

Nutzenbewertung



**Gemeinsamer
Bundesausschuss**

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

**Bewertung von Arzneimitteln für seltene Leiden
nach § 35a Absatz 1 Satz 10 i.V.m.**

5. Kapitel § 12 Nr. 1 Satz 2 VerfO

Wirkstoff: Nusinersen

Datum der Veröffentlichung: 2. Oktober 2017

Inhaltsverzeichnis

Tabellenverzeichnis.....	3
Abbildungsverzeichnis.....	5
Abkürzungsverzeichnis	6
Hintergrund.....	8
1 Einführung	9
2 Nutzenbewertung	11
2.1 Fragestellung.....	11
2.2 Zulassungsbegründende Studie(n).....	11
2.3 Liste der verwendeten Quellen.....	14
2.4 Studiencharakteristika	14
2.5 Endpunkte und statistische Methoden	24
2.5.1 In der Nutzenbewertung berücksichtigte Endpunkte.....	24
2.5.2 Verzerrungspotential auf Studien- und Endpunktebene.....	28
2.5.3 Charakterisierung der berücksichtigten Endpunkte und Einschätzung der Patientenrelevanz	30
2.5.4 Statistische Methoden.....	42
2.6 Ergebnisse zum Zusatznutzen	45
2.6.1 Charakterisierung der Studienpopulation	46
2.6.2 Mortalität.....	50
2.6.3 Morbidität.....	51
2.6.4 Lebensqualität	60
2.6.5 Sicherheit	61
2.6.6 Subgruppenanalysen	65
3 Methodische Anmerkungen und Bewertung der Unterlagen.....	67
3.1 Zulassungsstatus und Zulassungspopulation von Nusinersen (Spinraza®).....	67
3.2 Design und Methodik der Studien ENDEAR und CHERISH	67
3.3 Wirksamkeit	69
3.4 Lebensqualität	72
3.5 Sicherheit.....	73
4 Anforderung an eine qualitätsgesicherte Anwendung	75
5 Zusammenfassung der Bewertung	76
Referenzen	78

Tabellenverzeichnis

Tabelle 1:	Übersicht supportiver Studien mit dem Wirkstoff Nusinersen	13
Tabelle 2:	Charakterisierung der Studie ENDEAR	14
Tabelle 3:	Protokolländerungen der Studie ENDEAR	18
Tabelle 4:	Charakterisierung der Studie (CHERISH)	19
Tabelle 5:	Protokolländerungen der Studie (CHERISH)	22
Tabelle 6:	Charakterisierung der Intervention in der Studie ENDEAR	24
Tabelle 7:	Charakterisierung der Intervention in der Studie (CHERISH)	24
Tabelle 8:	In der Nutzenbewertung berücksichtigte Endpunkte der Studien ENDEAR und (CHERISH)	25
Tabelle 9:	Verzerrungspotential der Studien ENDEAR und (CHERISH) auf Studienebene	28
Tabelle 10:	Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte der Studien ENDEAR und (CHERISH)	29
Tabelle 11:	Beschreibung der Endpunkte der Kategorie Mortalität in der Studie ENDEAR	31
Tabelle 12:	Beschreibung der Endpunkte der Kategorie Morbidität in den Studien ENDEAR und (CHERISH)	31
Tabelle 13:	Beschreibung der Endpunkte der Kategorie Lebensqualität in der Studie (CHERISH)	38
Tabelle 14:	Beschreibung der Endpunkte der Kategorie Sicherheit in den Studien ENDEAR und (CHERISH)	40
Tabelle 15:	Allgemeine Angaben der Studien ENDEAR und (CHERISH)	47
Tabelle 16:	Charakterisierung der Studienpopulation der Studien ENDEAR und (CHERISH)	48
Tabelle 17:	Ergebnisse zum Endpunkt Zeit bis zum Tod in der Studie ENDEAR – ITT-Population	51
Tabelle 18:	Ergebnisse zum Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung in der Studie ENDEAR – ITT-Population	52
Tabelle 19:	Ergebnisse zum Endpunkt Zeit bis zur dauerhaften Beatmung in der Studie ENDEAR – ITT-Population	52
Tabelle 20:	Ergebnisse zum Endpunkt Anteil der HINE-Responder (Subskala 2) in der Studie ENDEAR – ITT-Population	54
Tabelle 21:	Ergebnisse zum Endpunkt Anteil der HINE-Responder (Subskala 2) in der Studie ENDEAR – ES-Population	54
Tabelle 22:	Ergebnisse zum Endpunkt Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline in der Studie ENDEAR – ITT-Population	55
Tabelle 23:	Ergebnisse zum Endpunkt Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline in der Studie ENDEAR – ES-Population	55

