachfolgenden Angaben zur statistischen Auswertung basieren primär auf den Angaben im Bericht RPT-806 und dem Nutzendossier (Modul 4). Ein SAP für den indirekten Vergleich der Zulassungsstudien mit den externen Kontrollen wurde nicht eingereicht.

Der pU sieht die Voraussetzung der Vergleichbarkeit der Studienpopulationen anhand der definierten Einschlusskriterien (siehe Abschnitt 2.2.2) für einen indirekten Vergleich als erfüllt an. Um dies zu bestätigen führt der pU Analysen anhand individueller Patientendaten für die Studie CL-101 im Vergleich mit den externen Kontrollen PNCR und NeuroNEXT durch. So wurden alle 12 Personen der Studie CL-101 ausgewählt und für diese die am besten passenden Kontrollen aus den Studien PNCR und NeuroNEXT anhand bestimmter Baselinekriterien (bialele SMN1-Deletion, 2 SMN2-Genkopien, Alter bei Krankheitsbeginn, Ernährungs- und Beatmungsunterstützung im Alter von 6 Monaten und Gesamtwert auf der CHOP-INTEND-Skala zu Baseline) gematcht. Trotz dieser Matching-Vorgehensweise bestehen weiterhin Unterschiede innerhalb der gebildeten Matching-Paare (CL-101 vs. PNCR, CL-101 vs. NeuroNEXT) bezogen auf u. a. CHOP-INTEND-Gesamtwert zu Baseline sowie Ernährungs- und Beatmungsunterstützung. Die bestehenden Unterschiede im Matching lassen sich nicht anhand der vorliegenden Studienunterlagen

beurteilen. Eine Erklärung hierfür fehlt und wird vom pU weder klinisch noch inhaltlich diskutiert. Darüber hinaus scheint für die verbliebenen Patientinnen und Patienten in den Studien PNCR und NeuroNEXT (PNCR: N = 11; NeuroNEXT: N = 4) eine Vergleichbarkeit angenommen zu werden, da diese Personen ebenfalls für die Auswertungen berücksichtigt werden.

Angaben zu den durchgeführten Endpunktauswertungen für die nicht-adjustierten indirekten Vergleiche sind primär dem Nutzendossier (Modul 4) zu entnehmen. Demnach wurden für binäre Endpunkte Effektschätzer (Relatives Risiko (RR), Odds Ratio (OR) und Risikodifferenz (RD)) berechnet. Für die Berechnungen der zugehörigen p-Werte wurde ein exakter Test gewählt (Fisher's exakter Test). Für berechnete Mittelwertdifferenzen wurde zur abschließenden Einschätzung der klinischen Relevanz zusätzlich noch Hedges' g bestimmt. Es wurden ausschließlich nicht-adjustierte Effektschätzer ohne Berücksichtigung der unterschiedlichen Beobachtungszeiten vorgelegt. Nach Angaben im Nutzendossier (Modul 4) wurde für einen Vergleich der frühere Vergleichszeitpunkt 13–14 Monate in den Studien CL-101 und CL-303 gewählt, um einen adäquateren Vergleich mit den Studien zum natürlichen Verlauf zu ermöglichen (14 Monate in den Studien PNCR und NeuroNEXT sowie finale Daten in der Studie ENDEAR mit einer maximalen Beobachtungszeit von 13 Monaten).

Für CHOP INTEND wurde beschrieben, dass dieser in der Studie NeuroNEXT zu jeder Visite (alle 3 Monate) erfasst wurde, jedoch nur bei den Personen, die im TIMPSI einen Wert von weniger als 41 Punkten erreichten [25,26]. Die Auswertung der CHOP-INTEND-Werte basierte auf zeitlichen Veränderungen (d. h. Steigungen), wenn mindestens zwei verfügbare Werte vorlagen. Die Analyse ist beschränkt auf Personen, für die Daten im Alter ≤ 12 Monate vorliegen. Die Analyse basiert auf der Veränderung der CHOP-INTEND-Werte von der ersten bis zur letzten vorhandenen Messung. Die Veränderung pro Zeiteinheit wurde mit den beobachteten Tagen zwischen den Datenpunkten berechnet und durch Multiplikation mit 30 auf einen Wert pro Monat hochgerechnet [7].

Für die Sicherheitsendpunkte lagen aus den Studien PNCR und NeuroNEXT keine Daten vor.

Für die Studie ENDEAR liegen keine näheren Informationen vor. Hier kann ausschließlich auf die Angaben im Nutzendossier (Modul 4) verwiesen werden, da es sich um eine vergleichende Post-hoc-Auswertung handelt.

Die Darstellung der nicht-adjustierten indirekten Vergleiche erfolgte im Nutzendossier für die gepoolte Studienpopulation der Studien PNCR und NeuroNEXT separat gegenüber der Studien CL-101 und CL-303 als auch gegenüber der gepoolten Studienpopulation der Zulassungsstudien. Dieses Poolen war a priori nicht geplant und wird insbesondere vor dem Hintergrund der Unterschiede zwischen den Studienpopulationen (siehe Abschnitt 2.2.2) und der Operationalisierung der Endpunkte (siehe 2.3.5) kritisch gesehen.

Tabelle 14: Relevante Analysepopulationen in den für die Nutzenbewertung berücksichtigten Studien

Analysepopulation	CL-101	CL-303	CL-302	CL-304
ITT-Population	Alle Personen, die eine Gentherapie-Infusion erhielten.	Alle Personen, die im Alter von höchstens 180 Tagen eine Gentherapie-Infusion erhielten.	Alle Personen, die im Alter von höchstens 180 Tagen eine Gentherapie-Infusion erhielten.	Alle Personen mit bialleler Deletion von SMN1 und 2–4 Genkopien des SMN2-Gens ohne bekannte Genmodifikator-Mutation (c.859G>C).
FAS	Alle Personen, die eine Gentherapie-Infusion erhielten und mindestens eine Visite nach der Gentherapie-Infusion hatten.	n. a.	n. a.	n. a.
Sicherheitspopulation	Entspricht der ITT-Population.	Alle Personen, die eine Gentherapie-Infusion erhielten.	Alle Personen, die eine Gentherapie-Infusion erhielten.	Alle Personen, die eine SMN1-Punktmutation und eine bekannte Genmodifikator-Mutation (c.859G>C) aufwiesen.

Abkürzungen: FAS: Full Analysis Set; ITT: Intention-to-Treat; n. a.: nicht adressiert; SMN: Survival Motor Neuron.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Die Studien CL-303 und CL-101 sind beendet und werden für die Nutzenbewertung herangezogen. Ergänzend werden in der Nutzenbewertung die Studien LT-001 (Langzeit-Sicherheitsstudie von CL-101), CL-302 und CL-304 für die Endpunkte „Mortalität“, „Tod oder dauerhafte Beatmung (Überleben)“ sowie die aggregierten Sicherheitsendpunkte herangezogen.

Da es sich bei den genannten Studien um unkontrollierte Studien handelt, wird von einem hohen Verzerrungspotential auf Studien- und Endpunktebene ausgegangen.

Indirekter Vergleich der Studien CL-303 und CL-101 vs. externe Kontrolle (PNCr, NeuroNEXT, ENDEAR)

Aufgrund des einarmigen Studiendesigns der Studien CL-303 und CL-101 legt der pU nicht-adjustierte indirekte Vergleiche für externe Kontrollpopulationen der Studien PNCr, NeuroNEXT und ENDEAR (BSC-Arm) vor. Die externen Kontrollen werden in den Abschnitten 2.2.2 (Einschlusskriterien), 2.3.5 (Endpunkte) und 2.4.4 (Statistik) ausführlich beschrieben. Der indirekte Vergleich der Studien CL-303 und CL-101 mit den externen Kontrollen war nicht a priori geplant, jedoch dienten Referenzwerte (Cut-offs) aus der Studie PNCr für die Hypothesenbildung und den Vergleich bestimmter Endpunkte (u. a. Mortalität) in den Zulassungsstudien (siehe Abschnitt 2.4). Für den Vergleich der Studien CL-303 und CL-101 mit den externen Kontrollen wurde vom pU ein separater Bericht (RPT-806) eingereicht [7]. Jedoch finden sich in diesem keine Angaben zu der a priori festgelegten Auswertungsmethodik der Zulassungsstudien mit den externen Kontrollen. Ein entsprechender SAP liegt nicht vor. Die Darstellung der Ergebnisse und der Auswertungsmethodik für die nicht-adjustierten indirekten Vergleiche sind primär dem Nutzendossier (Modul 4) zu entnehmen.

Bezogen auf die Einschlusskriterien der Zulassungs- und externen Kontrollstudien konnten bereits Unterschiede (z. B. Alter bei Symptombeginn, vgl. 2.2.2) aufgezeigt werden, die insbesondere für das Poolen der Studienpopulationen PNCr+NeuroNEXT bei der Endpunktauswertung, wie z. T. im Nutzendossier (Modul 4) vorgenommen, kritisch gesehen werden. Während in den Studien CL-101, PNCr und ENDEAR ausschließlich symptomatische Patientinnen und Patienten eingeschlossen wurden, konnten in die Studien CL-303 und NeuroNEXT auch asymptomatische Patientinnen und Patienten aufgenommen werden. Darüber hinaus waren in der Studie NeuroNEXT im Vergleich zu den Studien PNCr und ENDEAR Patientinnen und Patienten ausgeschlossen, die mehr als 12 Stunden/Tag eine nicht-invasive ventilatorische Unterstützung benötigten (z. B. BiPAP). In den externen Kontrollen finden sich keine Angaben zur erlaubten und nicht-erlaubten Begleitmedikation. Lediglich in der Studie PNCr finden sich Angaben darüber, dass vier Personen begleitend Valproinsäure einnahmen. Die in Abschnitt 2.4.4 beschriebene Matching-Vorgehensweise der Studie CL-101 vs. die Studien PNCr bzw. NeuroNEXT konnte die identifizierten Unterschiede (u. a. Alter bei Krankheitsbeginn, Ernährungs- und Beatmungsunterstützung, motorische Funktion anhand des CHOP-INTEND-Werts) nicht ausgleichen, sodass Unsicherheiten bezogen auf die Strukturgleichheit zwischen den Populationen der Studien bestehen bleiben und eine Verzerrung nicht ausgeschlossen werden kann. Für die Studien CL-303 und ENDEAR wurde kein Matching durchgeführt, weshalb die Strukturgleichheit nicht beurteilt werden kann. Da bei vergleichenden nicht-randomisierten Studien grundsätzlich nicht von einer Strukturgleichheit der Gruppen ausgegangen werden kann (sondern eine hohe Wahrscheinlichkeit für systematische Unterschiede zwischen den Gruppen besteht), sind insbesondere relevante Einflussfaktoren zu berücksichtigen. Eine entsprechende Confounder-Adjustierung wurde jedoch nicht durchgeführt.

Vor diesem Hintergrund der in den Abschnitten 2.2.2, 2.3.5, und 2.4.4 dargelegten inhaltlichen und methodischen Limitationen, werden die nicht-adjustierten indirekten Vergleiche als nicht ausreichend valide erachtet und als hoch verzerrt eingestuft. Ergebnisse werden aufgrund der Größenordnung der Effektschätzer (Grenze 5–10 Punkte) in der vorliegenden Nutzenbewertung für die Endpunkte „Mortalität“ und „Tod oder dauerhafte Beatmung (Überleben)“ sowie die Teilkomponente „Dauerhafte Beatmung“ berichtet. Da es sich bei diesen Endpunkten um objektiv bewertete Endpunkte handelt und diese hinsichtlich ihrer Operationalisierung weitestgehend vergleichbar sind, werden die Ergebnisse trotz der gravierenden methodischen Limitationen ergänzend dargestellt.

Es liegen ausschließlich für Mortalität vergleichbare Operationalisierungen für alle Zulassungs- und externen Kontrollstudien vor. Hierbei ist zu berücksichtigen, dass sich die Beobachtungszeiten in den einzelnen Studien unterscheiden. So sind in Studie CL-303 die Personen bis zu einem Alter von mindestens 18 Monaten und in Studie CL-101 bis 24 Monate nach Gentransfer beobachtet worden. Für die Studien PNCR und NeuroNEXT finden sich nur Angaben über die geplante Beobachtungsdauer (PNCR bis zu 36 Monate, NeuroNEXT bis zu 24 Monate). Angaben zur tatsächlichen medianen Beobachtungsdauer liegen sowohl für die Zulassungs- als auch für die externen Kontrollstudien nicht vor. Für die Vergleichbarkeit der externen Kontrollen mit den Zulassungsstudien CL-101 und CL-303 wurde vom pU das Alter 13–14 Monate als Zeitpunkt für die Auswertung festgelegt. Zusätzlich erfolgte die Auswertung für den gesamten Beobachtungszeitraum. Letztere wird aufgrund unterschiedlicher Beobachtungszeiten und fehlender Beobachtungszeit-adjustierter Effektschätzer als hoch verzerrt eingeschätzt und in der vorliegenden Nutzenbewertung nicht berücksichtigt. Darüber hinaus sei angemerkt, dass in der Studie ENDEAR zwar eine Beobachtungszeit von 13 Monaten geplant war, jedoch wurde die Studie vorzeitig beendet, sodass die Studiendauer im BSC-Arm im Median 187 Tage (ca. 6,2 Monate) betrug. Auch hier liegt kein geeigneter Effektschätzer vor, welcher die unterschiedlichen Beobachtungszeiten berücksichtigt.

Für den kombinierten Endpunkt „Tod oder dauerhafte Beatmung (Überleben)“ liegen unterschiedliche Operationalisierungen für die Teilkomponente „Dauerhafte Beatmung“ vor. Keine mit den Studien CL-303, CL-101 und PNCR vergleichbare Operationalisierung lag in den Studien NeuroNEXT und ENDEAR vor. Der pU stellt für den Endpunkt im Nutzendossier ausschließlich die gepoolten Ergebnisse der Studien PNCR+NeuroNEXT im Vergleich zu den Zulassungsstudien dar, welche als verzerrt angesehen werden. Die Auswertung erfolgte ebenfalls zum Alter 13–14 Monate und für den gesamten Beobachtungszeitraum. Letztere wird aufgrund der unterschiedlichen Beobachtungszeiten und fehlender Beobachtungszeit-adjustierter Effektschätzer als hoch verzerrt eingeschätzt und in der vorliegenden Nutzenbewertung nicht berücksichtigt. Wie bereits oben unter dem Endpunkt „Mortalität“ dargelegt, betrug die tatsächliche Studiendauer im BSC-Arm im Median nur 187 Tage (ca. 6,2 Monate). Auch für diesen Endpunkt liegen keine geeigneten Effektschätzer vor, welcher die unterschiedlichen Beobachtungszeiten berücksichtigt, weshalb von einem hohen Verzerrungspotential auf Endpunktebene ausgegangen wird.

3 Ergebnisse der eingeschlossenen Studien

3.1 Studiencharakteristika und Studienmedikation

Studie CL-101 (LT-001), CL-303, CL-302 und CL-304

Von Relevanz für die vorliegende Nutzenbewertung sind die Zulassungsstudien CL-101 und CL-303. Die Extensionsstudie LT-101 sowie die Studien CL-302 und CL-304 werden ergänzend dargestellt; die Studien sind laufend, es liegen keine abschließenden Ergebnisse vor.

In die Zulassungsstudien waren Patientinnen und Patienten mit infantiler SMA (Typ 1) mit einem Durchschnittsalter von 3,7 bzw. 3,4 Monaten (CL-303 und CL-101) eingeschlossen. In Studie CL-303 wiesen alle 22 Studienteilnehmenden 2 SMN2-Genkopien sowie eine biallelische Deletion des SMN1-Gens auf, in Studie CL-101 wurden nur Patientinnen und Patienten mit 2 SMN2-Genkopien aufgenommen. In beiden Studien konnten Patientinnen und Patienten, entsprechend der Einschlusskriterien, mit biallelischer Mutation des SMN1-Gens und ein oder zwei SMN2-Genkopien in die Studien aufgenommen werden. Das mediane Gestationsalter lag in beiden Studien bei 39,0 Wochen. Der Anteil weiblicher Kleinkinder beträgt in Studie CL-303 55 % und in Studie CL-101 58 %, der Anteil der männlichen entsprechend 45 % und 42 %. Alle Studienteilnehmenden erhielten begleitend Prednisolon über eine Dauer von durchschnittlich 74 Tagen (CL-303 und CL-101). In beiden Studien wiesen nahezu alle Patientinnen und Patienten zu Studienbeginn die SMA-spezifischen Symptome Hypotonie sowie Schwäche der Gliedmaßen auf. In Studie CL-303 konnten alle Patientinnen und Patienten dünnflüssige Nahrung schlucken, in Studie CL-101 waren es nur 33,3 % der Studienpopulation. In Studie CL-101 erhielten 41,7 % der Studienpopulation eine nicht-orale Nahrungsunterstützung, in Studie CL-303 war niemand davon betroffen. Bei einer Person in Studie CL-101 war zu Studienbeginn zudem eine Beatmungsunterstützung erforderlich. Von den 22 in Studie CL-303 eingeschlossenen Patientinnen und Patienten schlossen 19 die Studie ab, drei Personen brachen die Studie vorzeitig ab, aufgrund von UE, Tod oder Entzug der Einverständniserklärung. Die mediane Studiendauer pro Patientin / Patient ist in beiden Studien nicht bekannt. In Studie CL-101 schlossen alle 12 eingeschlossenen Patientinnen und Patienten die Studie ab.

Tabelle 15: Allgemeine Angaben für die Studien CL-101, CL-303, CL-302, CL-304 und LT-001

Studie Allgemeine Angaben	CL-303 ¹⁾ N = 22	CL-101 ²⁾³⁾ N = 12	CL-302 ⁴⁾ N = 33	CL-304 ⁵⁾		LT-001 ⁶⁾ N = 10
				SMN2: 2 Kopien N = 14	SMN2: 3 Kopien N = 15	
Personen in die Studie eingeschlossen	22	12	33	14	15	10
ITT-Population, n (%)	22 (100)	12 (100)	32 (97,0)	14 (100,0)	15 (100,0)	10 (100)
Sicherheitspopulation, n (%)	22 (100)	12 (100)	33 (100,0)	14 (100,0)	15 (100,0)	10 (100)
Studie abgeschlossen, n (%)	19 (86,4)	12 (100)	19 (57,6)7)	1 (7,1)	0	0
Studie abgebrochen, n (%)	3 (13,6)	0	1 (3,0)	0	0	0
Unterstützende Beatmung	-		0			
Aufgrund von UE	1 (4,5)		0			
Tod	1 (4,5)		1 (3,0)			
Entzug der Einverständniserklärung	1 (4,5)		0			
Andere	-		0			
Protokollverletzungen	0 (0,0)	k. A.	k. A.	k. A.	k. A.	k. A.
Studiendauer (Monate)	k. A.	k. A.				k. A.
MW (SD)			10,6 (3,4)	10,5 (5,1)	8,7 (3,3)	
Median (min; max)			11,9 (1,8;15,4)	9,9 (5,1;18)	9,0 (2;13,9)	

¹⁾ CL-303: Finaler Datenschnitt zum 12.12.2019.

²⁾ CL-101: Finaler Datenschnitt zum 14.12.2017.

³⁾ CL-101: Ergebnisse dargestellt für Kohorte 2 mit der Dosierung 2,0 x 10¹⁴vg/kg.

⁴⁾ CL-302: Datenschnitt: 31.12.2019; laufende Studie.

⁵⁾ CL-304: Datenschnitt 31.12.2019; laufende Studie.

⁶⁾ LT-001: Datenschnitt 31.12.2019; laufende Studie. Angaben zu den 10 Personen, die aus Studie CL101 in die Extensionsstudie LT-001 aufgenommen wurden.

⁷⁾ Anzahl bezogen auf die primäre Analyse nach 14 Monaten.

Abkürzungen: ITT: Intention-to-Treat; k. A.: keine Angabe; MW: Mittelwert; SD: Standardabweichung; SMN: Survival Motor Neuron; UE: Unerwünschte/s Ereignis/se; vg: Vektorgenom/e.

Tabelle 16: Charakterisierung der Studienpopulationen zu Studienbeginn: Studien CL-101, CL-303, CL-302, CL-304 und LT-001 – Sicherheitspopulationen

Studien Merkmal	CL-303 ¹⁾ N = 22	CL-101 ²⁾³⁾ N = 12	Studie CL-302 ⁴⁾ Kohorte 2 N = 33	Studie CL-304 ⁵⁾		LT-001 ⁶⁾ N = 10
				SMN2: 2 Kopien N = 14	SMN2: 3 Kopien N = 15	
<i>Geschlecht, n (%)</i>						
weiblich	12 (54,5)	7 (58,3)	19 (57,6)	10 (71,4)	9 (60,0)	5 (50)
männlich	10 (45,5)	5 (41,7)	14 (42,4)	4 (28,6)	6 (40,0)	5 (50)
<i>Alter (Monate, CL-304: Tage)⁷⁾</i>						
MW (SD)	3,7 (1,6)	3,4 (2,1)	4,1 (1,3)	20,6 (7,9)	28,7 (11,7)	33,7 (7,7)
Median (min; max)	3,5 (0,5; 5,9)	3,1 (0,9; 7,9)	4,1 (1,8; 6,0)	21,0 (8; 34)	32,0 (9; 43)	31,9 (25,4; 46,3)
<i>Altersgruppe, n (%)</i>		k. A.				
0–27 Tage	1 (4,5)		0	11 (78,6)	6 (40,0)	-
28 Tage – 23 Monate	21 (95,5)		33 (100)	3 (21,4)	9 (60,0)	-
24–59 Monate	-		-	-	-	10 (100)
<i>Gestationsalter zu Geburt (Wochen)</i>		n = 10	n = 32			n = 8
MW (SD)	39,0 (1,0)	38,5 (1,4)	39,0 (1,4)	38,2 (1,42)	38,8 (1,5)	38,4 (1,51)
Median (min; max)	39,0 (37,0; 41,0)	39,0 (36,0; 40,0)	39,0 (35,0; 41,0)	38,0 (36; 41)	39,0 (35; 41)	39 (36; 40)
<i>Ethnie (genetisch), n (%)</i>			k. A.			
asiatisch	2 (9,1)	0		2 (14,3)	2 (13,3)	-
schwarz oder afroamerikanisch	3 (13,6)	0		2 (14,3)	0	-
kaukasisch/weiß	11 (50,0)	11 (91,7)		7 (50,0)	10 (66,7)	9 (90)
amerikanisch indianisch / Ureinwohner v. Alaska	0	0		0	1 (6,7)	-
andere	6 (27,3)	1 (8,3)		3 (21,4)	2 (13,3)	1 (10)
<i>Region, n (%)</i>	k. A.	k. A.	k. A.	k. A.	k. A.	k. A.
EMEA						
Nordamerika						
Lateinamerika						
Asien/Pazifik						
<i>Gewicht zu Baseline (kg)</i>						
MW (SD)	5,8 (1,1)	5,7 (1,3)	5,8 (1,0)	3,6 (0,4)	4,1 (0,5)	11,9 (1,34)
Median (min; max)	5,8 (3,9; 7,5)	5,5 (3,6; 8,4)	5,8 (4,2; 8,4)	3,7 (3; 4,3)	4,1 (3,1; 5,2)	12 (10; 15)

Studien Merkmal	CL-303 ¹⁾ N = 22	CL-101 ²⁾³⁾ N = 12	Studie CL-302 ⁴⁾ Kohorte 2 N = 33	Studie CL-304 ⁵⁾		LT-001 ⁶⁾ N = 10
				SMN2: 2 Kopien N = 14	SMN2: 3 Kopien N = 15	
BMI (kg/m ²) MW (SD) Median (min; max)	15,4 (1,7) 15,1 (13,0; 20,0)	k. A.	15,1 (2,0) 14,4 (12,1; 20,4)	k. A.	k. A.	14,8 (1,3) 14,7 (12,5; 16,7)
Anzahl SMN-Genkopien, n (%)		k. A. ⁸⁾				k. A.
1	0 (0,0)		0 (0,0)	-	-	
2	22 (100,0)		33 (100,0)	14 (100)	0	
3	-		-	0	15 (100)	
SMN1 biallelische Deletion, n (%)		k. A.		k. A.	k. A.	k. A.
präsent	22 (100,0)		33 (0,0)			
abwesend	0 (0,0)		0 (0,0)			
SMN1-Punktmutation, n (%)		k. A.		k. A.	k. A.	k. A.
präsent	0 (0,0)		0 (0,0)			
abwesend	22 (100,0)		33 (0,0)			
Alter bei Symptombeginn (Monate)			k. A.	k. A.	k. A.	k. A.
MW (SD)	1,9 (1,2)	1,4 (1,0)				
Median (min; max)	1,8 (0; 4)	1,0 (0; 3)				
Alter bei genetischer Diagnose (Tage)	k. A.	k. A.	k. A.	k. A.	k. A.	k. A.
MW (SD)						
Median (min; max)						
SMA-Symptome zu Baseline			k. A.	k. A.	k. A.	k. A.
Hypotonie	21 (95,5)	12 (100,0)				
Schwäche in den Gliedmaßen	21 (95,5)	11 (91,7)				
Pneumonie oder respiratorische Symptome	12 (54,5)	0 (0,0)				
Schlucken/Probleme bei Nahrungsaufnahme	4 (18,2)	5 (41,7)				
Zungenfaszikulation	17 (77,3)	0 (0,0)				
Verzögerte Entwicklung	14 (63,6)	-				
Konstipation	6 (27,3)	-				
Andere	14 (63,6)	1 (8,3)				

Studien Merkmal	CL-303 ¹⁾ N = 22	CL-101 ²⁾³⁾ N = 12	Studie CL-302 ⁴⁾ Kohorte 2 N = 33	Studie CL-304 ⁵⁾		LT-001 ⁶⁾ N = 10
				SMN2: 2 Kopien N = 14	SMN2: 3 Kopien N = 15	
<i>Schlucken dünnflüssiger Nahrung, n (%)</i>						
ja	22 (100,0)	4 (33,3)	31 (93,9)	k. A.	k. A.	k. A.
nein	0 (0,0)	8 (66,7)	2 (6,1)			
<i>Nicht-orale Nahrungsunterstützung, n (%)</i>						
ja	0 (0,0)	5 (41,7)	10 (30,3)	0	0	k. A.
nein	22 (100)	7 (58,3)	23 (69,7)	14 (100)	15 (100)	
<i>Unterstützung durch ein Beatmungsgerät, n (%)</i>						
ja	0 (0,0)	1 (8,3) ⁹⁾	9 (27,3)	0	0	k. A.
nein	22 (100)	11 (91,7)	24 (72,7)	14 (100)	15 (100)	
<i>Verabreichung von Prednisolon (Tage)</i>			k. A.	k. A.	k. A.	k. A.
MW (SD)	73,7 (39,5)	73,8 (33,04)				
Median (min; max)	62,5 (37; 203)	61,0 (44; 161)				
<i>CHOP INTEND</i>			32		k. A.	k. A.
MW (SD)	32 (9,7)	28,2 (12,3)	27,6 (8,2)	46,1 (8,8)		
Median (min; max)	33,5 (18; 52)	29,0 (12; 50)	27,5 (14; 55)	48,5 (28; 57)		
<i>Hospitalisierungen¹⁰⁾, n (%)</i>		k. A.				
ja	17 (77,3)		20 (60,6)	4 (28,6)	2 (13,3)	4 (40)
nein	5 (22,7)		13 (39,4)	10 (71,4)	13 (86,7)	6 (60)

¹⁾ CL-303: Finaler Datenschnitt zum 12.12.2019 (bis zu einem Alter von mindestens 18 Monaten).

²⁾ CL-101: Finaler Datenschnitt zum 14.12.2017 (24 Monate nach Genttransfer).

³⁾ CL-101: Ergebnisse dargestellt für Kohorte 2 mit der Dosierung 2,0 x 10¹⁴vg/kg.

⁴⁾ CL-302: Datenschnitt 31.12.2019; laufende Studie.

⁵⁾ CL-304: Datenschnitt 31.12.2019; laufende Studie.

⁶⁾ LT-001: Datenschnitt 31.12.2019; laufende Studie. Angaben zu den 10 Personen, die aus Studie CL101 in die Extensionsstudie LT-001 aufgenommen wurden.

⁷⁾ CL-304: Angaben in Tagen.

⁸⁾ In die Studie CL-101 wurden nur Personen mit 2 SMN2-Genkopien eingeschlossen.

⁹⁾ Tägliche Beatmungsunterstützung > 0 bis ≤ 12 Stunden/Tag.

¹⁰⁾ Hospitalisierung vor Studienbeginn.

Abkürzungen: CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; EMEA: Europa-Arabien-Afrika (Wirtschaftsraum); k. A.: keine Angabe; MW: Mittelwert; SD: Standardabweichung; SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron; vg: Vektorgenom/e.

Begleitmedikation

In Studie CL-303 nahmen alle 22 Patientinnen und Patienten mindestens eine Begleitmedikation ein. Die am häufigsten eingenommenen Begleitmedikamente nach ATC-Klassifikation waren Impfstoffe (n = 21; 95,5 %), Analgetika (n = 16; 72,7 %) und Medikamente gegen säureassoziierte Störungen (n = 15; 68,2 %).

In Studie CL-101 nahmen alle 12 Patientinnen und Patienten mindestens eine Begleitmedikation ein. Am häufigsten wurden folgende Medikamente verabreicht: Paracetamol (n = 11; 91,7 %), Grippeimpfstoffe (n = 9; 75,0 %), Ranitidinhydrochlorid (n = 8; 66,7 %), Natriumchlorid (n = 9; 75,0 %), Ibuprofen (n = 8; 66,7 %), Pneumokokken-Impfstoff (n = 7, 58,3 %), Amoxicillin (n = 6; 50,0 %), Impfstoff gegen Diphtherie/Keuchhusten/Tetanus (n = 6; 50,0 %), Impfstoff gegen Haemophilus-Influenza Typ B (Hib) (n = 7; 58,3 %).

3.2 Mortalität

Mortalität (Todesfälle)

Studie CL-303 (finaler Datenschnitt: 12. Dezember 2019)

Von den 22 eingeschlossenen Patientinnen und Patienten in der Studie CL-303 ist eine Person im Alter von 7,8 Monaten verstorben (ohne Zusammenhang mit der Studienmedikation) (siehe Tabelle 21). Angaben zur medianen Beobachtungsdauer liegen nicht vor.

Studie CL-101 (finaler Datenschnitt: 14. Dezember 2017)

Alle 12 Patientinnen und Patienten in der Studie CL-101 (Kohorte 2) überlebten bis zu einem Alter von $\geq 13,6$ Monaten (Zeitpunkt der primären Wirksamkeitsanalyse; Datenschnitt: 20. Januar 2017) und 24 Monate nach Gentransfer (Datenschnitt: 14. Dezember 2017). Angaben zur medianen Beobachtungsdauer liegen nicht vor (siehe Tabelle 21).

Studien PNCR, NeuroNEXT und ENDEAR

Angaben zur Mortalität in den externen Kontrollen im Alter 13–14 Monate sind Tabelle 17 zu entnehmen. Die mediane Beobachtungsdauer (min; max) ist ausschließlich für die Studie ENDEAR bekannt (187 Tage (20; 423)). Für die Studien PNCR und NeuroNEXT finden sich nur Angaben über die geplante Beobachtungsdauer (PNCR: bis zu 36 Monate, NeuroNEXT: bis zu 24 Monate).

Tabelle 17: Ergebnisse zum Endpunkt „Mortalität“ in den Studien CL-303 und CL-101 und den externen Kontrollstudien (PNCR, NeuroNEXT, ENDEAR) (Alter: 13–14 Monate)

Studien Endpunkt „Mortalität“	CL-303 ¹⁾		CL-101 ²⁾		PNCR ³⁾		NeuroNEXT ⁴⁾		ENDEAR (BSC) ⁵⁾	
	N ⁶⁾	Personen mit Ereignis n (%)	N ⁶⁾	Personen mit Ereignis n (%)	N ⁶⁾	Personen mit Ereignis n (%)	N ⁶⁾	Personen mit Ereignis n (%)	N ⁶⁾	Personen mit Ereignis n (%)
Mortalität (Alter: 13–14 Monate)	22	1 (2,9)	12	0	23	7 (30,4)	16	7 (43,8)	41	16 (39,0)

¹⁾ ITT-Population. Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).

²⁾ ITT-Population. Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung 2,0 x 10¹⁴vg/kg.

³⁾ Geplante Beobachtungsdauer bis zu 36 Monate.

⁴⁾ Geplante Beobachtungsdauer bis zu 24 Monate.

⁵⁾ ITT-Population. Für die Studie ENDEAR war eine Beobachtungszeit von 13 Monaten geplant. Die Studie wurde wegen der guten Wirksamkeit jedoch frühzeitig beendet. Die mediane Beobachtungsdauer (min; max) betrug 187 Tage (20; 423), das entspricht in etwa 6,2 Monaten, und ist damit kürzer als in den Zulassungsstudien.

⁶⁾ Anzahl Patientinnen und Patienten in der Auswertung.

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; vg: Vektorgenom/e.

Vergleich Studien CL-303 und CL-101 vs. Studien PNCR, NeuroNEXT und ENDEAR

Angaben zur Mortalität der Zulassungsstudien im Vergleich zu den externen Kontrollen im Alter 13–14 Monate sind in den nachfolgenden zwei Tabellen adressiert. Für die gepoolten Studiendaten der Studien CL-101+CL-303 im Vergleich mit den Studien PNCR+NeuroNEXT zeigte sich im Alter 13–14 Monate ein RR von 12,21 (95%-KI: [1,692; 88,04], p < 0,001; siehe Tabelle 18). Für die gepoolten Studiendaten der Studien CL-101+CL-303 im Vergleich mit der Studie ENDEAR zeigte sich im Alter 13–14 Monate ein RR von 13,27 (95%-KI: [1,853; 95], p < 0,001; siehe Tabelle 19). In der ENDEAR-Studie betrug die mediane Beobachtungszeit 6,2 Monate und ist damit kürzer als in den Zulassungsstudien.

Tabelle 18: Ergebnisse zum Endpunkt „Mortalität“ in den Studien CL-101+CL-303 im Vergleich zu den externen Kontrollstudien PNCR+NeuroNEXT (Alter: 13–14 Monate)

Studien Endpunkt „Mortalität“	CL-303+CL-101 ¹⁾ N = 34	PNCR+NeuroNEXT N = 39
Personen mit Ereignis, n (%)	1 (2,9)	14 (35,9)
Zensierungen ²⁾ , n (%)	33 (97,1)	25 (64,1)
RR [95%-KI]	12,21 [1,692; 88,04]	
p-Wert ³⁾	< 0,001	

¹⁾ Angaben beziehen sich auf die ITT-Population. Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung 2,0 x 10¹⁴vg/kg.

²⁾ Personen ohne Ereignis.

³⁾ Exakter Test nach Fisher (Fisher's exakter Test).

Abkürzungen: ITT: Intention-to-Treat; KI: Konfidenzintervall; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; RR: Relatives Risiko; vg: Vektorgenom/e.

Tabelle 19: Ergebnisse zum Endpunkt „Mortalität“ in den Studien CL-101+CL-303 im Vergleich zur externen Kontrollstudie ENDEAR (Alter: 13–14 Monate)

Studien Endpunkt „Mortalität“	CL-101+CL-303 ¹⁾ N = 34	ENDEAR ²⁾ N = 41
Personen mit Ereignis, n (%)	1 (2,9)	16 (39,0)
Zensurierungen ³⁾ , n (%)	33 (97,1)	25 (61,0)
RR [95%-KI]	13,27 [1,853; 95]	
p-Wert ⁴⁾	< 0,001	

¹⁾ Angaben beziehen sich auf die ITT-Population. Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

²⁾ Angaben aus Modul 4 des Nutzendossiers von Nusinersen. In der Studie ENDEAR betrug die geplante Beobachtungsdauer 13 Monate. Aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht in etwa 6,2 Monaten, und ist damit kürzer als in den Zulassungsstudien.

³⁾ Personen ohne Ereignis.

⁴⁾ Exakter Test nach Fisher (Fisher's exakter Test).

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; KI: Konfidenzintervall; RR: Relatives Risiko; vg: Vektor-genom/e.

Studien LT-001, CL-302 und CL-304 (Datenschnitt: 31. Dezember 2019)

Zum Datenschnitt 31. Dezember 2019 beträgt die mediane Zeit (min; max) in der Studie CL-302 11,9 Monate (1,8; 15,4) und in der Studie CL-304 in der Kohorte mit 2 SMN2-Genkopien 9,9 Monate (5,1; 18) sowie in der Kohorte mit 3 SMN2-Genkopien 9,0 Monate (2; 13,9). Für Studie LT-001 liegen keine Angaben vor.

In der Extensionsstudie (LT-001) haben alle 10 Patientinnen und Patienten aus der Studie CL-101 (Kohorte 2) zum Datenschnitt 31. Dezember 2019 überlebt. In der Studie CL-302 haben 32 (97 %) von 33 eingeschlossenen Patientinnen und Patienten überlebt. Eine Person der Studienpopulation ist im Alter von 6,9 Monaten verstorben. In der Studie CL-304 haben alle 14 Personen der Kohorte 1 (2 SMN2-Genkopien) und alle 15 Personen in Kohorte 2 (3 SMN2-Genkopien) überlebt.

Tod oder dauerhafte Beatmung (Überleben)

Studie CL-303 (finaler Datenschnitt: 12. Dezember 2019)

Das Überleben ist definiert als das Vermeiden von entweder a) Tod oder b) dauerhafter Beatmung. Dauerhafte Beatmung ist definiert als die Notwendigkeit entweder einer Tracheostomie (invasiv) oder ventilatorischer Unterstützung von ≥ 16 Stunden pro Tag (nicht-invasiv) über einen Zeitraum von ≥ 14 aufeinanderfolgenden Tagen in Abwesenheit einer akuten reversiblen Erkrankung, ausgenommen perioperative Beatmung.

Zum Alter von $\geq 13,6$ Monaten und bis zu einem Alter von mindestens 18 Monaten (geplantes Studienende) überlebten 20 von 22 Patientinnen und Patienten ereignisfrei. Eine Person verstarb im Alter von 7,8 Monaten aufgrund von Atemstillstand, ohne Zusammenhang mit der Studienmedikation. Für die andere Person wurde eine nicht-invasive Beatmung (BiPAP) für ≥ 16 Stunden über einen Zeitraum von über 14 aufeinanderfolgenden Tagen berichtet; aufgrund fehlender BiPAP-Daten konnte dies nicht abschließend bestätigt werden, weshalb diese Person mit Ereignis „Dauerhafte Beatmung“ gezählt wurde. Diese Person zog ihre Einverständniserklärung im Alter von 11,9 Monaten zurück. Unter der Annahme, dass 25 % der Patientinnen und Patienten in der PNCR-Kohorte ein Alter von 13,6 Monaten ereignisfrei überleben, zeigte sich ein statistisch signifikanter Unterschied zugunsten von Onasemnogen-Abeparvovec (100 %, 97,5%-KI: [67,86; 99,23],

$p < 0,0001$ (einseitiger exakter Binomialtest)). Angaben für den finalen Datenschnitt (bis zu einem Alter von mindestens 18 Monaten) liegen nicht vor (siehe Tabelle 20).

Studie CL-101 (finaler Datenschnitt: 14. Dezember 2017)

Das Überleben in der Studie CL-101 war definiert als Zeit von der Geburt bis a) Tod oder b) dauerhafte Beatmung, definiert als Notwendigkeit einer invasiven Beatmungsunterstützung ≥ 16 Stunden pro Tag kontinuierlich für ≥ 14 Tage, bei Abwesenheit einer akuten reversiblen Erkrankung, ausgenommen perioperative Beatmung.

Alle 12 Patientinnen und Patienten (Kohorte 2) überlebten ereignisfrei bis zu einem Alter von $\geq 13,6$ Monaten (Zeitpunkt der primären Wirksamkeitsanalyse, Datenschnitt: 20. Januar 2017) und 24 Monate nach Gentransfer (Datenschnitt: 14. Dezember 2017). Unter der Annahme, dass 25 % der Patientinnen und Patienten in der PNCR-Kohorte ein Alter von 13,6 Monaten und 8 % ein Alter von 24 Monaten ereignisfrei überleben, zeigte sich ein statistisch signifikanter Unterschied zugunsten von Onasemnogen-Abepravovec zu 13,6 und 24 Monaten (jeweils 90,9 %, 95%-KI: [73,54; 100,00], $p < 0,001$ (einseitiger exakter Binomialtest)) (siehe Tabelle 20).

Tabelle 20: Ergebnisse zum Endpunkt „Überleben“ (Tod oder dauerhafte Beatmung) in den Studien CL-101 und CL-303

Studie Endpunkt „Überleben“	CL-303 ¹⁾ N = 22		CL-101 ²⁾ N = 12	
	Alter $\geq 13,6$ Monate	Bis zu einem Alter von mindestens 18 Monaten	Alter $\geq 13,6$ Monate	24 Monate nach Gentransfer
Personen mit Ereignis, n (%)	2 (9,1)	2 (9,1)	0	0
Zensurierungen ³⁾ , n (%)	20 (90,9)	20 (90,9)	12 (100)	12 (100)
[KI] p-Wert	[67,86; 99,23] ⁴⁾ < 0,0001	k. A. k. A.	[73,54; 100,00] ⁵⁾ < 0,001	

¹⁾ ITT-Population; berechnet mit einem einseitigen exakten Binomialtest auf Überlegenheit unter Zugrundelegung einer 25%igen Überlebenswahrscheinlichkeit bis zum Alter von 14 Monaten.

²⁾ ITT-Population; Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg. Berechnet mit einem einseitigen exakten Binomialtest auf Überlegenheit unter Zugrundelegung einer 25%igen Überlebenswahrscheinlichkeit bis zum Alter von 13,6 Monaten und 8 % bis zum Alter von 20 / 24 Monaten Überlebenswahrscheinlichkeit, die in der PNCR-Studie beobachtet wurde.

³⁾ Personen ohne Ereignis.

⁴⁾ Zweiseitiges 97,5%-KI basierend auf der exakten Binomialverteilung.

⁵⁾ Zweiseitiges 95%-KI basierend auf der exakten Binomialverteilung.

Abkürzungen: ITT: Intention-to-Treat; k. A.: keine Angabe; KI: Konfidenzintervall; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; vg: Vektorgenom/e.

PNCR-, NeuroNEXT- und ENDEAR-Kohorten

Angaben zum Überleben (Tod oder dauerhafte Beatmung) in den Zulassungsstudien und in den externen Kontrollen zum Alter 13–14 Monate sind Tabelle 21 zu entnehmen.

Tabelle 21: Ergebnisse zum Endpunkt „Überleben“ (Tod oder dauerhafte Beatmung) in den Studien CL-303 und CL-101 und den externen Kontrollen (PNCR, NeuroNEXT, ENDEAR) (Alter: 13–14 Monate)

Studie Endpunkt „Überleben“	CL-303 ¹⁾		CL-101 ²⁾		PNCR ³⁾⁴⁾		NeuroNEXT ⁵⁾⁶⁾		ENDEAR (BSC) ⁷⁾⁸⁾	
	N ⁹⁾	Personen mit Ereignis n (%)	N ⁹⁾	Personen mit Ereignis n (%)	N ⁹⁾	Personen mit Ereignis n (%)	N ⁹⁾	Personen mit Ereignis n (%)	N ⁹⁾	Personen mit Ereignis n (%)
Tod oder dauerhafte Beatmung (Alter: 13–14 Monate)	22	2 (9,1)	12	0	23	16 (69,6)	16	8 (50,0)	41	28 (68,3)

- ¹⁾ ITT-Population. Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).
²⁾ ITT-Population. Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung 2,0 x 1014 vg/kg.
³⁾ Geplante Beobachtungsdauer bis zu 36 Monate.
⁴⁾ Operationalisierung des Endpunkts entspricht denen in den Studien CL-101 und CL-303. Operationalisierung in den Studien CL-101, CL-303 und PNCR: Tracheostomie oder dauerhafte nicht-invasive Beatmung (≥ 16 Stunden/Tag und ≥ 14 aufeinanderfolgende Tage).
⁵⁾ Geplante Beobachtungsdauer bis zu 24 Monate.
⁶⁾ Abweichende Operationalisierung in der Studie NeuroNEXT: Tod oder invasive Beatmung.
⁷⁾ ITT-Population. Für die Studie ENDEAR war eine Beobachtungszeit von 13 Monaten geplant. Die Studie wurde wegen der guten Wirksamkeit jedoch frühzeitig beendet. Die mediane Beobachtungsdauer (min; max) betrug 187 Tage (20; 423), das entspricht in etwa 6,2 Monaten, und ist damit kürzer als in den Zulassungsstudien.
⁸⁾ Abweichende Operationalisierung in der Studie ENDEAR: Überleben definiert als Tod oder dauerhafte Beatmung (≥ 16 Stunden/Tag durchgehend für ≥ 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses oder Notwendigkeit einer Tracheostomie).
⁹⁾ Anzahl Personen in der Auswertung.

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; vg: Vektorgenom/e.

Vergleich Studien CL-303 und CL-101 vs. Studien PNCR, NeuroNEXT und ENDEAR

Angaben zum Überleben (Tod oder dauerhafte Beatmung) der Zulassungsstudien im Vergleich zu den externen Kontrollen im Alter 13–14 Monate sind in den nachfolgenden Tabellen adressiert. Für die gepoolten Studiendaten der Studien CL-101+CL-303 im Vergleich zu den gepoolten Studiendaten der Studien PNCR+NeuroNEXT zeigte sich im Alter 13–14 Monate ein RR von 10,46 (95%-KI: [2,666; 41,06], p < 0,001; siehe Tabelle 22). Die gepoolten Studiendaten der Studien CL-101+CL-303 im Vergleich zu Studie ENDEAR zeigten im Alter 13–14 Monate ein RR von 11,61 (95%-KI: [2,978; 45,26], p < 0,001; siehe Tabelle 23).

Tabelle 22: Ergebnisse zum Endpunkt „Überleben“ (Tod oder dauerhafte Beatmung) in den Studien CL-101+CL-303 im Vergleich zur externen Kontrolle PNCR+NeuroNEXT (Alter: 13–14 Monate)

Studien Endpunkt „Überleben“	CL-303+CL-101 ¹⁾ N = 34	PNCR+NeuroNEXT ²⁾³⁾ N = 39
Personen mit Ereignis, n (%)	2 (5,9)	24 (61,5)
Zensierungen ⁴⁾ , n (%)	32 (94,1)	15 (38,5)
RR [95%-KI]	10,46 [2,666; 41,06]	
p-Wert ⁵⁾	< 0,001	

¹⁾ Angaben beziehen sich auf die ITT-Population. Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

²⁾ Operationalisierung des Endpunktes in der Studie PNCR entspricht denen in den Studien CL-101 und CL-303.

³⁾ In der Studie NeuroNEXT wird das Überleben definiert als Zeit bis a) Tod oder b) dauerhafte invasive Beatmung (Intubation). Tracheostomie und Intubation werden jedoch als hinreichend vergleichbar im Sinne einer invasiven Beatmungsunterstützung angesehen. Im Vergleich zu den Studien CL-101, CL-303 und PNCR ist die nicht-invasive Beatmung in der Studie NeuroNEXT nicht mit umfasst. Zudem fehlen Angaben zur Dauer der invasiven Beatmung, wie sie für die Studien CL-101, CL-303 und PNCR vorliegen (siehe Abschnitt 2.3.5).

⁴⁾ Personen ohne Ereignis.

⁵⁾ Exakter Test nach Fisher (Fisher's exakter Test).

Abkürzungen: ITT: Intention-to-Treat; KI: Konfidenzintervall; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; RR: Relatives Risiko; vg: Vektorgenom/e.

Tabelle 23: Ergebnisse zum Endpunkt „Überleben“ (Tod oder dauerhafte Beatmung) in den Studien CL-101+CL-303 im Vergleich zur externen Kontrolle ENDEAR (Alter: 13–14 Monate)

Studien Endpunkt „Überleben“	CL-303+CL-101 ¹⁾ N = 34	ENDEAR ²⁾³⁾ N = 41
Personen mit Ereignis, n (%)	2 (5,9)	28 (68,3)
Zensierungen ⁴⁾ , n (%)	32 (94,1)	13 (31,7)
RR [95%-KI]	11,61 [2,978; 45,26]	
p-Wert ⁵⁾	< 0,001	

¹⁾ Angaben beziehen sich auf die ITT-Population. Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

²⁾ Abweichende Operationalisierung in der Studie ENDEAR im Vergleich zu den Studien CL-101 und CL-303: Überleben war definiert als Tod oder dauerhafte Beatmung (≥ 16 Stunden/Tag durchgehend für > 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses oder Notwendigkeit einer Tracheostomie) (siehe Abschnitt 2.3.5).