Tabelle 24:	Ergebnisse zum Endpunkt Anteil der CHOP-INTEND-Responder in der Studie ENDEAR – ITT-Population.....	56
Tabelle 25:	Ergebnisse zum Endpunkt Anteil der CHOP-INTEND-Responder in der Studie ENDEAR – ES-Population	56
Tabelle 26:	Ergebnisse zum Endpunkt Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline in der Studie ENDEAR – ITT-Population.....	57
Tabelle 27:	Ergebnisse zum Endpunkt Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline in der Studie ENDEAR – ES-Population	57
Tabelle 28:	Ergebnisse zum Endpunkt Veränderung des HFMSE-Scores im Vergleich zu Baseline in der Studie CHERISH – ITT-Population	58
Tabelle 29:	Ergebnisse zum Endpunkt Anteil der HFMSE-Responder in der Studie CHERISH – ITT-Population.....	58
Tabelle 30:	Ergebnisse zum Endpunkt Häufigkeit schwerwiegender respiratorischer Ereignisse in der Studie ENDEAR – ITT-Population	59
Tabelle 31:	Ergebnisse zum Endpunkt Häufigkeit schwerwiegender respiratorischer Ereignisse in der Studie CHERISH – ITT-Population	59
Tabelle 32:	Ergebnisse zum Endpunkt Häufigkeit von Hospitalisierungen in der Studie ENDEAR – ITT-Population.....	60
Tabelle 33:	Ergebnisse zum Endpunkt Häufigkeit von krankheitsbedingten Hospitalisierungen in der Studie CHERISH – ITT-Population.....	60
Tabelle 34:	Ergebnisse zum Endpunkt Unerwünschte Ereignisse in der Studie ENDEAR – SAS-Population	63
Tabelle 35:	Ergebnisse zum Endpunkt Unerwünschte Ereignisse in der Studie CHERISH – SAS-Population.....	64
Tabelle 36:	Zusammenfassende Darstellung der Ergebnisse der Studien ENDEAR und CHERISH.....	76

Abbildungsverzeichnis

Abbildung 1: Klinischer Entwicklungsplan Nusinersen12

Abkürzungsverzeichnis

ACEND	Assessment of Caregiver Experience with Neuromuscular Disease
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Analysis of Covariance
BSC	Best Supportive Care
CGI	Clinical Global Impression of Change
CHOP INTEND	Children's Hospital of Philadelphia Infant Test for Neuromuscular Disease
CMAP	Compound Muscle Action Potential
EAP	Expanded Access Program
EMA	European Medicines Agency
EPAR	European Public Assessment Report
ES	Efficacy-Set
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
HR	Hazard Ratio
HFMSE	Hammersmith Functional Motor Scale Expanded
HINE	Hammersmith Infant Neurological Examination
ICC	Interkorelationskoeffizient
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-treat
IVRS	Interactive Voice Response System
KI	Konfidenzintervall
LP	Lumbalpunktion
LS	Least Square
MCID	Minimal Clinical Important Difference
MedDRA	Medical Dictionary for Regulatory Activities
MW	Mittelwert
n / N	Anzahl Patienten
PedsQL	Pediatric Quality of Life Inventory
PK	Pharmakokinetik
PP	Per-Protokoll
pU	pharmazeutischer Unternehmer
RCT	randomisierte kontrollierte Studie/n (Randomized Controlled Trial/s)
RR	Relatives Risiko
RULM	Revised Upper Limb Module

SAS	Safety-Analysis-Set
SD	Standardabweichung
SGB	Sozialgesetzbuch
SMA	Spinale Muskelatrophie
SMN	Survival Motor Neuron
SOC	Systemorganklasse/n
SUE	schwerwiegende/s unerwünschte/s Ereignis/se
UE	unerwünschte/s Ereignis/se
VerfO	Verfahrensordnung des G-BA
ZNS	Zentralnervensystem

Hintergrund

Nusinersen (Spinraza®) 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 Abs. 1 Satz 10 SGB V gilt der medizinische Zusatznutzen durch die Zulassung als belegt.

Gemäß § 35a Abs. 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 Abs. 1 Satz 10 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 Abs. 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 Abs. 1 Satz 11 SGB V).