³⁾ In der Studie ENDEAR betrug die geplante Beobachtungsdauer 13 Monate. Aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht in etwa 6,2 Monaten, und ist damit kürzer als in den Zulassungsstudien.

⁴⁾ Personen ohne Ereignis.

⁵⁾ Exakter Test nach Fisher (Fisher's exakter Test).

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; KI: Konfidenzintervall; RR: Relatives Risiko; vg: Vektorgenom/e.

Studien LT-001, CL-302 und CL-304 (Datenschnitt: 31. Dezember 2019)

In der Extensionsstudie (LT-001) sind alle 10 Personen zum Datenschnitt am Leben und frei von dauerhafter Beatmung. In der Studie CL-302 haben 32 (97 %) von 33 eingeschlossenen Personen bis zum Datenschnitt am 31. Dezember 2019 ereignisfrei überlebt (31/32 Personen der ITT-Population). In der Studie CL-304 haben ebenfalls alle 14 Personen der Kohorte 1 (2 SMN2-Genkopien) und 15 Personen der Kohorte 2 (3 SMN2-Genkopien) zum Datenschnitt ereignisfrei überlebt.

3.3 Morbidität

Dauerhafte Beatmung

Studie CL-303 (Datenschnitt: 12. Dezember 2019)

Bis zum finalen Datenschnitt (bis zu einem Alter von mindestens 18 Monaten) benötigte eine Person (4,5 %) eine dauerhafte Beatmung. Für diese Person wurde eine nicht-invasive Beatmung (BiPAP) für ≥ 16 Stunden über einen Zeitraum von über 14 aufeinanderfolgenden Tagen berichtet.

Studie CL-101 (finaler Datenschnitt: 14. Dezember 2017)

Bis zum finalen Datenschnitt (24 Monate nach Gentransfer) benötigte keine Person in Kohorte 2 (N = 12) eine dauerhafte Beatmung.

Vergleich der Studien CL-101 und CL-303 vs. Studien PNCR, NeuroNEXT und ENDEAR

In der Studie PNCR entspricht die Operationalisierung des Endpunktes der Definition in den Studien CL-101 und CL-303, wodurch die Daten direkt vergleichbar sind. In einem Alter von 14 Monaten benötigten 13 Personen (56,5 %) in der PNCR-Studie eine dauerhafte Beatmung, in der Studie CL-101 war dies zu einem vergleichbaren Beobachtungszeitpunkt bei keiner Person (0 %) notwendig. In der Studie CL-303 wurde eine Person (4,5 %) dauerhaft beatmet (≥ 16 Stunden über mindestens 14 aufeinanderfolgende Tage) (siehe Tabelle 24).

Tabelle 24: Ergebnisse zum Endpunkt „Dauerhafte Beatmung“ in den Studien CL-303 und CL-101 im Vergleich zur externen Kontrolle PNCR (Alter: 13–14 Monate)

Studien Endpunkt „Dauerhafte Beatmung“	CL-303 ¹⁾ N = 22	PNCR ²⁾ N = 23	CL-101 ¹⁾ N = 12	PNCR ²⁾ N = 23	CL-101+ CL-303 ¹⁾ N = 34	PNCR ²⁾ N = 23
Personen mit Ereignis, n (%)	1 (4,5)	13 (56,5)	0	13 (56,5)	1 (2,9)	13 (56,5)
Zensierungen ³⁾ , n (%)	21 (95,5)	10 (43,5)	12 (100)	10 (43,5)	33 (97,1)	10 (43,5)
RR [95%-KI]	12,43 [1,772; 87,24]		14,625 [0,9; 226,7]		19,22 [2,696; 137]	
p-Wert ⁴⁾	< 0,001		< 0,001		< 0,001	

¹⁾ Angaben beziehen sich auf die ITT-Population. Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

²⁾ Der Endpunkt „Dauerhafte Beatmung“ ist in der Studie PNCR gleich operationalisiert (siehe Abschnitt 2.3.5.).

³⁾ Personen ohne Ereignis.

⁴⁾ Exakter Test nach Fisher (Fisher's exakter Test).

Abkürzungen: ITT: Intention-to-Treat; KI: Konfidenzintervall; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; RR: Relatives Risiko; vg: Vektorgenom/e.

In der Studie NeuroNEXT ist „Dauerhafte Beatmung“ auf Personen mit langfristiger invasiver Beatmung (Intubation) beschränkt, weshalb für den Vergleich mit den Studien CL-101 und CL-303 nur Personen berücksichtigt wurden, für die ebenfalls eine Intubation bis zum Alter 13–14 Monate berichtet wurden. Das traf auf keine Person in den Zulassungsstudien zu (siehe Tabelle 25).

Tabelle 25: Ergebnisse zum Endpunkt „Intubation“ in den Studien CL-101 und CL-303 im Vergleich zur externen Kontrolle NeuroNEXT (Alter: 13–14 Monate)

Studien Endpunkt „Intubation“	CL-303 ¹⁾ N = 22	NeuroNEXT ³⁾ N = 16	CL-101 ¹⁾²⁾ N = 12	NeuroNEXT ³⁾ N = 16	CL-101+ CL-303 ¹⁾ N = 34	NeuroNEXT ³⁾ N = 16
Personen mit Ereignis, n (%)	0	2 (12,5)	0	2 (12,5)	0	2 (12,5)
Zensurierungen ⁴⁾ , n (%)	22 (100)	14 (87,5)	12 (100)	14 (87,5)	34 (100)	14 (87,5)
RR [95%-KI]	6,765 [0,3467; 132]		3,824 [0,2; 73,0]		10,29 [0,5226; 202,8]	
p-Wert ⁵⁾	0,1369		0,492		0,0541	

¹⁾ Angaben beziehen sich auf die ITT-Population. Für die Vergleichbarkeit mit der Studie NeuroNEXT wurden nur Personen gezählt, die eine invasive Beatmung mittels Intubation aufweisen.

²⁾ Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

³⁾ In der Studie NeuroNEXT wurden nur Personen erfasst, für die eine Intubation (invasive Beatmung) berichtet wurden (siehe Abschnitt 2.3.5).

⁴⁾ Personen ohne Ereignis.

⁵⁾ Exakter Test nach Fisher (Fisher's exakter Test).

Abkürzungen: ITT: Intention-to-Treat; KI: Konfidenzintervall; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; RR: Relatives Risiko; vg: Vektorgenom/e.

In der Studie ENDEAR wurde sowohl invasive als auch nicht-invasive Beatmung für mindestens 16 Stunden am Tag als „dauerhaft“ gewertet. Die Operationalisierungen unterscheiden sich jedoch darin, ab wie vielen aufeinanderfolgenden Tagen dieses Kriterium erfüllt sein muss: In den Studien CL-101 und PNCR ist der Zeitraum mit 14 Tagen kürzer als in der Studie ENDEAR mit 21 Tagen. In der Studie ENDEAR betrug die mediane Beobachtungszeit 6,2 Monate und ist damit kürzer als in den Zulassungsstudien (siehe Tabelle 26).

Tabelle 26: Ergebnisse zum Endpunkt „Dauerhafte Beatmung“ in den Studien CL-303 und CL-101 im Vergleich zur externen Kontrolle ENDEAR (Alter: 13–14 Monate)

Studien Endpunkt „Dauerhafte Beatmung“	CL-303 ¹⁾ N = 22	ENDEAR ²⁾ N = 41	CL-101 ¹⁾ N = 12	ENDEAR ²⁾ N = 41	CL-101+ CL-303 ¹⁾ N = 34	ENDEAR ²⁾ N = 41
Personen mit Ereignis, n (%)	1 (4,5)	13 (31,7)	0	13 (31,7)	1 (2,9)	13 (31,7)
Zensurierungen ³⁾ , n (%)	21 (95,5)	28 (68,3)	12 (100)	28 (68,3)	33 (97,1)	28 (68,3)
RR [95%-KI]	6,976 [0,9758; 49,86]		8,4 [0,5; 131,2]		10,78 [1,485; 78,28]	
p-Wert ⁴⁾	0,0232		0,026		0,001	

¹⁾ Auswertung bezieht sich auf die ITT-Population. Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

²⁾ In der Studie ENDEAR war „Dauerhafte Beatmung“ definiert als ≥ 16 Stunden/Tag durchgehende Unterstützung (Tracheostomie oder nicht-invasive Beatmung) für > 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses oder bei Notwendigkeit einer Tracheostomie).

³⁾ Personen ohne Ereignis.

⁴⁾ Exakter Test nach Fisher (Fisher's exakter Test).

Abkürzungen: ITT: Intention-to-Treat; KI: Konfidenzintervall; RR: Relatives Risiko; vg: Vektorgenom/e.

Studien LT-001, CL-302 und CL-304 (Datenschnitt: 31. Dezember 2019)

Gemäß Nutzendossier (Modul 4) benötigte in der Extensionsstudie LT-001 und in der Studie CL-304 keine Person zum Datenschnitt 31. Dezember 2019 eine dauerhafte Beatmung. Für Studie CL-302 liegen keine Informationen vor.

Invasive Beatmung und ventilatorische Unterstützung (Beatmungsfreiheit)

Studie CL-303 (Datenschnitt: 12. Dezember 2019)

In den Studienunterlagen finden sich keine Angaben über den Anteil an Personen, die eine invasive Beatmungsunterstützung benötigten. Die Person in der Studie, die eine dauerhafte Beatmung benötigte, erhielt eine nicht-invasive Beatmungsunterstützung.

Der Anteil Personen, die von der Beatmungsunterstützung unabhängig waren, definiert als Personen, die im Alter von 18 Monaten keine tägliche Beatmungsunterstützung benötigten, ausgenommen akute reversible Erkrankungen und perioperative Beatmung, definiert durch die Bewertung der tatsächlichen Nutzungsdaten, die vom BiPAP-Gerät (nicht-invasiv) erfasst wurden, umfasste 18 von 22 Personen (81,8 %).¹⁰

Der Anteil der Personen, die zu jedem Zeitpunkt während der Studie keine nicht-invasive Beatmungsunterstützung (BiPAP) benötigten, umfasste 15 von 22 Personen (68,2 %).

Studie CL-101 (finaler Datenschnitt: 14. Dezember 2017)

Bis zum finalen Datenschnitt (24 Monate nach Gentransfer) benötigten 5 von 12 Personen (41,7 %) eine einmalige invasive Beatmung mittels Endotrachealtubus über Mund/Nase. Die Dauer der invasiven Beatmung lag zwischen 1 bis 9 Tagen. Somit benötigten 7 Personen (58,3 %) keine invasive Beatmung während der Studie.

Insgesamt benötigten zum Zeitpunkt des finalen Datenschnitts (24 Monate nach Gentransfer) 7 Personen (58,3 %) keine tägliche Beatmungsunterstützung. Für eine dieser Personen wurde eine tägliche Atmungsunterstützung angegeben, jedoch wurde zur letzten Studienvisite berichtet, dass kein täglicher Gebrauch des Beatmungsgerätes stattfand.

Erreichen motorischer Meilensteine

Studie CL-303 (Datenschnitt: 12. Dezember 2019)

In Tabelle 27 sind die erreichten Meilensteine zum finalen Datenschnitt, Alter ≥ 18 Monate dargelegt.

Zusätzlich wurde für die motorischen Meilensteine das mediane Alter beim Erreichen angegeben. Dieses wird nachfolgend nur für jene Meilensteine berichtet, die bei mehr als einer Person auftraten: So lag das mediane Alter (min; max) beim Erreichen der motorischen Meilensteine für „Freies Sitzen für ≥ 30 Sekunden“ (#26, Bayley-Skala) bei 12,6 Monaten (9,2; 18,6), für „Freies Sitzen für ≥ 10 Sekunden“ (WHO MGRS) bei 13,9 Monaten (10,2; 18,6), für „Rollen“ (#20, Bayley-Skala) bei 11,5 Monaten (2,9; 18,3) und für „Kopfkontrolle für mindestens 3 Sekunden ohne Unterstützung“ (#4, Bayley-Skala) bei 6,75 Monaten (2,0; 16,4).

¹⁰ Diskrepanz innerhalb des Studienberichtes zu CL-301; die geplante Auswertung für den Endpunkt erfolgt für die Anzahl der Personen, die zu jedem Zeitpunkt während der Studie keine ventilatorische Beatmungsunterstützung benötigten. Die geplante Auswertung mittels einseitigem Binomialtest liegt nicht vor.

Studie CL-101 (Datenschnitt: 14. Dezember 2017)

In Tabelle 27 sind die erreichten Meilensteine gemäß zentralem Review 24 Monate nach Gentransfer dargelegt. Angaben zum medianen Alter beim Erreichen der motorischen Meilensteine liegen für die Studie nicht vor.

Tabelle 27: Ergebnisse zum Endpunkt „Erreichen motorischer Meilensteine“ in den Studien CL-303 und CL-101 (gemäß zentralem Review)

Studien Endpunkt „Erreichen motorischer Meilensteine“	CL-303¹⁾ N = 22 <i>n (%)</i>	CL-101²⁾ N = 12 <i>n (%)</i>
Kopfkontrolle ohne Unterstützung (#4, Bayley-Skala)	17 (85,0) ³⁾⁴⁾	11 (91,7)
Rollen (#20, Bayley-Skala)	13 (59,1)	9 (75)
Sitzen mit Unterstützung (#19, Bayley-Skala)	k. A.	11 (91,7)
Freies Sitzen für ≥ 30 Sekunden (#26, Bayley-Skala)	13 (59,1) ⁵⁾	9 (75,0)
Freies Sitzen für ≥ 10 Sekunden (WHO MGRS)	14 (63,6)	10 (83,3)
Freies Sitzen für ≥ 15 Sekunden ⁶⁾	n. e.	9 (75,0)
Freies Sitzen für ≥ 5 Sekunden (#22, Bayley-Skala)	k. A.	11 (91,7) ⁷⁾
Gewicht für mindestens 2 Sekunden halten	1 (4,5) ⁸⁾	n. e.
Krabbeln	1 (4,5)	n. e.
Hochziehen in eine stehende Position (#35, Bayley-Skala)	1 (4,5)	n. e.
Stehen mit Unterstützung (#33, Bayley-Skala)	1 (4,5)	2 (16,7) ⁸⁾
Gehen mit Unterstützung (#37, Bayley-Skala)	1 (4,5)	2 (16,7) ⁸⁾
Stehen ohne Unterstützung (#40, Bayley-Skala)	1 (4,5)	2 (16,7) ⁸⁾
Gehen ohne Unterstützung (#43, Bayley-Skala)	1 (4,5)	2 (16,7) ⁸⁾

¹⁾ ITT-Population. Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).

²⁾ FAS-Population. Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14. Dezember 2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

³⁾ Zwei Personen hatten den Meilenstein bereits zur Screening-Visite, d. h. vor Behandlung mit Onasemnogen-Abepravovec erreicht; daher gingen nur 20 der 22 Personen in die Auswertung ein.

⁴⁾ Hier liegen inkonsistente Angaben in den Studienunterlagen vor: 18/20 vs. 17/20 Personen, die den Kopf für mindestens 3 Sekunden ohne Unterstützung halten können.

⁵⁾ Eine Person hat den Meilenstein „Freies Sitzen für 30 Sekunden“ in einem Alter von 16 Monaten erreicht, der Meilenstein wurde in der Studienvsiste in einem Alter von 18 Monaten nicht bestätigt.

⁶⁾ Definiert nach Nationwide Children's Hospital.

⁷⁾ Wurde im Studienbericht als „Freies Sitzen < 10 Sekunden“ bezeichnet. Die externe Begutachtung bestätigte, dass dieser Meilenstein als "Freies Sitzen ≥ 5 Sekunden" definiert und erfasst wurde.

⁸⁾ Nicht präspezifiziert.

Abkürzungen: n. e.: nicht ermittelt; FAS: Full Analysis Set; ITT: Intention-to-Treat; k. A.: keine Angabe; vg: Vektorgenom/e; WHO MGRS: World Health Organization Multicenter Growth Reference Study.

Tabelle 28: Ergebnisse zum Endpunkt „Unabhängiges Sitzen (≥ 30 Sekunden)“ in den Studien CL-101 und CL-303

Studien Endpunkt „Unabhängiges Sitzen (≥ 30 Sekunden)“	CL-303 ¹⁾ N = 22	CL-101 ²⁾ N = 12	
	Bis zu einem Alter von mindestens 18 Monaten	Alter $\geq 13,6$ Monate	24 Monate nach Gentransfer
Personen mit Ereignis, n (%)	13 (59,1)	5 (41,7)	9 (75)
Zensierungen ³⁾ , n (%)	9 (40,9)	7 (58,3)	3 (25)
[KI] p-Wert	[33,6; 81,4] ⁴⁾ < 0,0001	[15,17; 72,33] ⁵⁾ < 0,001	[42,81; 94,51] ⁵⁾ < 0,001

¹⁾ ITT-Population. Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).

²⁾ FAS-Population. Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14. Dezember 2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

³⁾ Personen ohne Ereignis.

⁴⁾ Zweiseitiges 97,5%-KI basierend auf der exakten Binomialverteilung.

⁵⁾ Zweiseitiges 95%-KI basierend auf der exakten Binomialverteilung.

Abkürzungen: FAS: Full Analysis Set; ITT: Intention-to-Treat; KI: Konfidenzintervall; vg: Vektorgenom/e.

Bayley-Skala (Grob- und Feinmotorik)

Studie CL-303 (Datenschnitt: 12. Dezember 2019)

In den Tabelle 29 und Tabelle 30 finden sich die Ergebnisse zu den Subskalen „Grobmotorik“ und „Feinmotorik“ der Bayley-Skala der Studie CL-303 über den Studienverlauf.

Tabelle 29: Ergebnisse zum Endpunkt „Bayley-Skala – Grobmotorik“ in der Studie CL-303

Studie Endpunkt „Bayley-Skala – Grobmotorik ¹⁾ “	CL-303 ²⁾ N = 22			
	n	Gesamtwert MW (SD)	n	Veränderung MW (SD)
Baseline	22	1,5 (1,87)	-	-
Monat 1 – Veränderung gegenüber Baseline	22	3,8 (3,57)	22	2,3 (2,75)
Monat 2 – Veränderung gegenüber Baseline	22	5,0 (4,56)	22	3,5 (4,10)
Monat 3 – Veränderung gegenüber Baseline	22	6,5 (4,90)	22	5,0 (3,94)
Monat 4 – Veränderung gegenüber Baseline	21	8,8 (5,89)	21	7,3 (5,20)
Monat 5 – Veränderung gegenüber Baseline	21	9,3 (6,25)	21	7,7 (5,27)
Monat 6 – Veränderung gegenüber Baseline	21	10,6 (6,54)	21	9,0 (5,84)
Monat 7 – Veränderung gegenüber Baseline	20	12,2 (6,72)	20	10,8 (5,88)
Monat 8 – Veränderung gegenüber Baseline	16	15,1 (5,77)	16	13,3 (5,28)
Monat 9 – Veränderung gegenüber Baseline	19	14,8 (7,08)	19	13,1 (6,60)
Monat 10 – Veränderung gegenüber Baseline	17	15,1 (8,12)	17	13,4 (7,42)
Monat 11 – Veränderung gegenüber Baseline	17	15,4 (8,60)	17	13,6 (7,88)
Monat 12 – Veränderung gegenüber Baseline	17	16,6 (8,34)	17	15,2 (7,40)
Monat 13 – Veränderung gegenüber Baseline	15	15,2 (6,93)	15	13,5 (6,67)

Studie Endpunkt „Bayley-Skala – Grobmotorik ¹⁾ “	CL-303 ²⁾ N = 22			
	n	Gesamtwert MW (SD)	n	Veränderung MW (SD)
Monat 14 – Veränderung gegenüber Baseline	13	19,8 (9,48)	13	18,5 (8,75)
Monat 15 – Veränderung gegenüber Baseline	8	21,0 (11,19)	8	19,6 (10,66)
Monat 16 – Veränderung gegenüber Baseline	5	25,0 (11,51)	5	22,4 (10,62)
Monat 17 – Veränderung gegenüber Baseline	2	34,5 (16,26)	2	32,5 (13,44)
Monat 18 – Veränderung gegenüber Baseline	n. e.	n. e.	n. e.	n. e.

¹⁾ Die Grobmotorik-Subskala der Bayley-Skala besteht aus 72 Items.

²⁾ ITT-Population. Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).

Abkürzungen: ITT: Intention-to-Treat; MW: Mittelwert; n. e.: nicht ermittelt; SD: Standardabweichung.

Tabelle 30: Ergebnisse zum Endpunkt „Bayley-Skala – Feinmotorik“ in der Studie CL-303

Studie Endpunkt „Bayley-Skala – Feinmotorik ¹⁾ “	CL-303 ²⁾ N = 22			
	n	Gesamtwert MW (SD)	n	Veränderung MW (SD)
Baseline	22	6,2 (3,65)	-	-
Monat 1 – Veränderung gegenüber Baseline	22	8,9 (3,88)	22	2,7 (2,59)
Monat 2 – Veränderung gegenüber Baseline	22	11,9 (4,63)	22	5,7 (3,23)
Monat 3 – Veränderung gegenüber Baseline	22	13,3 (4,24)	22	7,0 (3,67)
Monat 4 – Veränderung gegenüber Baseline	21	14,6 (5,71)	21	8,5 (4,97)
Monat 5 – Veränderung gegenüber Baseline	21	16,7 (5,16)	21	10,6 (4,89)
Monat 6 – Veränderung gegenüber Baseline	21	19,3 (5,41)	21	12,9 (5,42)
Monat 7 – Veränderung gegenüber Baseline	20	21,4 (4,80)	20	14,9 (5,10)
Monat 8 – Veränderung gegenüber Baseline	16	24,1 (3,81)	16	18,8 (3,12)
Monat 9 – Veränderung gegenüber Baseline	19	25,6 (3,58)	19	19,2 (3,36)
Monat 10 – Veränderung gegenüber Baseline	17	26,9 (4,05)	17	20,5 (3,95)
Monat 11 – Veränderung gegenüber Baseline	17	27,4 (3,37)	17	21,0 (3,46)
Monat 12 – Veränderung gegenüber Baseline	17	27,8 (4,02)	17	21,6 (3,57)
Monat 13 – Veränderung gegenüber Baseline	15	29,5 (3,48)	15	22,9 (3,84)
Monat 14 – Veränderung gegenüber Baseline	13	28,1 (6,46)	13	23,5 (6,58)
Monat 15 – Veränderung gegenüber Baseline	8	30,3 (3,54)	8	26,4 (5,42)
Monat 16 – Veränderung gegenüber Baseline	5	32,4 (1,52)	5	29,2 (2,77)
Monat 17 – Veränderung gegenüber Baseline	2	34,5 (0,71)	2	32,5 (2,12)
Monat 18 – Veränderung gegenüber Baseline	n. e.	n. e.	n. e.	n. e.

¹⁾ Die Feinmotorik-Subskala der Bayley-Skala besteht aus 66 Items.

²⁾ ITT-Population. Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).

Abkürzungen: ITT: Intention-to-Treat; MW: Mittelwert; n. e.: nicht ermittelt; SD: Standardabweichung.

Studie CL-101 (Datenschnitt: 14. Dezember 2017)

In der Studie CL-101 wurden nur diejenigen Patientinnen und Patienten mit der Bayley-Skala bewertet, die einen CHOP-INTEND-Wert von ≥ 60 Punkten erreichten. Für Patientinnen und Patienten, die an mehr als zwei aufeinanderfolgenden Studienvisiten einen CHOP-INTEND-Wert von 62 Punkten überschritten, wurde ausschließlich die Bayley-Skala erhoben. Insgesamt erreichten vier Personen (25 %) einen CHOP-INTEND-Wert ≥ 60 Punkte und wurden mit der Bayley-Skala bewertet.

CHOP INTEND

Studie CL-303 (Datenschnitt: 12. Dezember 2019)

Die Veränderungen des CHOP INTEND-Gesamtwertes über die Zeit für Studie CL-303 finden sich in Tabelle 31. Die Rücklaufquoten lagen bis Monat 12 bei über 70 %.

Tabelle 31: Ergebnisse zum Endpunkt „CHOP INTEND¹⁾“ in der Studie CL-303

Studie Endpunkt „CHOP INTEND ¹⁾ “	CL-303 ²⁾ N = 22			
	n	Gesamtwert MW (SD)	n	Veränderung MW (SD)
Baseline	22	32,0 (9,69)	-	-
Tag 7 – Veränderung gegenüber Baseline	22	33,2 (9,89)	22	1,2 (4,34)
Tag 14 – Veränderung gegenüber Baseline	22	35,2 (9,36)	22	3,2 (6,28)
Tag 21 – Veränderung gegenüber Baseline	22	37,1 (9,38)	22	5,1 (5,89)
Tag 30 – Veränderung gegenüber Baseline	22	38,9 (8,26)	22	6,9 (5,35)
Monat 2 – Veränderung gegenüber Baseline	22	41,2 (9,03)	22	9,2 (5,88)
Monat 3 – Veränderung gegenüber Baseline	22	43,7 (8,35)	22	11,7 (6,40)
Monat 4 – Veränderung gegenüber Baseline	21	44,1 (8,45)	21	12,9 (6,18)
Monat 5 – Veränderung gegenüber Baseline	21	46,7 (7,55)	21	14,3 (6,61)
Monat 6 – Veränderung gegenüber Baseline	20	46,3 (8,39)	20	14,6 (7,04)
Monat 7 – Veränderung gegenüber Baseline	18	48,8 (6,30)	18	16,4 (6,71)
Monat 8 – Veränderung gegenüber Baseline	15	49,1 (6,09)	15	16,9 (7,56)
Monat 9 – Veränderung gegenüber Baseline	18	48,9 (6,96)	18	16,6 (7,31)
Monat 10 – Veränderung gegenüber Baseline	16	49,3 (6,27)	16	16,9 (8,77)
Monat 11 – Veränderung gegenüber Baseline	16	50,3 (6,49)	16	17,8 (7,00)
Monat 12 – Veränderung gegenüber Baseline	16	48,3 (5,36)	16	16,4 (8,04)
Monat 13 – Veränderung gegenüber Baseline	14	49,8 (5,62)	14	18,8 (9,04)
Monat 14 – Veränderung gegenüber Baseline	11	53,5 (5,66)	11	23,5 (8,49)
Monat 15 – Veränderung gegenüber Baseline	6	51,2 (6,82)	6	21,8 (11,03)
Monat 16 v Veränderung gegenüber Baseline	4	55,0 (3,83)	4	24,8 (5,74)
Monat 17 – Veränderung gegenüber Baseline	1	58,0	1	37,0
Monat 18 – Veränderung gegenüber Baseline	n. e.	n. e.	n. e.	n. e.

¹⁾ Der CHOP INTEND besteht aus 16 Items, die mit 0 (ohne Funktion), 1, 2, 3 oder 4 (volle Funktion) bewertet werden. Es kann maximal ein Wert von 64 Punkten erreicht werden; höhere Werte weisen auf eine bessere motorische Funktion hin.

²⁾ ITT-Population. Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).

Abkürzungen: CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; ITT: Intention-to-Treat; MW: Mittelwert; n. e.: nicht ermittelt; SD: Standardabweichung.

Studie CL-101 (Datenschnitt: 14. Dezember 2017)

Die Veränderungen des CHOP-INTEND-Gesamtwertes über die Zeit für Studie CL-303 finden sich in Tabelle 32. Die Rücklaufquoten lagen bis Monat 12 bei über 70 %.

Tabelle 32: Ergebnisse zum Endpunkt „CHOP INTEND“ in der Studie CL-101

Studie Endpunkt „CHOP INTEND ¹⁾ “	CL-101 ²⁾ N = 12				
	n	Gesamtwert MW (SD)	n	Veränderung MW (%) ³⁾	Veränderung LS MW (SE) ³⁾ p-Wert
Baseline	12	28,2 (12,3)	-	-	-
Monat 1 – Veränderung gegenüber Baseline	12	37,9 (12,7)	12	9,8 (42,0)	9,77 (1,8) < 0,001
Monat 2 – Veränderung gegenüber Baseline	11	42,8 (13,4)	11	14,7 (66,2)	14,92 (1,8) < 0,001
Monat 3 – Veränderung gegenüber Baseline	12	43,6 (13,5)	12	15,4 (66,7)	15,55 (1,8) < 0,001
Monat 4 – Veränderung gegenüber Baseline	11	46,8(13,55)	11	17,6 (74,4)	18,08 (1,9) < 0,001
Monat 5 – Veränderung gegenüber Baseline	11	50,7 (8,6)	11	21,1 (90,9)	22,42 (1,9) < 0,001
Monat 6 – Veränderung gegenüber Baseline	11	49,5 (14,0)	11	21,4 (95,5)	21,79 (1,8) < 0,001
Monat 7 – Veränderung gegenüber Baseline	9	52,3 (8,0)	9	22,7 (100,6)	23,93 (2,1) < 0,001
Monat 8 – Veränderung gegenüber Baseline	10	49,2 (13,9)	10	20,8 (93,4)	21,21 (1,9) < 0,001
Monat 9 – Veränderung gegenüber Baseline	10	54,6 (6,3)	10	24,9 (109,9)	26,95 (2,0) < 0,001
Monat 10 – Veränderung gegenüber Baseline	10	53,8 (5,6)	10	26,2 (118,2)	27,09 (1,9) < 0,001
Monat 11 – Veränderung gegenüber Baseline	11	49,5 (12,2)	11	23,4 (107,7)	23,46 (1,8) < 0,001
Monat 12 – Veränderung gegenüber Baseline	9	53,6 (5,5)	9	25,7 (116,8)	26,62 (2,0) < 0,001
Monat 13 – Veränderung gegenüber Baseline	7	53,7 (6,4)	7	23,1 (90,7)	24,85 (2,5) < 0,001
Monat 14 – Veränderung gegenüber Baseline	3	53,7 (2,3)	3	24,0 (84,3)	26,25 (4,2) < 0,001
Monat 15 – Veränderung gegenüber Baseline	7	56,1 (2,7)	7	29,6 (128,0)	30,17 (2,3) < 0,001

Studie Endpunkt „CHOP INTEND ¹⁾ “	CL-101 ²⁾ N = 12				
	n	Gesamtwert MW (SD)	n	Veränderung MW (% ³⁾)	Veränderung LS MW (SE) ³⁾ p-Wert
Monat 16 – Veränderung gegenüber Baseline	5	52,0 (4,8)	5	25,8 (116,1)	26,15 (2,7) < 0,001
Monat 17 – Veränderung gegenüber Baseline	5	56,6 (5,4)	5	27,0 (110,2)	27,26 (2,7) < 0,001
Monat 18 – Veränderung gegenüber Baseline	6	55,7 (3,7)	6	30,7 (144,0)	30,07 (2,5) < 0,001
Monat 19 – Veränderung gegenüber Baseline	2	58,5 (0,7)	2	31,5 (117,7)	32,42 (5,0) < 0,001
Monat 20 – Veränderung gegenüber Baseline	2	57,0 (1,4)	2	28,0 (96,6)	-
Monat 24 – Veränderung gegenüber Baseline	6	55,5 (1,8)	6	30,7 (145,6)	-

¹⁾ Der CHOP INTEND besteht aus 16 Items, die mit 0 (ohne Funktion), 1, 2, 3 oder 4 (volle Funktion) bewertet werden. Es kann maximal ein Wert von 64 Punkten erreicht werden; höhere Werte weisen auf eine bessere motorische Funktion hin.

²⁾ FAS-Population. Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

³⁾ Auswertung basierend auf MMRM: Das Modell beinhaltet die Veränderung gegenüber dem Baseline-Wert als abhängige Variable sowie fixe Effekte (Kohorte, Visite), Kovariate (Baseline-Wert) und Interaktion (Kohorten-Visiten, Baseline-Visiten).

Abkürzungen: CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; FAS: Full Analysis Set; LS: Least Square; MMRM: Mixed Model of Repeated Measures; MW: Mittelwert; SD: Standardabweichung; SE: Standardfehler; vg: Vektorgenom/e.

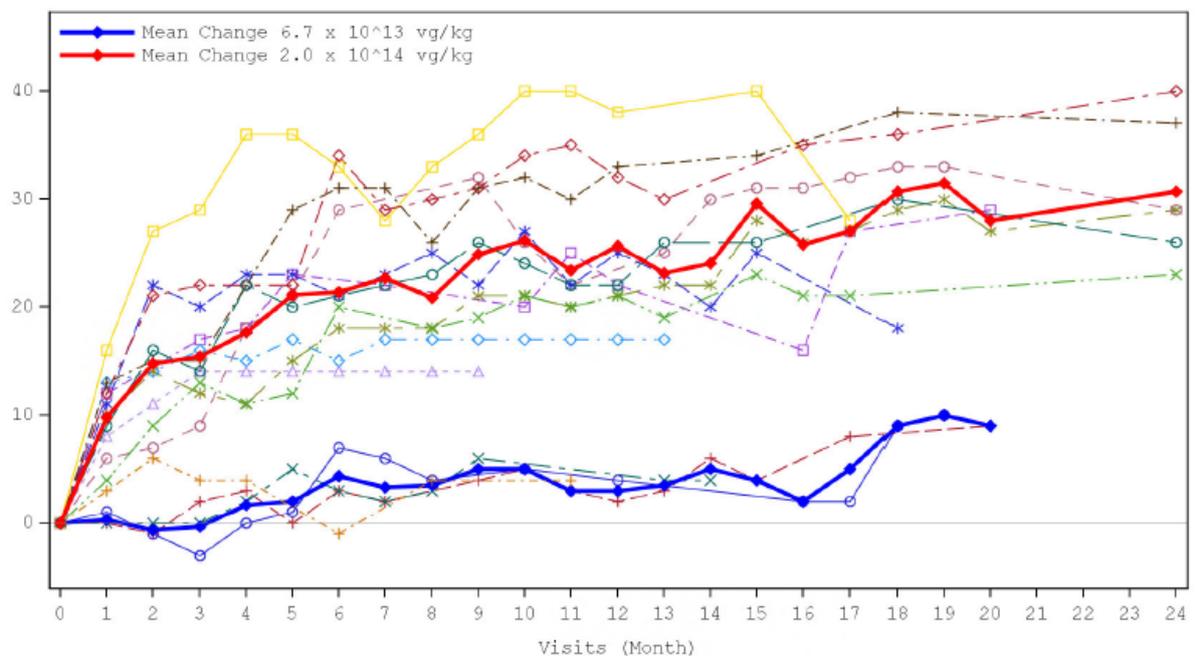


Abbildung 9: CHOP INTEND – Werte-Entwicklung in Studie CL-101 nach Kohorten (Kohorte 2 in rot)

Tabelle 33 kann die Anzahl Personen entnommen werden, die in den Studien CL-303 und CL-101 einen CHOP-INTEND-Gesamtwert von ≥ 40 , ≥ 50 und ≥ 60 Punkten erreichten. Der Gesamtwert konnte in der Studie CL-303 zu einer beliebigen Visite nach Baseline erreicht werden; für die Studie CL-101 liegen diesbezüglich keine Angaben vor.

Tabelle 33: Ergebnisse zum Endpunkt „CHOP-INTEND-Schwellenwerte“ in den Studien CL-303 und CL-101

Studien Endpunkt „CHOP-INTEND-Schwellenwerte“	CL-303 ¹⁾²⁾ N = 22	CL-101 ³⁾⁴⁾ N = 12
CHOP-INTEND-Gesamtwert ≥ 40 Punkte, n (%)	21 (95,5)	11 (91,7)
CHOP-INTEND-Gesamtwert ≥ 50 Punkte, n (%)	14 (63,6)	11 (91,7)
CHOP-INTEND-Gesamtwert ≥ 60 Punkte, n (%) ⁵⁾	5 (22,7)	4 (33,3)

¹⁾ ITT-Population. Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).

²⁾ Angabe der erreichten Meilensteine bezieht sich auf irgendeinen Zeitpunkt innerhalb der Beobachtungsdauer.

³⁾ FAS-Population. Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

⁴⁾ Es ist unklar, ob sich die Angabe der erreichten Meilensteine auf irgendeinen Zeitpunkt innerhalb der Beobachtungsdauer oder auf den finalen Datenschnitt 24 Monate nach Gentransfer bezieht.

⁵⁾ Im Studienprotokoll der Studie CL-303 ist als Endpunkt das Erreichen eines Wertes von ≥ 58 Punkten vorgesehen, derzeit liegt ausschließlich die Auswertung für einen Schwellenwert von 60 Punkten vor.

Abkürzungen: CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; FAS: Full Analysis Set; ITT: Intention-to-Treat; vg: Vektorgenom/e.

3.4 Lebensqualität

Es wurden keine Daten zur Lebensqualität in den Studien erhoben.

3.5 Sicherheit

Studie CL-303 (Datenschnitt: 12.12.2019) und Studie CL-101 (Datenschnitt: 14.12.2017)

Nachfolgend werden die Ergebnisse der Sicherheitsendpunkte für die bereits abgeschlossenen Studien CL-101 und CL-303 dargestellt. Für die Studien CL-302, CL-304 und LT-001 werden in der vorliegenden Nutzenbewertung ausschließlich die aggregierten Sicherheitsendpunkte (siehe Tabelle 34) berichtet.

Die Sicherheitspopulation in den Studien CL-101 und CL-303 umfasste alle Personen, die eine Gentherapie-Infusion erhielten. In den Studien CL-303 und CL-101 war bei jeder Person mindestens ein UE aufgetreten (siehe Tabelle 34).

Die Kinder wurden über 18 (Studie CL-303) bzw. 24 Monate nach Behandlung (Studie CL-101) in der Studie beobachtet. Angaben zur medianen Beobachtungsdauer liegen nicht vor.

Tabelle 34: Ergebnisse der Nebenwirkungen: Zusammenfassung der UE in den Studien CL-303, CL-101, CL-302, CL-304 und LT-001 – Sicherheitspopulation

Studien Personen mit mindestens einem ...	CL-303 ¹⁾ N = 22 n (%)	CL-101 ²⁾ N = 12 n (%)	CL-302 ³⁾⁴⁾ N = 33 n (%)	CL-304 ³⁾⁴⁾		LT-001 ⁵⁾ (Kohorte 2) N = 10 n (%)
				SMN2: 2 Kopien N = 14 n (%)	SMN2: 3 Kopien N = 15 n (%)	
UE	22 (100)	12 (100)	32 (97)	14 (100)	15 (100)	7 (70)
UE CTCAE-Grad ≥ 3	10 (45,5)	10 (83,3)	13 (39,4)	4 (28,6)	2 (13,3)	7 (70)
CTCAE-Grad 3	6 (27,3)	10 (83,3)	10 (30,3)	k. A.	k. A.	k. A.
CTCAE-Grad 4	3 (13,6)	1 (8,3)	2 (6,1)	k. A.	k. A.	k. A.
CTCAE-Grad 5	1 (4,5)	0	1 (3,0)	k. A.	k. A.	k. A.
SUE	10 (45,5)	10 (83,3)	19 (57,6)	4 (28,6)	2 (13,3)	7 (70)
Therapieabbruch aufgrund von UE	2 (9,1)	0	0	0	0	0
UE, das zum Tod führte	1 (4,5)	0	1 (3)	0	0	0

¹⁾ Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019). SUE werden mit der Unterzeichnung der Einverständniserklärung und UE mit Beginn der Gensatztherapie und bis zur letzten Studienvisite erfasst. Der Erfassungszeitraum von AESI ist nicht definiert.

²⁾ Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg. SUE und UE werden mit Studienbeginn oder Verabreichung der Studienmedikation bis zu 30 Tage nach der letzten Studienvisite erfasst. Der Erfassungszeitraum von AESI ist nicht definiert.

³⁾ Interimsdatenschnitt der Studien CL-302 und CL-304 zum 31.12.2019. Zum Datenschnitt beträgt die mediane Zeit (min; max) in der Studie CL-302 11,9 Monate (1,8; 15,4) sowie in der Studie CL-304 in der Kohorte mit 2 SMN2-Genkopien 9,9 Monate (5,1; 18) und in der Kohorte mit 3 SMN2-Genkopien 9,0 Monate (2; 13,9).

⁴⁾ UE werden mit der Unterzeichnung der Einverständniserklärung und bis zur letzten Studienvisite erfasst. SUE und AESI werden mit der Unterzeichnung der Einverständniserklärung bis zu 30 Tage nach der letzten Studienvisite erfasst.

⁵⁾ Interimsdatenschnitt zum 31.12.2019. Für die Studie LT-001 (Kohorte 2 aus der Studie CL-101 (zulassungskonforme Behandlung)) liegen keine Angaben zur medianen Beobachtungszeit vor. SUE und AESI werden mit der Unterzeichnung der Einverständniserklärung bis zur letzten Studienvisite erfasst (Follow-up bis 15 Jahre nach Gentransfer).

Abkürzungen: AESI: Unerwünschte Ereignisse von besonderem Interesse; CTCAE: Common Terminology Criteria for Adverse Events; k. A.: keine Angabe; SMN: Survival Motor Neuron; (S)UE: (Schwerwiegende/s) Unerwünschte/s Ereignis/se; vg: Vektorgenom/e.

Unerwünschte Ereignisse

In Tabelle 35 sind UE mit einer Inzidenz $\geq 10\%$ für die Studien CL-101 und CL-303 dargestellt.

In der Studie CL-101 traten am häufigsten (mehr als 50 %) Erkrankungen folgender Preferred Terms auf: Infektion der oberen Atemwege (83,3 %), Erbrechen (66,7 %), Pyrexie (58,3 %), Pneumonie (58,3 %) sowie Obstipation (50 %) und verstopfte Nase (50 %).

In der Studie CL-303 traten am häufigsten (mehr als 50 %) Erkrankungen folgender Preferred Terms auf: Pyrexie (54,5 %) und Infektion der oberen Atemwege (50 %).

Tabelle 35: UE mit Inzidenz $\geq 10\%$ in den Studien CL-303 und CL-101 – Sicherheitspopulation

Studien MedDRA-Systemorganklasse¹⁾ Preferred Term¹⁾	CL-303²⁾ N = 22 n (%)	CL-101³⁾ N = 12 n (%)
Erkrankungen des Blutes und des Lymphsystems	4 (18,2)	-
Herzerkrankungen	4 (18,2)	3 (25)
Tachykardie	-	2 (16,7)
Kongenitale, familiäre und genetische Erkrankungen	9 (40,9)	-
Trichterbrust (Pectus excavatum)	3 (13,6)	-
Erkrankungen des Gastrointestinaltrakts	19 (86,4)	9 (75)
Obstipation	9 (40,9)	6 (50)
Zahnen	5 (22,7)	n. e.
Diarrhö	4 (18,2)	3 (25)
Gastroösophageale Refluxkrankheit	4 (18,2)	5 (41,7)
Erbrechen	4 (18,2)	8 (66,7)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	14 (63,6)	7 (58,3)
Pyrexie	12 (54,5)	7 (58,3)
Erkrankungen des Immunsystems	-	2 (16,7)
Infektionen und parasitäre Erkrankungen	18 (81,8)	11 (91,7)
Infektion der oberen Atemwege	11 (50)	10 (83,3)
Bronchiolitis	3 (13,6)	3 (25)
Konjunktivitis	3 (13,6)	2 (16,7)
Otitis media	3 (13,6)	2 (16,7)
Pneumonie	3 (13,6)	7 (58,3)
RSV-Bronchiolitis	3 (13,6)	2 (16,7)
Adenovirus Infektion	-	2 (16,7)
Ohrinfektion	-	2 (16,7)
Enterovirus-Infektion	-	4 (33,3)
Virale Gastroenteritis	-	5 (41,7)
Rhinovirusinfektion	-	4 (33,3)
Parainfluenzavirus-Infektion	-	3 (25)
Streptokokken-Pharyngitis	-	2 (16,7)
RSV-Pneumonie	-	2 (16,7)
Virusbedingte Infektion der oberen Atemwege	-	3 (25)
Harnwegsinfektion	-	2 (16,7)

Studien MedDRA-Systemorganklasse¹⁾ Preferred Term¹⁾	CL-303²⁾ N = 22 n (%)	CL-101³⁾ N = 12 n (%)
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	9 (40,9)	6 (50)
Sturz	-	3 (25)
Kontusion	4 (18,2)	-
Spinnenbiss	3 (13,6)	-
Untersuchungen	15 (68,2)	5 (41,7)
Erhöhung der Aspartat-Aminotransferase	6 (27,3)	-
Erhöhung der Alanin-Aminotransferase	5 (22,7)	-
Erhöhung der Transaminasen	-	3 (25)
Positiver Test auf humanen Rhinovirus	-	3 (25)
Positiver Test auf Enteroviren	-	2 (16,7)
Stoffwechsel- und Ernährungsstörungen	8 (36,4)	3 (25)
Fütterungsstörung	3 (13,6)	-
Dehydration	-	2 (16,7)
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen	12 (54,5)	3 (25)
Skoliose	9 (40,9)	-
Erkrankungen des Nervensystems	5 (22,7)	-
Psychiatrische Erkrankungen	3 (13,6)	-
Erkrankungen der Atemwege, des Brustraums und Mediastinums	17 (77,3)	12 (100)
Husten	7 (31,8)	5 (41,7)
Atemnot	6 (27,3)	2 (16,7)
Abnormale Atmung	5 (22,7)	-
Verwendung der zusätzlichen Atemmuskeln	5 (22,7)	-
Verstopfte Nase	3 (13,6)	6 (50)
Schlafapnoe-Syndrom	3 (13,6)	-
Tachypnoe	3 (13,6)	-
Atelektase	-	4 (33,3)
Respiratorische Insuffizienz	-	3 (25)
Rhinorrhö	-	3 (25)
Aspirationspneumonie	-	2 (16,7)
Giemen	-	2 (16,7)
Erkrankungen der Haut und des Unterhautzellgewebes	13 (59,1)	8 (66,7)
Hautausschlag	5 (22,7)	5 (41,7)
Atopische Dermatitis	3 (13,6)	-
Ekzem	3 (13,6)	-
Dekubitus	-	2 (16,7)

¹⁾ Bei mehreren Ereignissen einer/eines Studienteilnehmenden in einer bestimmten Systemorganklasse oder einem Preferred Term wurde dies als einmaliges Ereignis in der Systemorganklasse oder im Preferred Term gezählt.

²⁾ Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019). UE werden mit Beginn der Gensatztherapie und bis zur letzten Studienvisite erfasst.

³⁾ Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg. UE werden mit Studienbeginn oder Verabreichung der Studienmedikation bis zu 30 Tage nach der letzten Studienvisite erfasst.

Abkürzungen: n. e.: nicht ermittelt; MedDRA: Medical Dictionary for Regulatory Activities; RSV: Respiratory Syncytial Virus; UE: Unerwünschte/s Ereignis/se; vg: Vektorgenom/e.

Unerwünschte Ereignisse des Schweregrades ≥ 3 (≥ 5 %)

In Tabelle 36 sind schwere UE mit einer Inzidenz ≥ 5 % für die Studien CL-101 und CL-303 dargestellt.

In der Studie CL-101 traten am häufigsten (mehr als 15 %) Erkrankungen folgender Preferred Terms auf: Atelektase (25 %) und Pneumonie (58,3 %). Alle anderen gelisteten Preferred Terms traten jeweils bei 2 von 12 Personen (16,7 %) auf.

In der Studie CL-303 traten am häufigsten (mehr als 10 %) Erkrankungen folgender Preferred Terms auf: RSV-Bronchiolitis und Atemnot (jeweils 13,6 %).