Der G-BA bestimmt gemäß 5. Kapitel § 12 Abs. 1 Nr. 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 Nusinersen zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers heran und bewertet die darin enthaltenen Daten nach Maßgabe der im 5. Kapitel § 5 Abs. 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Nusinersen in seiner Sitzung am 26. September 2017 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 30. Juni 2017 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pharmazeutischen Unternehmers in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

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

Die Nutzenbewertung wird am 2. Oktober 2017 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 Einführung

Als Spinale Muskelatrophie (SMA) ist eine Gruppe genetischer Störungen benannt. Diese sind durch eine fortschreitende Degeneration von Motorneuronen in Rückenmark und Hirnstamm gekennzeichnet. In der Folge treten niedriger Blutdruck (Hypotonie), Rückbildung (Atrophie) der Muskulatur und allgemeine Schwäche auf. Es gibt vier Arten von SMA. Typ I ist bereits zum Zeitpunkt der Geburt erkennbar. Die anderen Typen manifestieren sich typischerweise sukzessiv in späteren Altersgruppen. SMA wird durch einen Defekt im „survival motor neuron 1“ (SMN1) verursacht. In fast allen Fällen ergeben sich die Defekte aus Lösungen, Neuordnungen oder Mutationen im SMN1-Gen [15].

Menschen mit SMA sind in der Regel durch Hypotonie und Schwäche im Alltagsleben eingeschränkt. Es treten Muskelschwäche und Muskelschwund, einschließlich der Lungenmuskulatur auf. Eine geschwächte Atemmuskulatur ist eine häufige Konsequenz der SMA und kann zu respiratorischen Versagem, zu Aspirationspneumonie oder rezidivierenden respiratorischen Infektionen, die die Muskelschwäche verstärken können, führen. Die Muskelschwäche zeigt sich typischerweise symmetrisch, progressiv, eher proximal als distal und stärker in den Beinen als in den Armen. Die Erkrankung ist auf einen Defekt des Chromosoms 5q zurückzuführen und die Symptome treten in der Regel kurz nach der Geburt auf. Der Zeitpunkt für das Auftreten der Symptome legt den SMA-Typ fest. Typ I zeigt sich bei Säuglingen im Alter unter sechs Monaten mit fortschreitender Hypotonie, symmetrischer Schwäche der Extremitäten, geringer Kopfkontrolle und reduzierten oder fehlenden Reflexen (Areflexie). Typ II zeigt sich bei Säuglingen oder Kindern mit zunehmender proximaler Beinschwäche, Hypotonie und Areflexie. Typ III zeigt sich bei Kindern und Erwachsenen mit zunehmender proximaler Schwäche häufiger in den Beinen als in den Armen. Typ IV zeigt sich mit leichter proximaler Schwäche der Extremitäten, die sich im Erwachsenenalter entwickelt. Bei Verdacht auf eine SMA kann durch die genetische Untersuchung der Exon-7-Deletion auf dem SMN1-Gen im Chromosom 5 die Diagnose gesichert werden.

SMA kann nicht geheilt werden. Ziele der Therapie sind a) die Reduktion der Belastung durch die Erkrankung für die Betroffenen, b) Reduktion von Hospitalisierungen und c) Versorgung bei Komplikationen. Zur Behandlung der Komplikationen kommen folgende Interventionen zum Einsatz: Unterstützung der Atmung oder Beatmung, Optimierung des Ernährungszustandes, Erhaltung der Fitness und Ausdauer durch Physiotherapie und Ergotherapie für den Umgang mit der Bewegungsschwäche. In schweren Fällen kann eine palliative Therapie Unterstützung bieten. Um Infektionen zu vermeiden sind verfügbare Impfungen empfohlen.

Nusinersen

Nusinersen (Spinraza®) ist ein synthetisches Antisense-Oligonukleotid (eine synthetische, kurzket-tige, einzelsträngige Nukleinsäure). Es befähigt das SMN2-Gen zu einer normalen Funktion, d. h. es werden funktionierende vollständige Proteinie produziert. Das durch den SMN1-Genedefekt fehlende Protein wird ersetzt, die Symptome der Erkrankung dadurch gelindert [18].

Nusinersen wird zur Behandlung der SMA angewendet. Die empfohlene Dosis beträgt 12 mg (5 ml) pro Anwendung. Es wird von einem Arzt oder medizinischen Fachpersonal als intrathekale Injektion (als Lumbalpunktion) in den Rückenmarkskanal verabreicht. Eine Sedierung des Patienten zur Beruhigung vor Verabreichung von Nusinersen ist zu erwägen. Die Verabreichung der ersten Dosis sollte sobald als möglich nach Diagnosestellung mit nachfolgenden Aufsättigungsdosen an Tag 14, 28 und 63 erfolgen. Im Anschluss ist eine Erhaltungsdosis alle 4 Monate zu verabreichen [12].