Tabelle 36: Schwere UE (Grad ≥ 3) in den Studien CL-303 und CL-101 – Sicherheitspopulation

Studien MedDRA-Systemorganklasse¹⁾ Preferred Term¹⁾	CL-303²⁾ N = 22 n (%)	CL-101³⁾ N = 12 n (%)
Infektionen und parasitäre Erkrankungen	7 (31,8)	8 (66,7)
RSV-Bronchiolitis	3 (13,6)	-
Enterovirus-Infektion	-	2 (16,7)
Parainfluenzavirus-Infektion	-	2 (16,7)
Pneumonie	-	7 (58,3)
RSV-Pneumonie	-	2 (16,7)
Rhinovirusinfektion	-	2 (16,7)
Infektion der oberen Atemwege	-	2 (16,7)
Untersuchungen	3 (13,6)	3 (25)
Positiver Test auf humanen Rhinovirus	-	2 (16,7)
Stoffwechsel- und Ernährungsstörungen	2 (9,1)	-
Erkrankungen der Atemwege, des Brustraums und Mediastinums	6 (27,3)	6 (50)
Atemnot	3 (13,6)	-
Atemstillstand	2 (9,1)	-
Atelektase	-	3 (25)
Aspirationspneumonie	-	2 (16,7)
Atemstillstand	-	2 (16,7)
Gefäßkrankungen	2 (9,1)	-

¹⁾ Bei mehreren Ereignissen einer/eines Studienteilnehmenden in einer bestimmten Systemorganklasse oder einem Preferred Term wurde dies als einmaliges Ereignis in der Systemorganklasse oder im Preferred Term gezählt.

²⁾ Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019). UE werden mit Beginn der Gensersatztherapie und bis zur letzten Studienvisite erfasst.

³⁾ Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg. UE werden mit Studienbeginn oder Verabreichung der Studienmedikation bis zu 30 Tage nach der letzten Studienvisite erfasst.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; RSV: Respiratory Syncytical Virus; UE: Unerwünschte/s Ereignis/se; vg: Vektorgenom/e.

Schwerwiegende unerwünschte Ereignisse

In Tabelle 37 sind SUE mit einer Inzidenz $\geq 5\%$ für die Studien CL-101 und CL-303 dargestellt.

In der Studie CL-101 traten am häufigsten (mehr als 15 %) Erkrankungen folgender Preferred Terms auf: Pneumonie (58,3 %) und Infektion der oberen Atemwege (25 %). Alle anderen gelisteten Preferred Terms traten jeweils bei 2 von 12 Personen (16,7 %) auf.

In der Studie CL-303 trat am häufigsten (mehr als 10 %) folgender Preferred Term auf: Atemnot (18,2 %).

Tabelle 37: SUE mit Inzidenz $\geq 5\%$ in den Studien CL-303 und CL-101 – Sicherheitspopulation

Studien MedDRA-Systemorganklasse¹⁾ Preferred Term¹⁾	CL-303³⁾ N = 22 n (%)	CL-101²⁾ N = 12 n (%)
Infektionen und parasitäre Erkrankungen	7 (31,8)	9 (75)
Bronchiolitis	2 (9,1)	-
Pneumonie	2 (9,1)	7 (58,3)
RSV-Bronchiolitis	2 (9,1)	2 (16,7)
Adenovirus-Infektion	-	2 (16,7)
Parainfluenzavirus-Infektion	-	2 (16,7)
RSV-Pneumonie	-	2 (16,7)
Rhinovirus-Infektion	-	2 (16,7)
Enterovirus-Infektion	-	2 (16,7)
Infektion der oberen Atemwege	-	3 (25)
Untersuchungen	3 (13,6)	4 (33,3)
Stoffwechsel- und Ernährungsstörungen	2 (9,1)	-
Erkrankungen der Atemwege, des Brustraums und Mediastinums	5 (22,7)	-
Atemnot	4 (18,2)	2 (16,7)
Atemstillstand	2 (9,1)	-
Aspirationspneumonie	-	2 (16,7)

¹⁾ Bei mehreren Ereignissen einer/eines Studienteilnehmenden in einer bestimmten Systemorganklasse oder einem Preferred Term wurde dies als einmaliges Ereignis in der Systemorganklasse oder im Preferred Term gezählt.

²⁾ Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt zum 12.12.2019). SUE werden mit der Unterzeichnung der Einverständniserklärung bis zur letzten Studienvisite erfasst.

³⁾ Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg. SUE werden mit Studienbeginn oder Verabreichung der Studienmedikation bis zu 30 Tage nach der letzten Studienvisite erfasst.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; RSV: Respiratory Syncytical Virus; SUE: Schwerwiegende/s unerwünschte/s Ereignis/se; vg: Vektorgenom/e.

Unerwünschte Ereignisse von besonderem Interesse (AESI)

In der Studie CL-101 wurden erhöhte Leberenzymwerte als AESI präspezifiziert. Bei drei Patienten (25 %) traten erhöhte Leberenzymwerte auf (Erhöhung der Transaminasen (N = 3), Erhöhung der Aspartat-Aminotransferase (N = 1)). Die beobachteten Erhöhungen der Leberwerte waren klinisch asymptomatisch, d. h. weder Gelbsucht noch Symptome einer eingeschränkten Leberfunktion wurden berichtet. Bei keinem der Patienten war Hy's Law erfüllt. Alle Ereignisse waren im Verlauf der Studie reversibel.

In der Studie CL-303 wurden „Hepatotoxizität“, „Thrombozytopenie“, „Kardiale Ereignisse“, „Auf Ganglionopathie hindeutende sensorische Anomalien“ als AESI präspezifiziert. Insgesamt wurden bei 13 von 22 Personen (59,1 %) AESI berichtet. Darunter waren 7 Personen (31,8 %) mit mindestens einem Ereignis in der Kategorie „Hepatotoxizität“, 8 (36,4 %) mit mindestens einem Ereignis in der Kategorie „Thrombozytopenie“, 4 (18,2 %) mit mindestens einem Ereignis in der Kategorie „Kardiale Ereignisse“ und 5 (22,7 %) mit mindestens einem Ereignis in der Kategorie „Auf Ganglionopathie hindeutende sensorische Anomalien“ (siehe Tabelle 38).

Tabelle 38: (Prä-)Spezifizierte unerwünschte Ereignisse von besonderem Interesse in der Studie CL-303 – Sicherheitspopulation

Studie Endpunkt „Unerwünschte Ereignisse von besonderem Interesse“	CL-303¹⁾ N = 22 <i>n (%)</i>
Hepatotoxizität	
Erhöhung der Aspartat-Aminotransferase	6 (27,3)
Erhöhung der Alanin-Aminotransferase	5 (22,7)
Erhöhung der Gammaglutamyltransferase	2 (9,1)
Erhöhung der Transaminasen	2 (9,1)
Ammoniak erhöht	1 (4,5)
Thrombozytopenie	
Kontusion	4 (18,2)
Blut im Stuhl	2 (9,1)
Thrombozytopenie	2 (9,1)
Blut im Urin	1 (4,5)
Thrombozytenzahl verringert	1 (4,5)
Hämaturie	1 (4,5)
Kardiale Ereignisse	
Kreatinkinase-MB im Blut erhöht	2 (9,1)
Tachykardie	1 (4,5)
Diastolischer Bluthochdruck	1 (4,5)
Systolischer erniedrigter Blutdruck	1 (4,5)
Ganglionopathie	
Ungewollte Muskelkontraktionen	2 (9,1)
Hydrozephalus	1 (4,5)
Hypotonie	1 (4,5)
Tremor	1 (4,5)
Extrapyramidale Erkrankung	1 (4,5)
Hyporeflexie	1 (4,5)
Lethargie	1 (4,5)

¹⁾ Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019). Der Erfassungszeitraum von AESI ist nicht definiert.

Abkürzungen: AESI: Unerwünschte Ereignisse von besonderem Interesse.

Hospitalisierung

Studie CL-303 (finaler Datenschnitt: 12. Dezember 2019)

Die mittlere annualisierte Hospitalisierungsrate (Standardfehler, SE) lag bei den Personen in der Studie CL-303 bei 0,8 (0,2). Gemäß den Angaben in den Studienunterlagen wurden sowohl ungeplante als auch geplante Hospitalisierungen erfasst. Es ist nicht ersichtlich, wie viele Hospitalisierungen davon geplant bzw. ungeplant waren.

Studie CL-101 (finaler Datenschnitt: 14. Dezember 2017)

Die mittlere annualisierte Hospitalisierungsrate (SE) lag bei den Personen in der Studie CL-101 bei 2,08 (0,6). Es ist nicht ersichtlich, ob sowohl geplante als auch ungeplante Krankenhausaufenthalte berücksichtigt wurden und wie viele Hospitalisierungen davon geplant bzw. ungeplant waren.

Tabelle 39: Ergebnisse zum Endpunkt „Hospitalisierung“ in den Studien CL-101 und CL-303

Studien Endpunkt „Hospitalisierung“	CL-303¹⁾²⁾ N = 22	CL-101³⁾⁴⁾ N = 12
Personen mit Hospitalisierung, n (%)	11 (50)	10 (83,3)
Anzahl an Hospitalisierungen (insgesamt)	21	51
Tage im Krankenhaus (insgesamt)	k. A.	397
<i>Anzahl Hospitalisierungen pro Person</i>		
MW (SE)	k. A.	4,25 (1,3)
Median (min; max)	k. A.	2,5 (0; 16)
<i>Verweildauer im Krankenhaus (in Tagen)</i>		
MW (SE)	k. A.	6,7 (1,2)
Median (min; max)	k. A.	5,8 (3,0; 12,1)
<i>Annualisierte Hospitalisierungsrate</i>		
MW (SE)	0,8 (0,2)5)	2,1 (0,6)
Median (min; max)	k. A.	1,3 (0; 7,6)
<i>Annualisierte Anzahl an Tagen im Krankenhaus</i>		
MW (SE)	k. A.	16,1 (6,1)
Median (min; max)	k. A.	8,4 (0; 66,8)

¹⁾ Geplantes Studienende: Bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019).

²⁾ Gemäß den Angaben in den Studienunterlagen wurden sowohl ungeplante als auch geplante Hospitalisierungen erfasst. Es ist nicht ersichtlich, wie viele Hospitalisierungen davon geplant bzw. ungeplant waren.

³⁾ Geplantes Studienende: 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben der Studie CL-101 beziehen sich auf Kohorte 2 mit der Dosierung $2,0 \times 10^{14}$ vg/kg.

⁴⁾ Aus den Studienunterlagen geht nicht hervor, ob sowohl geplante als auch ungeplante Krankenhausaufenthalte berücksichtigt wurden und wie viele Hospitalisierungen davon geplant bzw. ungeplant waren.

⁵⁾ Nicht-adjustierte annualisierte Rate.

Abkürzungen: k. A.: keine Angabe; MW: Mittelwert, SE: Standardfehler; vg: Vektorgenom/e.

4 Diskussion der Methodik und Ergebnisse

4.1 Zulassungsstatus und Zulassungspopulation von Onasemnogen-Abeparvovec (Zolgensma®)

Gemäß Zulassung wird Onasemnogen-Abeparvovec (Zolgensma®) angewendet zur Behandlung 5q13-assoziiertes spinale Muskelatrophie (SMA) mit biallelischer Deletion des SMN1-Gens und entweder einer klinischen Diagnose SMA Typ 1 oder bis zu drei Genkopien des SMN2-Gens [9]. Onasemnogen-Abeparvovec unterhält eine Zulassung als Arzneimittel zur Behandlung eines seltenen Leidens nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlamentes und des Rates vom 16. Dezember 1999 über Arzneimittel für seltene Leiden. Die Zulassung durch die EMA (European Medicines Agency) erfolgte unter „Besonderen Bedingungen“ (siehe Tabelle 1) und besteht seit dem 18. Mai 2020. Die vorliegende Nutzenbewertung zu Onasemnogen-Abeparvovec basiert auf den Zulassungsstudien CL-303 (STR1VE-US) und CL-101 (START) [5,6]. In den zulassungsbegründenden Studien liegen Daten für Patientinnen und Patienten mit infantiler SMA (Typ 1) und 2 SMN2-Genkopien vor. Für Patientinnen und Patienten mit 1 SMN2-Genkopie liegen keine Daten in den zulassungsbegründenden Studien vor. Patientinnen und Patienten mit, ebenfalls vom Anwendungsgebiet umfassten, 3 SMN2-Genkopien sind in die Studie CL-304 eingeschlossen. Die Studien CL-304 und CL-302 sowie die Langzeitnachbeobachtung LT-001 der Studie CL-101 sind laufende Studien; abschließende (finale) Ergebnisse liegen noch nicht vor.

Von einer Übertragbarkeit der Zulassungspopulation auf den deutschen Versorgungskontext ist auszugehen.

4.2 Design und Methodik der Studie

Studie CL-303 (STR1VE-US) ist eine einarmige, offene, multizentrische Phase-III-Studie. Das Ziel der Studie bestand darin, Wirksamkeit und Sicherheit von Onasemnogen-Abeparvovec als einmalige intravenöse Injektion bei symptomatischen und präsymptomatischen Kleinkindern mit SMA Typ 1 zu untersuchen. In die Studie eingeschlossen waren 22 Patientinnen und Patienten mit biallelischer Mutation des SMN1-Gens (Deletion oder Punktmutation) mit 1 oder 2 SMN2-Genkopien im Alter < 6 Monate (< 180 Tage) zum Zeitpunkt des Gentransfers. Die Studie gliederte sich in 3 Phasen: Screening, Gentransfer (stationär) und Nachbeobachtung bis zum Alter von 18 Monaten; die Studie wurde im Dezember 2019 beendet. Die Patientinnen und Patienten hatten die Möglichkeit im Anschluss an die Studie in die Langzeitbeobachtungsstudie (LT-002) überzugehen (Nachbeobachtung bis zu 15 Jahre).

Studie CL-101 (START) ist eine einarmige, offene, monozentrische Phase-I-Studie. Primäres Ziel der Studie war die Untersuchung der Sicherheit verschiedener Dosierungen von Onasemnogen-Abeparvovec als einmalige intravenöse Injektion bei symptomatischen Kleinkindern mit SMA Typ 1. In die Studie eingeschlossen waren Patientinnen und Patienten im Alter ≤ 6 Monate zum Zeitpunkt des Gentransfers und einer biallelischen Mutation des SMN1-Gens (Deletion oder Punktmutation) und 2 SMN2-Genkopien sowie Auftreten der ersten Symptome in einem Alter zwischen 0 und 6 Monaten. Die Studie wurde mit vier verschiedenen Kohorten mit niedriger (Kohorte 1), mittlerer (Kohorten 2A und 2B) und hoher Dosierung (Kohorte 3) der Studienmedikation geplant. Für die vorliegende Nutzenbewertung sind Kohorte 2A und 2B mit mittlerer Dosierung relevant, die der fachinformationskonformen Dosierung von Onasemnogen-Abeparvovec entspricht. Die Kohorten 2A und 2B wurden zusammengefasst und umfassen insgesamt 12 Personen. Studie

CL-101 gliederte sich ebenfalls in 3 Phasen: Screening, Gentransfer (stationär) sowie eine zwei-jährige Nachbeobachtungsphase für die Bewertung der Sicherheit. Die Studie wurde nach einer geplanten Laufzeit von 24 Monaten nach Gentransfer im Dezember 2017 beendet. Die Patientinnen und Patienten konnten im Anschluss zur kontinuierlichen Sicherheitsüberwachung über einen Zeitraum von bis zu 15 Jahren an der Langzeitbeobachtung der Studie CL-101 (Studie [LT-001](#)) teilnehmen. Dort erfolgen jährliche Studienvisiten über einen Zeitraum von 5 Jahren und jährliche Telefonkontakte über einen Zeitraum von 10 Jahren. Die Beendigung der Studie LT-001 ist für Dezember 2033 vorgesehen. Im aktuellen Zwischenbericht der Studie LT-001 vom 31. Dezember 2019 liegen Informationen für 10 Patientinnen und Patienten vor. Vier der eingeschlossenen Personen erhielten zu Studienbeginn der Studie LT-001 eine Nusinersen-Behandlung; nach einem Jahr waren es acht Personen. Warum eine Behandlung mit Nusinersen erfolgte und ob diese im Zusammenhang mit der Langzeitwirksamkeit von Onasemnogen-Abeparvovec zu betrachten ist, kann aufgrund fehlender Informationen nicht beurteilt werden. Hierzu wären Angaben wünschenswert.

In den Zulassungsstudien CL-303 und CL-101 wiesen alle Kinder mit infantiler SMA (Typ 1) 2 SMN2-Genkopien auf. Das mediane Alter zu Baseline (min; max) lag in Studie CL-303 bei 3,7 Monaten (0,5; 5,9) und in Studie CL-101 bei 3,4 Monaten (0,9; 7,9). Das Gestationsalter lag in beiden Studien im Median bei 39,0 Wochen. Der Anteil weiblicher Kleinkinder beträgt in Studie CL-303 54,5 %, in Studie CL-101 58,3 %, der Anteil der männlichen entsprechend 45,5 % und 41,7 %. In beiden Studien erhielten alle Studienteilnehmenden begleitend Prednisolon über eine mediane Dauer (min; max) von 62,5 (37; 203) bzw. 61,0 (44; 161) Tagen (CL-303 und CL-101). Und ebenfalls in beiden Studien wiesen nahezu alle Patientinnen und Patienten zu Studienbeginn SMA-spezifischen Symptome wie Hypotonie sowie Schwäche der Gliedmaßen auf und sind dementsprechend als symptomatisch einzustufen. In Studie CL-303 konnten alle Patientinnen und Patienten dünnflüssige Nahrung schlucken, in Studie CL-101 waren es nur 33,3 % der Studienpopulation. In Studie CL-101 erhielten 41,7 % der Studienpopulation eine nicht-orale Nahrungsunterstützung, während in Studie CL-303 niemand davon betroffen war. Bei einem Kleinkind in Studie CL-101 war zu Studienbeginn zudem eine Beatmungsunterstützung erforderlich, in Studie CL-303 war eine Beatmungsunterstützung für niemanden nötig. Von den 22 eingeschlossenen Patientinnen und Patienten der Studie CL-303 brachen drei die Studie vorzeitig ab – aufgrund von UE, Tod oder Entzug der Einverständniserklärung. Obwohl es per Einschlusskriterium möglich war, nahmen an der Studie CL-303 keine präsymptomatischen Patientinnen und Patienten teil. In Studie CL-101 haben alle 12 eingeschlossenen Patientinnen und Patienten die Studie abgeschlossen. Angaben zur medianen Studiendauer liegen für die Studien CL-303 und CL-101 nicht vor. Das Verzerrungspotential der einarmigen zulassungsbezüglichen Studien CL-303 und CL-101 wird mit hoch bewertet.

Weitere potentiell relevante Studien des pU zum Wirkstoff Onasemnogen-Abeparvovec sind die Studien [CL-302](#) und [CL-304](#).

Studie [CL-302 \(STR1VE-EU\)](#) ist eine einarmige, offene, multizentrische Phase-III-Studie und eine europäische Version der Studie CL-303 (STR1VE-US). In der Studie werden Wirksamkeit und Sicherheit von Onasemnogen-Abeparvovec als einmalige intravenöse Injektion bei 33 symptomatischen und präsymptomatischen Kleinkindern mit SMA Typ 1 mit biallelischer Mutation des SMN1-Gens und 1 oder 2 SMN2-Genkopien untersucht, die zum Zeitpunkt des Gentransfers ≤ 6 Monate alt sind. Die Studie endet, wenn die Kinder ein Alter von 18 Monaten erreichen. Das geschätzte

Studienende ist für das 3. Quartal 2020 geplant. Im Anschluss an die Studie haben die Patientinnen und Patienten die Möglichkeit an einer Langzeitnachbeobachtung (LT-002) teilzunehmen.

Studie CL-304 ist eine einarmige, offene, multizentrische Phase-III-Studie. Untersucht werden Wirksamkeit und Sicherheit von Onasemnogen-Abeparvovec als einmalige intravenöse Injektion bei genetisch diagnostizierten und präsymptomatischen Kleinkindern mit SMA Typ 1 mit biallelicher Mutation des SMN1-Gens mit 2 oder 3 SMN2-Genkopien, die zum Zeitpunkt des Gentransfers ≤ 6 Wochen alt sind. Patientinnen und Patienten werden abhängig von der Anzahl ihrer SMN2-Genkopien ausgewertet, 2 SMN2-Genkopien N = 14 und 3 SMN2-Genkopien N = 15. Die Studie endet, wenn die Kinder ein Alter von 18 (2 SMN2-Genkopien) bzw. 24 Monaten (3 SMN2-Genkopien) erreichen. Ein Vergleich mit der natürlichen Verlaufs-Kohorte der PNCR-Studie ist vorgesehen. Das Studienende ist für das 4. Quartal 2020 (2 SMN2-Genkopien) bzw. das 2. Quartal 2021 (3 SMN2-Genkopien) geplant. Auch für diese Patientinnen und Patienten besteht die Möglichkeit im Anschluss an die Studie an einer Langzeitnachbeobachtung (LT-002) teilzunehmen.

Für die Studien CL-302 und CL-304 sind aktuell lediglich Interimsanalysen mit letztem Datenschnitt vom 31. Dezember 2019 verfügbar; Ergebnisse werden für die Endpunkte „Mortalität“, „Tod oder dauerhafte Beatmung (Überleben)“ und die Sicherheitsendpunkte ergänzend berichtet.

4.3 Externe Kontrollen (PNCR, NeuroNEXT, ENDEAR)

Aufgrund des einarmigen Designs der zulassungsbegründenden Studien CL-101 und CL-303 wurden vom pU nicht-adjustierte indirekte Vergleiche (ohne Brückenkomparator) in den Endpunktkategorien Mortalität, Morbidität und Sicherheit vorgelegt. Hierfür wurden folgende Studien zum natürlichen Verlauf der SMA-Erkrankung eingereicht: PNCR [18] und NeuroNEXT [25,26] sowie die ENDEAR-Studie (BSC-Arm) zu Nusinersen [19,21]. Eine weitere Publikation von Finkel et al. (2014 b) [20] wurde vom pU ebenfalls als externe Kontrolle eingereicht, allerdings aufgrund fehlender Angaben zum Patientenkollektiv und einem vorrangig qualitativen Vergleich (ohne Effektschätzer) nicht in der Nutzenbewertung berücksichtigt.

Bei den Studien PNCR und NeuroNEXT handelt es sich um prospektive Studien zum natürlichen Verlauf der SMA. In der Studie PNCR war kein Vergleich mit einer Kontrollgruppe vorgesehen; in der Studie NeuroNEXT wurde eine Kontrollgruppe gesunder Kleinkinder berücksichtigt, welche für die Fragestellung der Nutzenbewertung jedoch nicht relevant ist. Bei der Studie ENDEAR handelt es sich um eine Phase-III-Studie zur Wirksamkeit und Sicherheit von Nusinersen im Vergleich zu einer Scheinintervention (unter BSC). Für die Beurteilung des natürlichen Verlaufs der Erkrankung wurde vom pU die BSC-Kohorte herangezogen.

Studie PNCR

In die Studie PNCR wurden sowohl bereits behandelte als auch neu-diagnostizierte Personen mit SMA Typ 1 und 2 an drei Studienzentren in den USA eingeschlossen. Die Diagnose SMA inkl. Klassifizierung des vorliegenden SMA-Typs erfolgte durch die Studienleitungen der Zentren (nähere Angaben siehe Tabelle 10). Die Einschlusskriterien umfassten primär die SMA-Diagnose, andere Ein- und Ausschlusskriterien wurden nicht definiert. Der Einschluss des Patientenkollektivs in die Studie erfolgte zwischen Mai 2005 und April 2009. Die Beobachtungsdauer war bis zu 36 Monate geplant. Studienvisiten erfolgten zur Baseline, in den Monaten 2, 4, 6, 9 und 12 sowie im Anschluss alle 6 Monate. Aus dem Studienbericht geht nicht hervor, welche und wie viele Patientendaten prospektiv und/oder retrospektiv erfasst wurden bzw. welche Daten retrospektiv in

die Analyse eingingen [18]. Die Datenbasis der Studie PNCr umfasste insgesamt 337 Personen mit SMA Typen 1–3 (RPT-806 [7]), in einer weiteren Publikation von Finkel et. al (2014 a) [18] wird eine Patientenpopulation von 79 Personen mit SMA Typen 1–2 angegeben. Es bestehen Diskrepanzen hinsichtlich Patienteneinschluss bzw. Patientenfluss für Personen mit SMA Typ 1 in die natürliche Vergleichskohorte. Hierbei stehen 81 von 337 Personen mit SMA Typ 1 (RPT-806, siehe Abbildung 6) 34 von 79 Personen mit SMA Typ 1 (Finkel et al. (2014 a)) gegenüber. Aufgrund fehlender Angaben in den Studienunterlagen ist nicht ersichtlich, welche zusätzlichen Einschluss- bzw. Selektionskriterien zu einer Abnahme der Population von 79 bzw. 81 auf 34 Personen führten. Für die beschriebenen Patientenpopulationen (N = 337, N = 81, N = 79 und N = 34) wurden vom pU keine Informationen zu soziodemographischen und krankheitsbezogenen Charakteristika eingereicht.

Aus der PNCr-Studie wurde eine Zielpopulation von 23 Personen für den Vergleich mit den Zulassungsstudien CL-101 und CL-303 ausgewählt, da diese die folgenden Charakteristika erfüllten:

- Alter zu Beginn der Symptome: ≤ 6 Monate,
- biallelische Deletion des SMN1-Gens und
- zwei SMN2-Genkopien.

Eine Untersuchung nach der genetischen Modifikation c.859G>C war nicht vorgesehen, demnach konnten auch Personen mit diesem Merkmal in die Studie eingeschlossen worden sein [7]. Nach Angaben des pU sei das Vorliegen dieses genetischen Modifikators mit einem geringeren klinischen Schweregrad assoziiert. Für eine Beurteilung dieses Zusammenhangs konnten keine Publikationen identifiziert werden.

Studie NeuroNEXT

In die Studie NeuroNEXT waren genetisch diagnostizierte Kleinkinder mit SMA (Typen nicht näher spezifiziert) und gesunde Kleinkinder in 14 Studienzentren in den USA eingeschlossen, die ein Alter von ≤ 6 Monaten aufwiesen und zwischen der 36. und 42. Schwangerschaftswoche geboren wurden. Die Diagnostik der SMA wurde entweder durch Studienärztinnen/Studienärzte oder der/dem behandelnden Neurologin/Neurologen vorgenommen (siehe Tabelle 10). Vor Aufnahme in die Studie hatte eine Bestätigung der SMA durch einen Gentest zu erfolgen. Auch asymptomatische Kinder mit genetisch diagnostizierter SMA konnten in die Studie eingeschlossen werden. Ausgeschlossen wurden Kinder, die mehr als 12 Stunden/Tag eine nicht-invasive ventilatorische Unterstützung benötigten (z. B. BiPAP), eine Komorbidität aufwiesen oder an einer anderen therapeutischen klinischen Studie teilnahmen. Angaben zu möglichen Komorbiditäten liegen nicht vor. Patientinnen und Patienten, die irgendeine Intervention mit Potential die SMN-Expression zu erhöhen (z. B. Valproinsäure) erhielten, waren ebenfalls von der Studie ausgeschlossen. Studienvisiten erfolgten zu Baseline, in den Monaten 3 (wenn möglich), 6, 9, 12, 18 und 24. Der Einschluss des Patientenkollektivs in die Studie erfolgte zwischen November/Dezember 2012 bis September 2014. Die Beobachtungsdauer war bis zu 24 Monate geplant. Insgesamt wurden 26 Kinder mit SMA Typ 1 in die NeuroNEXT-Studie eingeschlossen. Informationen zu soziodemographischen und krankheitsbezogenen Charakteristika der Studienpopulation finden sich in Anhang 2 (Tabelle 42).

Aus der NeuroNEXT-Studie wurden letztendlich 16 Personen (siehe Abbildung 7) für den indirekten Vergleich mit den Zulassungsstudien CL-101 und CL-303 ausgewählt, da diese die folgenden Charakteristika erfüllten:

- Alter von ≤ 6 Monaten,
- biallelische Deletion des SMN1-Gens,
- zwei SMN2-Genkopien und
- Abwesenheit von c.859G>C.

Studie ENDEAR

Die Studie ENDEAR ist eine randomisiert kontrollierte Phase-III-Studie zur Wirksamkeit und Sicherheit von Nusinersen im Vergleich zu einer Scheinintervention. Die Studie gliederte sich in eine 21-tägige Screeningphase, eine anschließende Behandlungsphase von 10 Monaten und eine Nachbeobachtung von 3 Monaten. Die Studie wurde in einem Zeitraum von August 2014 bis November 2016 in 31 Studienzentren in 13 Ländern durchgeführt. Nusinersen wurde gegenüber einer Scheinintervention in Kombination mit BSC verglichen. Die Vergleichsgruppe (BSC) wird als zusätzliche Kontrollgruppe für Onasemnogen-Abeprarvovec herangezogen. Der BSC-Arm der Studie ENDEAR erhielt an denselben Studientagen (Tag 1, 15, 29, 64, 183 und 302, über eine 10-monatige Behandlungsphase) wie der Nusinersen-Arm eine Scheininjektion in Form eines Nadelstichs am unteren Rücken (ohne Lumbalpunktion). Die Injektionsstelle wurde analog zur Nusinersen-Gruppe versorgt. Wesentliche Einschlusskriterien waren:

- Homozygote Deletion oder homozygote Mutation des SMN1-Gens,
- erste Symptome in einem Alter von ≤ 6 Monaten,
- zwei SMN2-Genkopien und
- Adäquate Nahrungs- und Flüssigkeitszufuhr (mit oder ohne Gastrostomie) nach Einschätzung des ärztlichen Prüfpersonals.

Anhand der zuvor beschriebenen, a priori definierten Einschlusskriterien der externen Kontrollstudien (PNCr, NeuroNEXT, ENDEAR), sollte eine möglichst homogene Population für einen indirekten Vergleich mit dem Patientenkollektiv der Studien CL-303 und CL-101 identifiziert werden. Jedoch zeigen sich sowohl bezogen auf diese Einschlusskriterien im direkten Vergleich der Zulassungsstudien als auch zwischen den externen Kontrollen Unterschiede. Während in den Studien CL-303 und NeuroNEXT Personen mit einer genetischen Modifikation c.859G>C im SMN2-Gen ausgeschlossen waren, wurden diese in den Studien CL-101 und PNCr nicht explizit ausgeschlossen. Sowohl für die Studie PNCr als auch die Studie ENDEAR liegen diesbezüglich keine Angaben vor, wie viele Personen eine genetische Modifikation c.859G>C im SMN2-Gen aufwiesen. Nach Angaben des pU sei das Vorliegen dieses genetischen Modifikators mit einem geringeren klinischen Schweregrad assoziiert [7]. Für eine Beurteilung dieses Zusammenhangs konnten keine Publikationen identifiziert werden. In der Studie NeuroNEXT waren Patientinnen und Patienten ausgeschlossen, die mehr als 12 Stunden/Tag eine nicht-invasive ventilatorische Unterstützung benötigten (z. B. BiPAP). Durch die Nicht-Berücksichtigung von Patientinnen und Patienten, die länger als 12 Stunden/Tag eine nicht-invasive ventilatorische Unterstützung benötigten, kann ein Selektionsbias bezogen auf den Einschluss von mildereren Verlaufsformen in der Studie NeuroNEXT im Vergleich zu den anderen externen Kontrollen nicht ausgeschlossen werden. Zudem durften Patientinnen und Patienten in den Studien NeuroNEXT und ENDEAR zu Studienbeginn keine SMA-behandelnden Wirkstoffe, wie z. B. Valproinsäure/Valporat und weitere, erhalten. Für die Studie PNCr wurden keine Ausschlusskriterien definiert.

Das Einschlusskriterium Alter ≤ 6 Monate bei Symptombeginn war ausschließlich für die Studien CL-101, PNCR und ENDEAR definiert, weshalb die entsprechenden Studienpopulationen symptomatische Patientinnen und Patienten umfassen, während in den Studien CL-303 und NeuroNEXT auch asymptomatische Personen berücksichtigt wurden. Vor dem Hintergrund des natürlichen Krankheitsverlaufes und den mit Symptombeginn einhergehenden irreversiblen Schädigungen von Muskelgewebe ist es möglich, dass die in die Studien CL-303 und NeuroNEXT eingeschlossenen asymptomatischen Patientinnen und Patienten zu Studienbeginn mildere Verlaufsformen aufweisen.

Der pU hat für den Vergleich der Studienpopulationen (CL-101 vs. PNCR und NeuroNEXT) Informationen zu soziodemographischen und krankheitsbezogenen Charakteristika in einem Bericht dargelegt (RPT-806 [7]). Angaben für die Studien CL-303 und ENDEAR sind in dem Bericht nicht enthalten. Trotz der weitgehend übereinstimmenden Einschlusskriterien zeigen sich zwischen den Studienpopulationen der Zulassungsbegründenden Studien und den externen Kontrollen (PNCR, NeuroNEXT, ENDEAR) zu Baseline erhebliche Unterschiede. Darüber hinaus fehlen relevante Angaben zum statistischen Lageparameter Median, um die Daten umfassend beurteilen zu können. Während das mittlere Alter zu Studieneinschluss in den Studien CL-101, CL-303, NeuroNEXT und ENDEAR in etwa vergleichbar war (CL-101: 3,5 Monate; CL-303: 3,7 Monate; NeuroNEXT: 4,1 Monate; ENDEAR: 5,5 Monate), lag dieses in der Studie PNCR im Mittel bei 29,0 Monaten; diese Angabe wird bei Betrachtung der Spannweite bestätigt, die zwischen 2 und 171 Monaten liegt. Hierfür wären Informationen zum Median hilfreich, um einschätzen zu können, ob es sich um einen Ausreißer handelt. Auch das Alter beim Auftreten der ersten Symptome unterscheidet sich zwischen den Studienpopulationen: Während bei Kindern in der Studie CL-101 Symptome erstmals zwischen 0 und 3 Monaten und in der Studie CL-303 zwischen 0 und 4 Monaten auftraten, zeigten sich diese in der Studie PNCR erstmals zwischen 0,5 bis 6 Monaten und in der Studie ENDEAR zwischen 0,25 bis 5 Monaten und damit etwas später als in den Zulassungsstudien. Für die Studie NeuroNEXT liegen diesbezüglich keine Angaben vor. Während in Studie CL-303 keine Person und in Studie CL-101 83,3 % bzw. 58,3 % keine Beatmungs- bzw. Ernährungsunterstützung benötigten, betraf dies zum Studieneinschluss in der Studie PNCR weniger als 50 % (Ernährung: 21,7 %, Beatmung: 47,8 %) und in der Studie NeuroNEXT weniger als 65 % (Ernährung: 56,3 %, Beatmung: 62,5 %). Patientinnen und Patienten, die mehr als 12 Stunden/Tag eine nicht-invasive ventilatorische Unterstützung benötigten (z. B. BiPAP), waren in der NeuroNEXT-Studie im Vergleich zur PNCR-Studie ausgeschlossen. Daraus ergibt sich unter Umständen der höhere Anteil beatmeter Personen vor dem 6. Lebensmonat in der PNCR-Kohorte gegenüber der NeuroNEXT-Kohorte (91,3 % vs. 62,5 %). Für die BSC-Kohorte der Studie ENDEAR liegen hinsichtlich der ventilatorischen Unterstützung zum 6. Lebensmonat keine Daten vor (siehe Anhang 1 bis 3).

Wie bereits ausgeführt zeigen sich Hinweise auf mögliche Unterschiede der Populationen bezogen auf die Einschlusskriterien und die soziodemographische und krankheitsbezogene Baselinecharakteristik. Darüber hinaus fehlen für die Studien PNCR und NeuroNEXT Angaben über die tatsächliche Beobachtungsdauer (geplant: 36 Monate PNCR-Studie, 24 Monate NeuroNEXT-Studie). Für die Studie ENDEAR war eine Beobachtungszeit von 13 Monaten geplant, aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie jedoch vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht ca. 6,2 Monaten. Durch die längere Beobachtungsdauer der Studien CL-101 und CL-303 im Vergleich zu ENDEAR ergibt sich eine potentielle Verzerrung, da die unterschiedlichen Beobachtungszeiten in den Analysen nicht berücksichtigt wurden. Gleiches gilt für die Studie PNCR im Vergleich zu den Zulassungsstudien.

Von einem analogen Verzerrungsmechanismus ist beim Vergleich der Studien CL-101 und CL-303 mit PNCR und NeuroNEXT aufgrund der möglicherweise längeren Beobachtungszeiten in den beiden Beobachtungsstudien auszugehen. Der pU adressiert im Nutzendossier (Modul 4), dass der frühere Vergleichszeitpunkt 13–14 Monate gewählt wurde, um einen adäquateren Vergleich mit den Studien zum natürlichen Verlauf zu ermöglichen (13,6 Monate in Studie CL-101, 14 Monate in den Studien PNCR und NeuroNEXT sowie finale Daten in der Studie ENDEAR mit einer maximalen Beobachtungszeit von 13 Monaten). Die zusätzliche Darstellung der Ergebnisse über die gesamte Beobachtungsdauer wird kritisch gesehen, da diese sich zwischen den Studien relevant unterscheiden und Angaben zur tatsächlichen Beobachtungsdauer sowohl in den zulassungsbegründenden Studien als auch in den externen Kontrollen fehlen. Hierfür wären Beobachtungszeit-adjustierte Effektschätzer adäquat, die diese Unterschiede in der Analyse berücksichtigen.

Für den Vergleich mit den externen Kontrollen wurden weder ein Studienprotokoll und ein SAP noch eine detaillierte Ergebnisdarstellung vorgelegt. Die Aufbereitung der Ergebnisse erfolgte primär im Nutzendossier (Modul 4). Dadurch fehlen beispielsweise detaillierte Angaben zu Studiendurchführung (Erhebungsmethodik), Datengrundlage (Vollständigkeit, Vollzähligkeit) und angewendete statistische Methoden. Eine systematische Erfassung von potentiellen Confoundern und deren Berücksichtigung in den Analysen für den Vergleich ist aus den Studienunterlagen nicht ersichtlich. Dies wird kritisch gesehen, da nicht-adjustierte indirekte Vergleiche grundsätzlich eine erhöhte Ergebnisunsicherheit besitzen und in der Regel keine hinreichend aussagekräftige Datengrundlage für eine Nutzenbewertung darstellen. Es ist kritisch anzumerken, dass für den Vergleich der Zulassungsstudien gegenüber den externen Kontrollen z. T. nur aggregierte Auswertungen durch Zusammenfassen (Poolen) der Studienpopulationen PNCR und NeuroNEXT vorliegen. Dies erscheint aufgrund der zuvor beschriebenen Unterschiede der Studienpopulationen nicht adäquat. Der pU sieht die Voraussetzung der Vergleichbarkeit der Studienpopulationen anhand der definierten Einschlusskriterien (siehe Abschnitt 2.2.2) für einen indirekten Vergleich als erfüllt an. Um dies zu bestätigen führt der pU im Bericht RPT-806 eine sogenannte „Sensitivitätsanalyse“ anhand individueller Patientendaten für die Studie CL-101 im Vergleich mit den externen Kontrollen PNCR und NeuroNEXT durch. Alle 12 Personen der Studie CL-101 wurden ausgewählt und für diese die am besten passenden Kontrollen aus den Studien PNCR und NeuroNEXT anhand folgender Baselinekriterien gegenübergestellt: u. a. Alter bei Krankheitsbeginn, Ernährungs- und Beatmungsunterstützung, motorische Funktion anhand des CHOP-INTEND-Wertes. Trotz dieses Matching-Vorgehens bestehen weiterhin Unterschiede zwischen den einzelnen Paaren (CL-101 vs. PNCR, CL-101 vs. NeuroNEXT). Dies zeigt sich u. a. bei der Ernährungs- und Beatmungsunterstützung. Es fehlen Angaben bezogen auf die Definition/Kriterien dieser Baselinevariablen, die für das Matching herangezogen wurden, um die bestehenden Unterschiede im Matching beurteilen zu können. Eine Erklärung hierfür fehlt und wird vom pU auch nicht klinisch/inhaltlich diskutiert. Darüber hinaus scheint für die verbliebenen Patientinnen und Patienten in den Studien PNCR und NeuroNEXT (PNCR: N = 11; NeuroNEXT: N = 4) vom pU eine Vergleichbarkeit angenommen zu werden, da diese für die Auswertungen berücksichtigt wurden. Eine Adjustierung für mögliche Confounder fand, wie oben beschrieben, ebenfalls keine Berücksichtigung. Dabei sollte die Auswahl geeigneter Adjustierungsverfahren und der Störfaktoren prospektiv erfolgen. Bei einem historischen Vergleich liegt keine Randomisierung vor, sodass die Studien hinsichtlich ihres Designs, der Population und weiterer Kriterien sehr homogen sein müssen. Das IQWiG spezifiziert dabei genauer, dass auch bei sehr ähnlichen Studien ein historischer Vergleich nur dann akzeptiert werden kann, wenn ein dramatischer Effekt vorliegt – ein Unterschied von mindestens einem Faktor 5 bis 10; d. h., der relative Effekt, z. B. die Hazard

Ratio, liegt unterhalb von 0,2 bis 0,1 [24]. Zwar liegen aufgrund der genannten methodischen Limitationen große Unsicherheiten bei den nicht-adjustierten indirekten Vergleichen vor, jedoch erreichen die beobachteten Effekte teilweise ein Ausmaß in Größenordnungen, die darauf hinweisen, dass die Effekte nicht allein auf systematischer Verzerrung oder Zufallsbefunden beruhen. Jedoch ist die Ergebnisunsicherheit aufgrund der geringen Präzision der Ergebnisse hierbei einschränkend zu beachten, da sehr breite Konfidenzintervalle vorliegen, die auch Effekte basierend auf systematischen Verzerrungen oder Zufallsbefunden umfassen können. Dennoch werden die Ergebnisse aufgrund der Größenordnung der Punktschätzer der Effekte in der vorliegenden Nutzenbewertung berücksichtigt. Wünschenswert wäre eine separate Ergebnisdarstellung der Studien CL-101 und CL-303 im Vergleich mit den Studien PNCr und NeuroNEXT, um eine bessere Einschätzung der gepoolten Ergebnisse vornehmen zu können.

Für den Nachweis der Wirksamkeit von Onasemnogen-Abepravovec zur Behandlung von 5q13-assoziiertes SMA mit biallelischer Deletion des SMN1-Gens und entweder einer klinischen Diagnose SMA Typ 1 oder bis zu drei Genkopien des SMN2-Gens reicht der pU Daten zu Mortalität, Morbidität und Sicherheit ein.

4.4 Mortalität

In der Endpunktkategorie Mortalität werden die Ergebnisse zu Mortalität und Überleben (Tod oder dauerhafte Beatmung) berichtet.

Für die beiden Endpunkte „Mortalität“ und „Tod oder dauerhafte Beatmung (Überleben)“ wurden die Daten aus den Studien CL-101 und CL-303 (getrennt und gepoolt) gegenüber der gepoolten Population der Studien PNCr und NeuroNEXT verglichen. Eine separate Ergebnisdarstellung (CL-101 vs. PNCr bzw. NeuroNEXT und CL-303 vs. PNCr bzw. NeuroNEXT) liegt nicht vor. Aufgrund der in Abschnitt 4.3 beschriebenen methodischen Limitation wird ein Poolen dieser Populationen kritisch gesehen.

Mortalität

In den Studien CL-101 und CL-303 wurden Todesfälle als Sicherheitsendpunkt erfasst. Darüber hinaus wurden Todesfälle als Teil des kombinierter Endpunkts „Tod oder dauerhafte Beatmung“ in der ITT-Population ausgewertet. Angaben zur medianen Beobachtungsdauer liegen sowohl für die zulassungsbegründenden Studien CL-303 und CL-101 sowie für die externen Kontrollstudien nicht vor. In der Studie ENDEAR betrug die geplante Beobachtungsdauer 13 Monate. Aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht etwa 6,2 Monaten, und ist damit deutlich kürzer als in den Zulassungsstudien.

In der Studie CL-303 verstarb bis zum finalen Datenschnitt am 31. Dezember 2019 (Alter: ≥ 18 Monate) von den 22 eingeschlossenen Person eine (4,5 %) im Alter von 7,8 Monaten. Zum Zeitpunkt des Gentransfers war die Person 2,1 Monate alt.

In der Studie CL-101 verstarb bis zum finalen Datenschnitt am 14. Dezember 2017 (24 Monate nach Gentransfer) keine Person.

In der Studie PNCr wurden die Patientinnen und Patienten über einen Zeitraum von 36 Monaten beobachtet. In dieser Zeit verstarben insgesamt 11 von 23 Personen, von denen 7 Personen (30,4 %) bereits in einem Alter von 14 Monaten verstarben.

In der Studie NeuroNEXT wurden die Patientinnen und Patienten über einen Zeitraum von 24 Monaten beobachtet. In dieser Zeit verstarben insgesamt 8 von 16 Personen, von denen 7 Personen (43,8 %) bereits in einem Alter von 14 Monaten verstarben.

In der Studie ENDEAR betrug die geplante Beobachtungsdauer 13 Monate. Aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht in etwa 6,2 Monaten, und ist damit deutlich kürzer als in den Zulassungsstudien. In dieser Zeit verstarben insgesamt 16 von 41 Personen (39,0 %).

Angaben zur medianen Beobachtungszeit und der medianen Überlebensdauer (jeweils in Monaten) liegen für die Studien CL-101 und CL-303 sowie für die Studien PNCR und NeuroNEXT nicht vor; auch wurden keine Kaplan-Meier-Schätzer mitgeliefert. Das mediane Überleben in der BSC-Kohorte der ENDEAR-Studie wurde nicht erreicht. Das Relative Risiko für den Endpunkt „Mortalität“ im Alter 13–14 Monate für den Vergleich der gepoolten Studienpopulation CL-101+CL-303 im Vergleich zur gepoolten Studienpopulation PNCR+NeuroNEXT liegt bei 12,21 (95%-KI: [1,692; 88,04]) und für den Vergleich der gepoolten Studienpopulation CL-101+CL-303 mit der Studie ENDEAR bei 13,27 (95%-KI: [1,853; 95]; jeweils $p < 0,001$). Ein ähnlicher Effekt bestätigt sich für den Vergleich der Studie CL-101 bzw. CL-303 im Vergleich zur gepoolten Studienpopulation PNCR+NeuroNEXT: 9,4 (95%-KI: [0,6; 147,3]) bzw. 7,9 (95%-KI: [1,1; 56,08]).

Tod oder dauerhafte Beatmung (Überleben)

Der kombinierte Endpunkt war in den Studien als primärer oder ko-primärer Wirksamkeitsendpunkt definiert.

Zum Alter von $\geq 13,6$ Monaten und bis zu einem Alter von mindestens 18 Monaten (geplantes Studienende der Studie CL-303) überlebten 20 von 22 Patientinnen und Patienten ereignisfrei. Eine Person verstarb im Alter von 7,8 Monaten aufgrund von Atemstillstand; für die andere Person wurde eine nicht-invasive Beatmung (BiPAP) für ≥ 16 Stunden über einen Zeitraum von über 14 aufeinanderfolgenden Tagen berichtet.

In der Studie CL-101 überlebten alle 12 Patientinnen und Patienten (Kohorte 2) ereignisfrei bis zum Alter von $\geq 13,6$ Monaten (Zeitpunkt der primären Wirksamkeitsanalyse, Datenschnitt: 20. Januar 2017) und 24 Monate nach Gentransfer (Datenschnitt: 14. Dezember 2017).

Darüber hinaus liegen für den kombinierten Endpunkt „Tod oder dauerhafte Beatmung (Überleben)“ unterschiedliche Operationalisierungen der Teilkomponente „Dauerhafte Beatmung“ vor. Eine mit den Studien CL-303 und CL-101 vergleichbare Operationalisierung ist lediglich für die Studie PNCR gegeben. So wurde in der Studie NeuroNEXT ausschließlich die invasive Beatmung (Intubation) berücksichtigt. Es fehlen Angaben über die Mindestdauer eines Ereignisses (invasive Beatmung) wie in den Zulassungsstudien, zudem wird die nicht-invasive Beatmung nicht erfasst. Zwar werden die Tracheostomie und die Intubation als hinreichend vergleichbar im Sinne einer invasiven Beatmungsunterstützung zu den Zulassungsstudien angesehen, dennoch können die relevanten Unterschiede in der Operationalisierung potentiell zu einer Überschätzung des Effektes zuungunsten von Onasemnogen-Abeparvovec führen, da in den Zulassungsstudien CL-101 und CL-303 zusätzlich die nicht-invasiven Fälle erfasst werden und diese im Vergleich zu der invasiven Beatmung eher mildere Formen der Beatmungssituation darstellen und womöglich häufiger erfasst werden. In der Studie ENDEAR wurde sowohl invasive als auch nicht-invasive Beatmung für mindestens 16 Stunden am Tag als dauerhaft gewertet. Die Operationalisierungen unterscheiden sich jedoch hinsichtlich des Kriteriums Mindestdauer für das Ereignis. In den Studien CL-303,

CL-101 und PNCR ist der Zeitraum mit 14 Tagen kürzer als in der Studie ENDEAR mit 21 Tagen. Auch hier besteht die Möglichkeit einer Verzerrung zuungunsten von Onasemnogen-Abeparvovec, da der Endpunkt in den Studien CL-101 und CL-303 früher erreicht wird als in der Studie ENDEAR. Für die Studie ENDEAR war eine Beobachtungszeit von 13 Monaten geplant, aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie vorzeitig beendet worden; die Studiendauer im BSC-Arm betrug im Median 187 Tage, das entspricht ca. 6,2 Monaten. Durch die längere Beobachtungsdauer in den Studien CL-101 und CL-303 ist das Ergebnis unter Berücksichtigung der ITT-Population potentiell zuungunsten von Onasemnogen-Abeparvovec verzerrt.

Das geschätzte Relative Risiko des Endpunktes „Tod oder dauerhafte Beatmung“ im Alter 13–14 Monate für den Vergleich der gepoolten Studienpopulation CL-303+CL-101 im Vergleich zur gepoolten Studienpopulation PNCR+NeuroNEXT lag bei 10,46 (95%-KI: [2,67; 41,06]) und für den Vergleich der gepoolten Studienpopulation CL-101+CL-303 mit der Studie ENDEAR bei 11,61 (95%-KI: [2,98; 45,26], jeweils $p < 0,001$).

Zusammenfassende Einschätzung zur Mortalität

Von den eingeschlossenen Patientinnen und Patienten verstarben bis zum finalen Datenschnitt 4,5 % in der Studie CL-303 bzw. 0 % in der Studie CL-101. Angaben zur medianen Beobachtungszeit und der medianen Überlebensdauer liegen für die Studien CL-101 und CL-303 nicht vor. Da es sich bei beiden Studien um unkontrollierte Studien handelt, wird von einem hohen Verzerrungspotential auf Studien- und Endpunktebene ausgegangen. Zudem ist bei der Ergebnissicherheit die geringe Stichprobengröße der Studien – 12 Personen in der Studie CL-101, 22 Personen in der Studie CL-303 – als einschränkend zu berücksichtigen.

Für den kombinierten Endpunkt „Tod oder dauerhafte Beatmung“ blieben in der Studie CL-303 bis zu einem Alter von mindestens 18 Monaten 20 von 22 Patientinnen und Patienten ereignisfrei. In der Studie CL-101 überlebten alle 12 Patientinnen und Patienten (Kohorte 2) ereignisfrei bis 24 Monate nach Gentransfer.

Aufgrund des einarmigen Designs der zulassungsbegründenden Studien CL-101 und CL-303 wurden vom pU nicht-adjustierte indirekte Vergleiche (ohne Brückenkomparator) vorgelegt. Zwar liegen aufgrund der methodischen Limitationen große Unsicherheiten für den nicht-adjustierten indirekten Vergleich vor, jedoch erreichen die beobachteten Effekte teilweise ein Ausmaß in Größenordnungen, die darauf hinweisen, dass die Effekte nicht allein auf systematischer Verzerrung oder Zufallsbefunden beruhen. Dennoch ist hierbei die Ergebnisunsicherheit aufgrund der geringen Präzision der Ergebnisse einschränkend zu beachten, da für die Endpunktkategorie Mortalität sehr breite Konfidenzintervalle vorliegen und nicht ausgeschlossen werden kann, dass die Effekte systematische Verzerrungen oder Zufallsbefunde umfassen. Darüber hinaus ist aufgrund fehlender Angaben zur tatsächlichen Beobachtungszeit in den zulassungsbegründenden Studien und externen Kontrollen zu berücksichtigen, dass keine Beobachtungszeit-adjustierten Effektschätzer vorliegen, weshalb sich die Ergebnisdarstellung auf die vom pU festgelegte vergleichbare Beobachtungsdauer von 13–14 Monaten bezieht. Es ist darauf hinzuweisen, dass die vorliegenden dramatischen Effekte ausschließlich bis zu diesem Zeitpunkt beurteilt werden können; Langzeitdaten liegen nicht vor. Die adressierten methodischen Limitationen des nicht-adjustierten indirekten Vergleichs sind zu berücksichtigen (siehe Abschnitt 4.3.)

4.5 Morbidität

In der Endpunktkategorie Morbidität werden Ergebnisse zu „Dauerhafte Beatmung“, „Erreichen motorischer Meilensteine“ sowie die Veränderungen über die Zeit für die „Bayley-Skala“ und „CHOP INTEND“ herangezogen und berichtet.

Dauerhafte Beatmung

In der Studie CL-303 erfüllte bis zum finalen Datenschnitt am 31. Dezember 2019 (Alter ≥ 18 Monate) eine Person (4,5 %) im Verlauf der Studie das Kriterium für eine dauerhafte Beatmung. In der Studie CL-101 wurde keine Person bis zum finalen Datenschnitt am 14. Dezember 2017 (24 Monate nach Gentransfer) per Definition dauerhaft beatmet.

Wie bereits unter Abschnitt 4.4 angemerkt, differierte die Operationalisierung der dauerhaften Beatmung in den externen Kontrollen, ausgenommen in der Studie PNCR. In der Studie NeuroNEXT wurde ausschließlich die invasive Beatmung (Intubation) berücksichtigt. Es fehlen Angaben über die Mindestdauer eines Ereignisses (invasive Beatmung) wie in den Zulassungsstudien. Die Tracheostomie (CL-303 und CL-101) und die Intubation (NeuroNEXT) werden als hinreichend vergleichbar im Sinne einer invasiven Beatmungsunterstützung angesehen. In der Studie ENDEAR wurde sowohl die invasive als auch nicht-invasive Beatmung für mindestens 16 Stunden am Tag als dauerhaft gewertet. Die Operationalisierungen der Studien unterscheiden sich jedoch hinsichtlich des Kriteriums „Mindestdauer für das Ereignis“. In den Studien CL-303 und CL-101 ist der Zeitraum mit 14 Tagen kürzer als in der Studie ENDEAR mit 21 Tagen.

In einem Alter von 13–14 Monaten benötigten in der Studie PNCR 13 Patientinnen und Patienten (56,5 %) eine dauerhafte Beatmung. In der Studie CL-101 benötigte keine Person eine dauerhafte Beatmung zum vergleichbaren Beobachtungszeitpunkt und in der Studie CL-303 eine Person (4,5 %). Für den Vergleich der gepoolten Studienpopulation CL-303+CL-101 mit der Studie PNCR zeigte sich ein RR von 19,22 (95%-KI: [2,696; 137], $p < 0,001$). Dieses Ergebnis ist konsistent mit der separaten Darstellung der Zulassungsstudien im Vergleich zu PNCR.

In der Studie NeuroNEXT ist die dauerhafte Beatmung auf Personen mit langfristiger invasiver Beatmung (Intubation) beschränkt, weshalb für den Vergleich mit den Studien CL-101 und CL-303 nur Personen berücksichtigt wurden, für die ebenfalls eine Intubation bis zu einem Alter von 13–14 Monaten berichtet wurde. Das traf auf keine Person in den Zulassungsstudien zu. Das Ergebnis im Alter 13–14 Monate für den Vergleich der gepoolten Studienpopulation CL-303+CL-101 mit der Studienpopulation NeuroNEXT zeigte einen nicht-signifikanten Effekt (RR: 10,29; 95%-KI: [0,5226; 202,8], $p = 0,0541$).

In der Studie ENDEAR unterschied sich die Endpunktoptimalisierung mit denen der Zulassungsstudien hinsichtlich des Kriteriums „Mindestdauer für das Ereignis“, die deutlich länger war. Zudem lag die mediane Beobachtungsdauer in der ENDEAR-Studie bei nur 6,2 Monaten. Für den Vergleich der gepoolten Studienpopulation CL-101+CL-303 mit der Studie ENDEAR lag das geschätzte RR im Alter 13–14 Monate bei 10,78 (95%-KI: [1,485; 78,28], $p = 0,001$) zugunsten von Onasemnogen-Abeparvovec. Beobachtungszeit-adjustierte Effektschätzer wurden nicht erreicht.

Erreichen motorischer Meilensteine

In der Studie CL-303 hatten bis zum finalen Datenschnitt am 31. Dezember 2019 (Alter ≥ 18 Monate) mehr als 50 % der Personen folgende Meilensteine erreicht: „Kopfkontrolle ohne Unterstützung“ (85,0%¹¹), „Freies Sitzen ≥ 10 Sekunden“ (63,6 %), „Rollen“ und „Freies Sitzen für ≥ 30 Sekunden“ (jeweils 59,1 %). Das mediane Alter (min; max) beim Erreichen der motorischen Meilensteine für „Freies Sitzen für ≥ 30 Sekunden“ lag bei 12,6 Monaten (9,2; 18,6), für „Freies Sitzen für ≥ 10 Sekunden“ bei 13,9 Monaten (10,2; 18,6), für „Rollen“ bei 11,5 Monaten (2,9; 18,3) und für „Kopfkontrolle ohne Unterstützung“ bei 6,75 Monaten (2,0; 16,4).

In der Studie CL-101 hatten bis zum finalen Datenschnitt am 14. Dezember 2017 (24 Monate nach Gentransfer) mehr als 50 % der Personen folgende Meilensteine erreicht: „Kopfkontrolle ohne Unterstützung“, „Sitzen mit Unterstützung“ und „Freies Sitzen ≥ 5 Sekunden“ (jeweils 91,7 %), „Freies Sitzen für ≥ 10 Sekunden“ (83,3 %), „Rollen“ und „Freies Sitzen für ≥ 30 Sekunden“ und „Freies Sitzen für ≥ 15 Sekunden“ (jeweils 75,0 %). Angaben zum medianen Alter beim Erreichen der motorischen Meilensteine liegen für die Studie nicht vor.

Bayley-Skala (Grob- und Feinmotorik)

Für die Studien CL-303 und CL-101 wurden die Veränderungen der grob- und feinmotorischen Subskalen der Bayley-Skala über die Zeit berichtet. Die Rücklaufquoten lagen in der Studie CL-303 bis Monat 13 bei annähernd 70 %. Für beide Subskalen zeigte sich im Verlauf ein Anstieg des Wertes im Vergleich zum Ausgangswert. Die klinische Relevanz der Veränderung kann nicht eingeschätzt werden. In der Studie CL-101 wurden nur diejenigen Patientinnen und Patienten mit der Bayley-Skala bewertet, die einen CHOP-INTEND-Wert von ≥ 60 Punkten erreichten. Insgesamt erreichten vier Patientinnen und Patienten einen CHOP-INTEND-Wert ≥ 60 Punkte und wurden mit der Bayley-Skala bewertet. Eine Auswertung für die gesamte Studienpopulation war nicht geplant und liegt nicht vor.

CHOP INTEND

Berichtet wurden die Veränderungen des CHOP-INTEND-Gesamtwertes über die Zeit in den Studien CL-303 und CL-101. Die Rücklaufquoten lagen in beiden Studien bis Monat 12 bei über 70 %. Für beide Studien zeigte sich im Verlauf ein Anstieg des Wertes im Vergleich zum Ausgangswert. Die klinische Relevanz der Veränderung kann nicht eingeschätzt werden.

In der Studie CL-303 erreichten innerhalb der Studiendauer von 18 Monaten 95,5 % der Patientinnen und Patienten einen CHOP-INTEND-Gesamtwert ≥ 40 Punkte, 63,6 % einen Gesamtwert von ≥ 50 Punkten und 22,7 % einen Gesamtwert von ≥ 60 Punkten. In der Studie CL-101 waren es im Verlauf der Studie von 24 Monaten nach Gentransfer 91,7 % mit Gesamtwerten von ≥ 40 und ≥ 50 Punkten sowie 33,3 % ≥ 60 Punkten.

¹¹ Zwei Personen hatten den Meilenstein bereits zur Screening-Visite erreicht.

Zusammenfassende Einschätzung zur Morbidität

Da es sich bei den Studien CL-303 und CL-101 um unkontrollierte Studien handelt wird von einem hohen Verzerrungspotential auf Studien- und Endpunktebene ausgegangen. Zudem sind bei der Ergebnissicherheit die kleinen Studienpopulationen – 12 Personen in der Studie CL-101, 22 Personen in der Studie CL-303 – als einschränkend zu berücksichtigen. Vor dem Hintergrund des fehlenden validen Vergleichs und des limitierten Stichprobenumfangs ist eine abschließende Bewertung zur Morbidität nicht abschließend möglich. Darüber hinaus kann die klinische Relevanz der Veränderung für die grob- und feinmotorischen Subskalen der Bayley-Skala und des CHOP INTEND nicht beurteilt werden.

Aufgrund des einarmigen Designs der zulassungsbegründenden Studien CL-101 und CL-303 wurden vom pU nicht-adjustierte indirekte Vergleiche (ohne Brückenkomparator) für u. a. den Endpunkt „Dauerhafte Beatmung“ vorgelegt. Zwar liegen aufgrund der methodischen Limitationen große Unsicherheiten für den nicht-adjustierten indirekten Vergleich vor, jedoch erreichen die beobachteten Effekte teilweise ein Ausmaß in Größenordnungen, die darauf hinweisen, dass die Effekte nicht allein auf systematischer Verzerrung oder Zufallsbefunden beruhen. Dennoch ist hierbei die Ergebnisunsicherheit aufgrund der geringen Präzision der Ergebnisse einschränkend zu beachten, da hier sehr breite Konfidenzintervalle vorliegen und nicht ausgeschlossen werden kann, dass die Effekte systematische Verzerrungen oder Zufallsbefunde umfassen. Darüber hinaus ist aufgrund fehlender Angaben zur tatsächlichen Beobachtungszeit in den zulassungsbegründenden Studien und externen Kontrollen zu berücksichtigen, dass keine Beobachtungszeit-adjustierten Effektschätzer vorliegen, weshalb sich die Ergebnisdarstellung auf die vom pU festgelegte vergleichbare Beobachtungsdauer 13–14 Monate bezieht. Es ist darauf hinzuweisen, dass die vorliegenden dramatischen Effekte ausschließlich bis zu diesem Zeitpunkt beurteilt werden können; Langzeitdaten liegen nicht vor. Die adressierten methodischen Limitationen des nicht-adjustierten indirekten Vergleichs sind zu berücksichtigen (siehe Abschnitt 4.3.)

4.6 Lebensqualität

Es wurden keine Daten zur gesundheitsbezogenen Lebensqualität erhoben.

4.7 Sicherheit

Die Ergebnisse der Studien CL-101 und CL-303 beziehen sich auf die Sicherheitspopulation, in der jede Person eine Gentherapie-Infusion erhalten hat. Für die Studie CL-101 wird sich auf den finalen Datenschnitt am 14. Dezember 2017 (≥ 24 Monate nach Gentransfer) und für die Studie CL-303 auf den finalen Datenschnitt am 12. Dezember 2019 (Studienende: Alter ≥ 18 Monate) bezogen. Es lässt sich nicht einschätzen, ob die aufgetretenen UE mit der Symptomatik der Grunderkrankung einhergehen oder als Nebenwirkungen der Behandlung auftraten (z. B. Atemnot).

In der Studie CL-303 wurden SUE mit der Unterzeichnung der Einverständniserklärung und UE mit Beginn der Gensatztherapie und bis zur letzten Studienvisite erfasst. In der Studie CL-101 wurden SUE und UE mit Studienbeginn oder Verabreichung der Studienmedikation bis zu 30 Tage nach der letzten Studienvisite erfasst. Unklar ist, warum zumindest SUE in der Studie CL-303 analog der Studie CL-101 nicht auch bis zu 30 Tage nach der letzten Studienvisite erfasst wurden. Die Erfassungszeitraum von AESI wurde in den Studien nicht angegeben.

In beiden Zulassungsstudien CL-101 und CL-303 traten UE jeglichen Grades bei allen Personen (100 %) auf. Die Häufigkeit von UE nach CTCAE-Grad ≥ 3 und SUE trat in beiden Studien bei jeweils 10 Personen auf (CL-101: 83,3 %; CL-303: 45,5 %). Ein UE, das zum Therapieabbruch führte, trat ausschließlich in Studie CL-303 auf (2/22 Personen (9,1 %)); Gründe waren jeweils Atemnot und Atemstillstand.

Unerwünschte Ereignisse nach Preferred Term (> 10 %) sowie UE CTCAE-Grad ≥ 3 (5 %)

In der Studie CL-101 waren die häufigsten Preferred Terms: Infektion der oberen Atemwege (83,3 %), Erbrechen (66,7 %), Pyrexie (58,3 %) und Pneumonie (58,3 %) sowie Obstipation (50 %) und verstopfte Nase (50 %). In der Studie CL-303 traten am häufigsten (mehr als 50 %) Erkrankungen folgender Preferred Terms auf: Pyrexie (54,5 %) und Infektion der oberen Atemwege (50 %).

Bei den UE CTCAE-Grad ≥ 3 nach Preferred Terms waren es in Studie CL-101 am häufigsten: Atelektase (25 %) und Pneumonie (58,3 %). In der Studie CL-303 traten am häufigsten (mehr als 10 %) Erkrankungen folgender Preferred Terms auf: RSV-Bronchiolitis (13,6 %) und Atemnot (13,6 %).

Schwerwiegende unerwünschte Ereignisse (5 %)

In der Studie CL-101 traten am häufigsten (mehr als 15 %) Erkrankungen folgender Preferred Terms auf: Pneumonie (58,3 %) und Infektion der oberen Atemwege (25 %). In der Studie CL-303 trat am häufigsten (mehr als 10 %) folgender Preferred Term auf: Atemnot (18,2 %).

Unerwünschte Ereignisse von besonderem Interesse (AESI)

In der Studie CL-101 wurden erhöhte Leberenzymwerte als AESI präspezifiziert. Bei 3 Personen (25 %) traten erhöhte Leberenzymwerte auf (Erhöhung der Transaminasen (N = 3), Erhöhung der Aspartat-Aminotransferase (N = 1)). Die beobachteten Erhöhungen der Leberwerte waren klinisch asymptomatisch, d. h. weder Gelbsucht noch Symptome einer eingeschränkten Leberfunktion wurden berichtet. Bei keiner Person war Hy's Law erfüllt. Alle Ereignisse waren im Verlauf der Studie reversibel. In der Studie CL-303 wurden „Hepatotoxizität“, „Thrombozytopenie“, „Kardiale Ereignisse“ sowie „Auf Ganglionopathie hindeutende sensorische Anomalien“ als AESI präspezifiziert. Insgesamt wurden bei 13 von 22 Personen (59,1 %) AESI berichtet. Darunter waren 7 Personen (31,8 %) mit mindestens einem Ereignis innerhalb der Kategorie „Hepatotoxizität“, 8 (36,4 %) mit mindestens einem Ereignis innerhalb der Kategorie „Thrombozytopenie“, 4 (18,2 %) mit mindestens einem Ereignis innerhalb der Kategorie „Kardiale Ereignisse“ und 5 (22,7 %) mit mindestens einem Ereignis innerhalb der Kategorie „Auf Ganglionopathie hindeutende sensorische Anomalien“.

Hospitalisierung

In den Zulassungsstudien wurde die Hospitalisierung unter den Sicherheitsendpunkten erfasst. Bis zum finalen Datenschnitt am 12. Dezember 2019 (Alter ≥ 18 Monate) mussten insgesamt 11 von 22 Patientinnen und Patienten (50 %) in der Studie CL-303 ins Krankenhaus. Die mittlere annualisierte Hospitalisierungsrate (SE) lag bei 0,8 (0,2). Angaben über u. a. die Anzahl an Krankenhaustagen sowie Verweildauer liegen für die Studie nicht vor.

Bis zum finalen Datenschnitt am 14. Dezember 2017 (24 Monate nach Gentransfer) wurden insgesamt 10 von 12 Patientinnen und Patienten (83,3 %) in der Studie CL-101 ins Krankenhaus aufgenommen. Insgesamt wurden 51 stationäre Aufnahmen berichtet. Jede/r Patient/in verbrachte im Median (min; max) 5,8 Tage (3,0; 12,1) im Krankenhaus. Die ermittelte mittlere annualisierte Hospitalisierungsrate (SE) lag bei bei 2,1 (0,6).

Zusammenfassende Einschätzung zur Sicherheit

Da es sich bei den Studien CL-303 und CL-101 um unkontrollierte Studien handelt, wird von einem hohen Verzerrungspotential auf Studien- und Endpunktebene ausgegangen. Zudem sind bei der Ergebnissicherheit die kleinen Studienpopulationen – 12 Personen in der Studie CL-101, 22 Personen in der Studie CL-303 – als einschränkend zu berücksichtigen. Vor dem Hintergrund des fehlenden validen Vergleichs und des limitierten Stichprobenumfangs ist eine abschließende Bewertung zur Sicherheit, insbesondere im Hinblick auf Aussagen zur Langzeitsicherheit, nicht möglich.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Vorgaben der Fachinformation sind zu berücksichtigen. Die Behandlung mit Zolgensma darf nur von Ärztinnen und Ärzten eingeleitet und überwacht werden, die über Erfahrung in der Therapie von Patienten mit spinaler Muskelatrophie verfügen.

6 Zusammenfassung der Nutzenbewertung

Onasemnogen-Abeparvovec (Zolgensma®) ist zugelassen bei Patientinnen und Patienten mit 5q13-assoziiertes spinaler Muskelatrophie (SMA) mit biallelischer Deletion des SMN1-Gens und entweder einer klinischen Diagnose SMA Typ 1 oder bis zu drei Kopien des SMN2-Gens. Die Nutzenbewertung von Onasemnogen-Abeparvovec basiert auf den zulassungsbegründenden Studien CL-101 und CL-303. Studie CL-303 (STRIVE-US) ist eine einarmige, offene, multizentrische Phase-III-Studie. Das Ziel der Studie bestand darin, die Wirksamkeit und Sicherheit von Onasemnogen-Abeparvovec als einmalige intravenöse Injektion zu untersuchen. Eingeschlossen waren 22 Patientinnen und Patienten mit biallelischer Mutation des SMN1-Gens (Deletion oder Punktmutation), ein oder zwei SMN2-Genkopien im Alter < 6 Monate (< 180 Tage) zum Zeitpunkt des Gentransfers. Studie CL-101 (START) ist eine einarmige, offene, monozentrische Phase-I-Studie. Das primäre Ziel der Studie war die Untersuchung der Sicherheit verschiedener Dosierungen von Onasemnogen-Abeparvovec als einmalige intravenöse Injektion. Eingeschlossen waren Patientinnen und Patienten im Alter ≤ 6 Monate zum Zeitpunkt des Gentransfers und infantiler SMA Typ 1 mit biallelischer Mutation des SMN1-Gens (Deletion oder Punktmutation) und ein oder zwei SMN2-Genkopien sowie Auftreten der ersten Symptome in einem Alter zwischen 0 und 6 Monaten. Von Relevanz sind in den zulassungsbegründenden Studien die zulassungskonforme Applikation gemäß Fachinformation [9].

Das Verzerrungspotential beider Studien wird auf Basis des Studiendesigns mit hoch bewertet. Die Ergebnisse der einarmigen Studien werden in der folgenden Tabelle zusammengefasst.

Tabelle 40: Zusammenfassende Darstellung der Ergebnisse für die Studien CL-101 und CL-303 (finaler Datenschnitt)

Endpunkt	Studie CL-303 ¹⁾ N = 22	Studie CL-101 ²⁾ N = 12
Mortalität		
Mortalität ³⁾ , n (%)	1 (4,5)	0
Überleben (Tod oder dauerhafte Beatmung) ³⁾ , n (%)	2 (9,1)	0
Morbidität		
Dauerhafte Beatmung ³⁾ , n (%)	1 (4,5)	0
<i>Erreichen motorischer Meilensteine⁴⁾, n (%)</i>		
Kopfkontrolle ohne Unterstützung	17 (85,0) ⁵⁾	11 (91,7)
Rollen	13 (59,1)	9 (75)
Sitzen mit Unterstützung	k. A.	11 (91,7)
Freies Sitzen für ≥ 30 Sekunden	13 (59,1) ⁶⁾	9 (75,0)
Freies Sitzen für ≥ 10 Sekunden (WHO MGRS)	14 (63,6)	10 (83,3)
Hält das eigene Gewicht für mindestens 2 Sekunden	1 (4,5)	n. e.
Krabbeln	1 (4,5)	n. e.
Hochziehen in eine stehende Position	1 (4,5)	n. e.
Stehen mit Unterstützung	1 (4,5)	2 (16,7) ⁹⁾
Gehen mit Unterstützung	1 (4,5)	2 (16,7) ⁹⁾
Stehen ohne Unterstützung	1 (4,5)	2 (16,7) ⁹⁾
Gehen ohne Unterstützung	1 (4,5)	2 (16,7) ⁹⁾

Endpunkt	Studie CL-303 ¹⁾ N = 22	Studie CL-101 ²⁾ N = 12
<i>Veränderung der Bayley-Skala zu Monat 13 im Vergleich zu Baseline</i>		
Grobmotorik ¹⁰⁾ , MW (SD)	13,5 (6,67) ¹²⁾	_ ¹³⁾
Feinmotorik ¹¹⁾ , MW (SD)	22,9 (3,84) ¹²⁾	_ ¹³⁾
<i>Veränderung des CHOP-INTEND-Gesamtwertes zu Monat 12 im Vergleich zu Baseline¹⁴⁾¹⁵⁾</i>		
MW (SD)	16,4 (8,04)	n. e.
LS MW (SE) ¹⁶⁾	n. e.	26,6 (2,0)
Sicherheit		
UE, n (%)	22 (100)	12 (100)
UE CTCAE-Grad ≥ 3, n (%)	10 (45,5)	10 (83,3)
Therapieabbruch aufgrund von UE, n (%)	2 (9,1)	0
SUE, n (%)	10 (45,5)	10 (83,3)
AESI		
Hepatotoxizität, n (%)	7 (31,8)	3 (25)
Thrombozytopenie, n (%)	8 (36,4)	n. e.
Kardiale Ereignisse, n (%)	4 (18,2)	n. e.
Ganglionopathie, n (%)	5 (22,7)	n. e.
Annualisierte Hospitalisierungsrate ⁹⁾ , MW (SE)	0,8 (0,2)	2,1 (0,6)

¹⁾ Studienende bis zu einem Alter von mindestens 18 Monaten (finaler Datenschnitt: 12.12.2019). Angaben zur medianen Beobachtungsdauer liegen nicht vor.

²⁾ Studienende 24 Monate nach Gentransfer (finaler Datenschnitt: 14.12.2017). Angaben zur medianen Beobachtungsdauer liegen nicht vor.

³⁾ Auswertung bezieht sich auf die ITT-Population.

⁴⁾ CL-303: ITT-Population; CL-101: FAS-Population.

⁵⁾ Zwei Personen hatten den Meilenstein bereits zur Screening-Visite, d. h. vor Behandlung mit Onasemnogen-Abeparvovec erreicht; in die Auswertung gingen demnach nur 20 Personen ein. Es liegen inkonsistente Angaben in den Studienunterlagen vor: 18/20 vs. 17/20 Personen, die den Kopf für mindestens 3 Sekunden ohne Unterstützung halten können.

⁶⁾ Eine Person hat den Meilenstein „Freies Sitzen für 30 Sekunden“ im Alter von 16 Monaten erreicht, der Meilenstein wurde in der Studiervisite in einem Alter von 18 Monaten nicht bestätigt.

⁷⁾ Wurde im Studienbericht als „Freies Sitzen < 10 Sekunden“ bezeichnet. Die externe Begutachtung bestätigte, dass dieser Meilenstein als "Freies Sitzen ≥ 5 Sekunden" definiert und erfasst wurde.

⁸⁾ Nicht präspezifiziert.

⁹⁾ Nicht-adjustierte Hospitalisierungsrate.

¹⁰⁾ Die Grobmotorik-Subskala der Bayley-Skala besteht aus 72 Items.

¹¹⁾ Die Feinmotorik-Subskala der Bayley-Skala besteht aus 66 Items.

¹²⁾ Die Rücklaufquote lag in Studie CL-303 bis Monat 13 bei annähernd 70 % (N = 15/22).

¹³⁾ In Studie CL-101 wurden nur jene Personen mit der Bayley-Skala bewertet, die einen CHOP-INTEND-Wert von ≥ 60 Punkten erreichten. Insgesamt erreichten 4 Personen (25 %) einen CHOP-INTEND-Wert ≥ 60 Punkte und wurden mit der Bayley-Skala bewertet. Rücklaufquote unter 70 %.

¹⁴⁾ Der CHOP INTEND besteht aus 16 Items, die mit 0 (ohne Funktion), 1, 2, 3 oder 4 (volle Funktion) bewertet werden. Es kann maximal ein Wert von 64 Punkten erreicht werden; höhere Werte weisen auf eine bessere motorische Funktion hin.

¹⁵⁾ Rücklaufquoten lagen in beiden Studien bis Monat 12 bei über 70 % (CL-303: N = 16/22; CL-101: N = 9/12).

¹⁶⁾ Auswertung basierend auf MMRM: Das Modell beinhaltete die Veränderung gegenüber dem Baseline-Wert als abhängige Variable sowie fixe Effekte (Kohorte, Visite), Kovariate (Baseline-Wert) und Interaktion (Kohorten-Visiten, Baseline-Visiten).

¹⁷⁾ Abkürzungen: AESI: Unerwünschte Ereignisse von besonderem Interesse; CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CTCAE: Common Terminology Criteria for Adverse Events; FAS: Full Analysis Set; ITT: Intention-to-Treat; k. A.: keine Angabe; LS: Least Square; MMRM: Mixed Model of Repeated Measures; MW: Mittelwert; n. e.: nicht ermittelt; SD: Standardabweichung; SE: Standardfehler; (S)UE: (Schwerwiegende/s) Unerwünschte/s Ereignis/se; WHO MGRS: World Health Organization Multicenter Growth Reference Study.

Referenzen

1. **Albers CA, Grieve AJ.** Test review: Bayley, N. (2006). Bayley Scales of Infant and Toddler Development - third edition. San Antonio, TX: Harcourt Assessment. Journal of Psychoeducational Assessment 2016;25(2):180-190.
2. **AveXis.** European, Phase 3, Open-Label, Single-Arm, Single-Dose Gene Replacement Therapy Clinical Trial for Patients With Spinal Muscular Atrophy Type 1 With One or Two SMN2 Copies Delivering AVXS-101 by Intravenous Infusion; study protocol CL-302 [unveröffentlicht]. 2019.
3. **AveXis.** A Global Study of a Single, One-Time Dose of AVXS-101 Delivered to Infants With Genetically Diagnosed and Pre-symptomatic Spinal Muscular Atrophy With Multiple Copies of SMN2; study protocol CL-304 [unveröffentlicht]. 2019.
4. **AveXis.** A Long Term Follow up Safety Study of Patients in the AVXS-101-CL-101 Gene Replacement Therapy Clinical Trial for Spinal Muscular Atrophy Type 1 Delivering AVXS 101; study protocol LT-001 [unveröffentlicht]. 2018.
5. **AveXis.** Phase 3, open-label, single-arm, single-dose gene replacement therapy clinical trial for patients with spinal muscular atrophy type 1 with one or two SMN2 copies delivering AVXS-101 by intravenous infusion; study report CL-303 [unveröffentlicht]. 2020.
6. **AveXis.** Phase I gene transfer clinical trial for spinal muscular atrophy type 1 delivering AVXS-101; study report CL-101 [unveröffentlicht]. 2018.
7. **AveXis.** PNCr and neuronext database report; final report (RPT-806) [unveröffentlicht]. 01.06.2018.
8. **AveXis EU.** Dossier zur Nutzenbewertung gemäß § 35a SGB V: Onasemnogen-Abeparvovec (Zolgensma); Modul 4 A: 5q-assozierte Spinale Muskelatrophie mit biallelischer Mutation von SMN1 und klinischer Diagnose SMA Typ 1 oder bis zu 3 SMN2-Kopien; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 15.05.2020.
9. **AveXis EU.** Zolgensma 2 x 10¹³ Vektorgenom/ml Infusionslösung: Fachinformation [online]. 05.2020. [Zugriff: 08.07.2020]. URL: <http://www.fachinfo.de>.
10. **Bach JR, Baird JS, Plosky D, Navado J, Weaver B.** Spinal muscular atrophy type 1: management and outcomes. Pediatr Pulmonol 2002;34(1):16-22.
11. **Bayley N.** Bayley scales of infant and toddler development; 3. ed. San Antonio: Harcourt Assessment; 2006.
12. **Biogen.** Dossier zur Nutzenbewertung gemäß § 35a SGB V; Nusinersen (Spinraza®); Modul 4 [online]. Berlin (GER): Gemeinsamen Bundesausschusses (G-BA); 2017. [Zugriff: 09.09.2020]. URL: https://www.g-ba.de/downloads/92-975-1991/2017-06-30_Modul4A_Nusinersen.pdf.
13. **De Sanctis R, Pane M, Coratti G, Palermo C, Leone D, Pera MC, et al.** Clinical phenotypes and trajectories of disease progression in type 1 spinal muscular atrophy. Neuromuscul Disord 2018;28(1):24-28.

14. **Deutsches Institut für Medizinische Dokumentation und Information (DIMDI).** Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme: 10. Revision; Version 2019. Mit Aktualisierungen vom 01.11.2019, 13.02.2020, 23.03.2020 und 22.04.2020 [online]. Köln (GER): Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM); 2019. [Zugriff: 03.09.2020]. URL: <https://www.dimdi.de/static/de/klassifikationen/icd/icd-10-who/kode-suche/htmlamtl2019/>.
15. **European Medicines Agency (EMA).** Zolgensma (onasemnogene abeparvovec): EPAR Produktinformation EMEA/H/C/004750/0000 [online]. Amsterdam (NED): EMA; 2020. [Zugriff: 08.07.2020]. URL: https://www.ema.europa.eu/en/documents/product-information/zolgensma-epar-product-information_de.pdf.
16. **European Medicines Agency (EMA).** Zolgensma (onasemnogene abeparvovec): European public assessment report EMEA/H/C/004750/0000 [online]. Amsterdam (NED): EMA; 26.03.2020. [Zugriff: 08.07.2020]. URL: https://www.ema.europa.eu/en/documents/assessment-report/zolgensma-epar-public-assessment-report_en.pdf.
17. **Finkel R, Bertini E, Muntoni F, Mercuri E, Enmc Sma Workshop Study Group.** 209th ENMC International Workshop: Outcome Measures and Clinical Trial Readiness in Spinal Muscular Atrophy 7-9 November 2014, Heemskerk, The Netherlands. *Neuromuscul Disord* 2015;25(7):593-602.
18. **Finkel RS, McDermott MP, Kaufmann P, Darras BT, Chung WK, Sproule DM, et al.** Observational study of spinal muscular atrophy type I and implications for clinical trials. *Neurology* 2014;83(9):810-817.
19. **Finkel RS, Mercuri E, Darras BT, Connolly AM, Kuntz NL, Kirschner J, et al.** Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *N Engl J Med* 2017;377(18):1723-1732.
20. **Finkel RS, Weiner DJ, Mayer OH, McDonough JM, Panitch HB.** Respiratory muscle function in infants with spinal muscular atrophy type I. *Pediatr Pulmonol* 2014;49(12):1234-1242.
21. **Gemeinsamen Bundesausschusses (G-BA).** Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB VBewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 10 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 Verfo Wirkstoff: Nusinersen; vom 2. Oktober 2017 [online]. Berlin (GER): G-BA; 2017. [Zugriff: 09.09.2020]. URL: https://www.g-ba.de/downloads/92-975-1992/2017-07-01_Nutzenbewertung-G-BA_Nusinersen_D-294.pdf.
22. **Glanzman AM, Mazzone E, Main M, Pelliccioni M, Wood J, Swoboda KJ, et al.** The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND): test development and reliability. *Neuromuscul Disord* 2010;20(3):155-161.
23. **Glanzman AM, McDermott MP, Montes J, Martens WB, Flickinger J, Riley S, et al.** Validation of the Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND). *Pediatr Phys Ther* 2011;23(4):322-326.

24. **Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG).** Allgemeine Methoden; Vers. 5.0 [online]. 10.07.2017. Köln (GER): IQWiG; 2017. [Zugriff: 14.09.2020]. URL: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf.
25. **Kolb SJ, Coffey CS, Yankey JW, Krosschell K, Arnold WD, Rutkove SB, et al.** Baseline results of the NeuroNEXT spinal muscular atrophy infant biomarker study. *Ann Clin Transl Neurol* 2016;3(2):132-145.
26. **Kolb SJ, Coffey CS, Yankey JW, Krosschell K, Arnold WD, Rutkove SB, et al.** Natural history of infantile-onset spinal muscular atrophy. *Ann Neurol* 2017;82(6):883-891.
27. **Liao PJ, Campbell SK.** Examination of the item structure of the Alberta infant motor scale. *Pediatr Phys Ther* 2004;16(1):31-38.
28. **Piper MC, Pinnell LE, Darrah J, Maguire T, Byrne PJ.** Construction and validation of the Alberta Infant Motor Scale (AIMS). *Can J Public Health* 1992;83 Suppl 2:S46-50.
29. **Wang CH, Finkel RS, Bertini ES, Schroth M, Simonds A, Wong B, et al.** Consensus statement for standard of care in spinal muscular atrophy. *J Child Neurol* 2007;22(8):1027-1049.
30. **Wijnhoven TM, de Onis M, Onyango AW, Wang T, Bjoerneboe GE, Bhandari N, et al.** Assessment of gross motor development in the WHO Multicentre Growth Reference Study. *Food Nutr Bull* 2004;25(1 Suppl):S37-45.
31. **World Health Organization Multicentre Growth Reference Study Group.** Reliability of motor development data in the WHO Multicentre Growth Reference Study. *Acta Paediatr Suppl* 2006;450:47-55.
32. **World Health Organization Multicentre Growth Reference Study Group.** WHO Motor Development Study: windows of achievement for six gross motor development milestones. *Acta Paediatr Suppl* 2006;450:86-95.

Anhang 1

Tabelle 41: Charakterisierung der Population mit SMA Typ 1 der PNCR-Kohorte [18]

Studie PNCR Merkmal	SMA Typ 1B ¹⁾		SMA Typ 1C ²⁾	
	Recent ³⁾ N = 8	Chronic ⁴⁾ N = 10	Recent ³⁾ N = 6	Chronic ⁴⁾ N = 10
Geschlecht, n (%)				
männlich	2 (25)	4 (40)	5 (83)	8 (80)
weiblich	6 (75)	6 (60)	1 (17)	2 (20)
Alter bei Symptombeginn (Monate)				
MW (SD)	k. A.	k. A.	k. A.	k. A.
Median (IQA)	1,5 (1; 2,5)	3,0 (2; 3)	4,0 (4; 5)	4,0 (4; 6)
Alter bei klinischer Diagnose (Monate)				
MW (SD)	k. A.	k. A.	k. A.	k. A.
Median (IQA)	3,5 (1,8; 6)	6,0 (4; 7)	5,5 (5; 7)	6,0 (6; 8)
Alter zum Studieneinschluss (Monate)				
MW (SD)	k. A.	k. A.	k. A.	k. A.
Median (IQA)	5,0 (2,5; 6,5)	59 (17; 184)	6,5 (6; 8)	30,5 (22; 78)
SMN2-Genkopien, n (%)				
2	7 (88)	6 (60)	3 (50)	7 (70)
3	1 (12)	3 (30)	3 (50)	2 (20)
Keine Angabe	0	1 (10)	0	1 (10)
Unterstützung bei Ernährung (Nasen- oder PEG-Sonde), n (%)	4 (50)	10 (100)	1 (17)	9 (90)
Unterstützung bei Beatmung (nicht- invasive Beatmung / Tracheostomie und Beatmungsgerät), n (%)	0	8 (80)	0	8 (80)
Unterstützung bei Ernährung und nicht-invasiver Beatmung, n (%)	0	8 (80)	0	8 (80)
CHOP-INTEND-Wert				
n	6	8	4	9
MW (SD)	k. A.	k. A.	k. A.	k. A.
Median (IQA)	32,5 (31; 33)	11,0 (6,5; 18)	28,5 (26,5; 31)	28,0 (11; 30)

¹⁾ Beginn der Symptome im Alter von weniger als 3 Monaten.

²⁾ Beginn der Symptome im Alter von mehr als 3 Monaten.

³⁾ Die Aufnahme in die Studie erfolgte innerhalb von 3 Monaten nach Diagnose.

⁴⁾ Die Aufnahme in die Studie erfolgte, wenn die Diagnose länger zurücklag.

Abkürzungen: CHOP INTEND: Children's Hospital of Philadelphia Infant Test for Neuromuscular Disease; IQA: Interquartilsabstand; k. A.: keine Angabe; MW: Mittelwert; PNCR: Pediatric Neuromuscular Clinical Research Network for SMA; SD: Standardabweichung; SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron.

Anhang 2

Tabelle 42: Charakterisierung der Population der NeuroNEXT-Kohorte [25]

Studie NeuroNEXT Merkmal	SMA N = 26
<i>Geschlecht, n (%)</i> männlich weiblich	11 (42) 15 (58)
<i>Ethnie (genetisch), n (%)</i> kaukasisch/weiß keine Angabe	24 (92) 2 (8)
<i>Ethnizität, n (%)</i> hispanisch keine Angabe	6 (23) 20 (77)
<i>Alter zum Studieneinschluss (Monate)</i> MW (SD) Median (min; max)	3,7 (1,7) k. A.
<i>Gestationsalter (Wochen)</i> MW (SD) Median (min; max)	38,8 (1,5) k. A.
<i>SMN2-Genkopien, n (%)</i> 1 2 3 4 unbekannt	0 16 (61,5) 5 (19,2) 1 (3,8) 4 (15,4)
Mutation c.859G>C im SMN2-Gen	0

Abkürzungen: k. A.: keine Angabe; MW: Mittelwert; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; SD: Standardabweichung; SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron.

Anhang 3

Studiencharakteristika der Studien CL-101 und CL-303 vs. externe Kontrollen (PNCr, NeuroNEXT, ENDEAR)

Für den Vergleich der Zulassungsbegründenden Studien CL-303 und CL-101 wurden 23 Personen aus der Studie PNCr, 16 aus der Studie NeuroNEXT und 41 (ITT-Population) aus der BSC-Kohorte der Studie ENDEAR herangezogen (siehe Abschnitt 2.2.2).

In Bezug auf die Patientenpopulation bestehen zwischen den Studien CL-101 und CL-303 und der Studie PNCr insbesondere Unterschiede bezogen auf das Alter bei Einschluss in die Studie und das Alter bei Auftreten der ersten Symptome. So lag das mittlere Alter (SD) zum Studieneinschluss in der Studie CL-303 bei 3,7 Monaten (1,6) und in der Studie CL-101 bei 3,4 Monaten (2,1), während es in der Studie PNCr im Mittel bei 29,0 Monaten (41,7) lag und somit auch ältere Personen als in den Zulassungsstudien umfasste. Demgegenüber lag das mittlere Alter bei Studieneinschluss in der NeuroNEXT-Population bei 4,1 Monaten (1,7) und in der BSC-Kohorte der ENDEAR-Studie bei 5,5 Monaten (1,6), was etwas höher als in den Zulassungsstudien ist, aber in etwa vergleichbar scheint. Während die Symptome in der Studie CL-303 erstmals zwischen 0,5 und 5,9 Monaten und in der Studie CL-101 zwischen 0 und 3 Monaten auftraten, lag dieser Zeitpunkt in der Studie PNCr bei 0,5 bis 6 Monaten und in der Studie ENDEAR bei 0,25 bis 5 Monaten und damit später als in den Zulassungsstudien. Für die Studie NeuroNEXT liegen diesbezüglich keine Daten vor (siehe Tabelle 43).

Während in der Studie CL-303 keine Person eine Beatmungs- und Ernährungsunterstützung benötigte und in der Studie CL-101 83,3 % bzw. 58,3 % keine Beatmungs- bzw. Ernährungsunterstützung benötigten, betraf dies in der Studie PNCr zum Studieneinschluss nur weniger als 50 % (Ernährung: 21,7 %, Beatmung: 47,8 %) und in der Studie NeuroNEXT zu Baseline weniger als 65 % (Ernährung: 56,3 %, Beatmung: 62,5 %).

Tabelle 43: Charakterisierung der Studienpopulationen CL-101 und CL-303 mit den externen Kontrollen (PNCr, NeuroNEXT, ENDEAR (BSC)) [7,12,21]

Studien Merkmal	CL-101 N = 12	CL-303 N = 22	PNCr N = 23	NeuroNEXT N = 16	ENDEAR N = 41
Geschlecht, n (%)					
männlich	5 (41,7)	10 (45,5)	11 (48,8)	8 (50,0)	17 (41,5)
weiblich	7 (58,3)	12 (54,5)	12 (52,2)	8 (50,0)	24 (58,5)
Ethnie (genetisch), n (%)					
kaukasisch/weiß	11 (91,7)	11 (50,0)	16 (69,6)	15 (93,8)	36 (87,8)
unbekannt bzw. andere	1 (8,3)	11 (50)	7 (30,4)	1 (6,2)	5 (12,2)
Alter zum Studieneinschluss (Monate)					
MW (SD)	3,4 (2,1)	3,7 (1,6)	29,0 (41,7)	4,1 (1,7)	5,5 (1,6) ²⁾
Median (min; max)	3,1 (0,9; 7,9)	3,5 (0,5; 5,9)	k. A. (2; 171)	k. A. (0; 6)	k. A. (0,7; 7,0) ²⁾
Alter bei Symptombeginn (Monate)					
MW (SD)	1,4 (1,0)	1,9 (1,2)	3,0 (1,6)	n. a.	2,4 (k. A.) ³⁾
Median (min; max)	1,0 (0; 3)	1,8 (0; 4)	k. A. (0,5; 6)	n. a.	k. A. (0,25; 5)
Alter bei genetischer SMA-Diagnose (Monate)					
MW (SD)	k. A.	k. A.	n. a.	n. a.	69,6 (k. A.) ⁴⁾
Median (min; max)	k. A.	k. A.	n. a.	n. a.	k. A. (14; 210)
Benötigt keine Unterstützung von¹⁾, n (%)					
Ernährung	7 (58,3)	22 (100)	5 (21,7)	9 (56,3)	36 (87,8)
Beatmung	10 (83,3)	22 (100)	11 (47,8)	10 (62,5)	35 (85,4)
Ernährung und Beatmung	10 (83,3)	22 (100)	11 (47,8)	14 (87,5)	k. A.
Beatmung vor dem 6. Lebensmonat	10 (83,3)	22 (100)	21 (91,3)	10 (62,5)	k. A.
CHOP-INTEND-Wert					
MW (SD)	28,2 (12,3)	32 (9,7)	24,6 (11,6)	20,3 (7,3)	28,4 (7,6) ⁵⁾
Median (min; max)	29,0 (12; 50)	33,5 (18; 52)	k. A. (5; 40)	k. A. (10; 33)	28,5 (10,5; 50,5)

¹⁾ Zu Baseline in der Studie CL-101 und CL-303 sowie NeuroNEXT. Zum Studieneinschluss in der Studie PNCr.

²⁾ Gemäß Angaben des pU im Nutzendossier (Modul 4): „Angaben aus dem Nusinersen-Dossier bzw. der Publikation von Finkel et al. (2017), dort in Tagen angegeben, für die Angabe der Monate wurde durch 30 geteilt.“

³⁾ Gemäß Angaben des pU im Nutzendossier (Modul 4): „Angaben aus dem Nusinersen-Dossier, dort in Wochen angegeben, für die Angabe der Monate wurde durch 4 geteilt.“

⁴⁾ Gemäß Angaben des pU im Nutzendossier (Modul 4): „Angaben aus dem Nusinersen-Dossier, dort in Wochen angegeben, für die Angabe in Tagen wurde mit 7 multipliziert.“

⁵⁾ Gemäß Angaben des pU im Nutzendossier (Modul 4): „Angaben aus der Publikation von Finkel et al. (2017).“

Abkürzungen: BSC: Best Supportive Care; CHOP INTEND: The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; k. A.: keine Angabe; MW: Mittelwert; n. a.: nicht verfügbar; NeuroNEXT: National Network for Excellence in Neuroscience Clinical Trials; PNCr: Pediatric Neuromuscular Clinical Research Network for SMA; pU: pharmazeutischer Unternehmer; SD: Standardabweichung; SMA: Spinale Muskelatrophie.