Die Einstufung als Arzneimittel für seltene Leiden erhielt Nusinersen (Spinraza®) am 2. April 2012, die Zulassung in der EU erfolgte am 30. Mai 2017.

2 Nutzenbewertung

2.1 Fragestellung

Nusinersen (Spinraza®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Abs. 1 Satz 10 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 Patienten in folgender Indikation eingeschlossen:

- Patienten mit 5q-assoziiertes spinaler Muskelatrophie.

2.2 Zulassungsbegründende Studie(n)

Pivotal Studie

- ISIS 396443-CS3B (ENDEAR). A Phase 3, Randomized, Double-blind, Sham-Procedure Controlled Study to Assess the Clinical Efficacy and Safety of ISIS 396443 Administered Intrathecally in Patients With Infantile-onset Spinal Muscular Atrophy [10].
- ISIS 396443-CS4 (CHERISH). A Phase 3, Randomized, Double-blind, Sham-Procedure Controlled Study to Assess the Clinical Efficacy and Safety of ISIS 396443 Administered Intrathecally in Patients With Later-onset Spinal Muscular Atrophy ISIS 396443-CS3A (CS3A) [11].

Supportive Studien

- 232SM201 (NURTURE). An Open-Label Study to Assess the Efficacy, Safety, Tolerability, and Pharmacokinetics of Multiple Doses of ISIS 396443 Delivered Intrathecally to Subjects With Genetically Diagnosed and Presymptomatic Spinal Muscular Atrophy [4].
- 232SM202 (EMBRACE). A Phase 2, Randomized, Double-blind, Sham-procedure Controlled Study to Assess the Safety and Tolerability and Explore the Efficacy of ISIS 396443 (BIIB058) Administered Intrathecally in Subjects With Spinal Muscular Atrophy Who Are Not Eligible to Participate in the Clinical Studies ISIS 396443-CS3B or ISIS 396443-CS4 [9].
- 232-SM-901 (EAP). Expanded Access Program (EAP) to Provide Nusinersen to Patients With Infantile-onset Spinal Muscular Atrophy (SMA) [2].
- ISIS 396443-CS1 (CS1). An Open-label, Escalating Dose Study to Assess the Safety, Tolerability and Dose-range Finding of a Single Intrathecal Dose of ISIS 396443 in Patients With Spinal Muscular Atrophy [8].
- ISIS 396443-CS2 (CS2). An Open-Label, Dose Escalation Study to Assess the Safety, Tolerability and Dose-Range Finding of Multiple Doses of ISIS 396443 Delivered Intrathecally to Patients With Spinal Muscular Atrophy [7].
- ISIS 396443-CS3A (CS3A). A Study to Assess the Efficacy, Safety, Tolerability, and Pharmacokinetics of Multiple Doses of ISIS 396443 Delivered Intrathecally to Patients With Infantile-Onset Spinal Muscular Atrophy [13].

- ISIS 396443-CS10 (CS10). An Open-label Study to Assess the Safety and Tolerability of a Single Intrathecal Dose of ISIS 396443 in Patients With Spinal Muscular Atrophy Who Previously Participated in ISIS 396443-CS1 [5].
- ISIS 396443-CS11 (CS11, SHINE). An Open-Label Extension Study for Patients With Spinal Muscular Atrophy Who Previously Participated in Investigational Studies of ISIS 396443 [3].
- ISIS 396443-CS12 (CS12). An Open-label Study to Assess the Safety and Tolerability of ISIS 396443 in Patients With Spinal Muscular Atrophy Who Previously Participated in 396443-CS2 or 396443-CS10 [6].

Neben den pivotalen Studien CS3B (ENDEAR) und CS4 (CHERISH) beinhaltet das Studienprogramm des pharmazeutischen Unternehmers (pU) zahlreiche supportive Studien (CS1, CS2, CS3A, CS10, CS12, EAP und 232SM201), die nicht für die vorliegende Nutzenbewertung von Nusinersen herangezogen werden. 232SM201 (NURTURE), 232SM202 (EMBRACE) und CS11 (SHINE) sind laufende Studien, Studienabschluss ist für 2020, 2019 bzw. 2022 geplant.

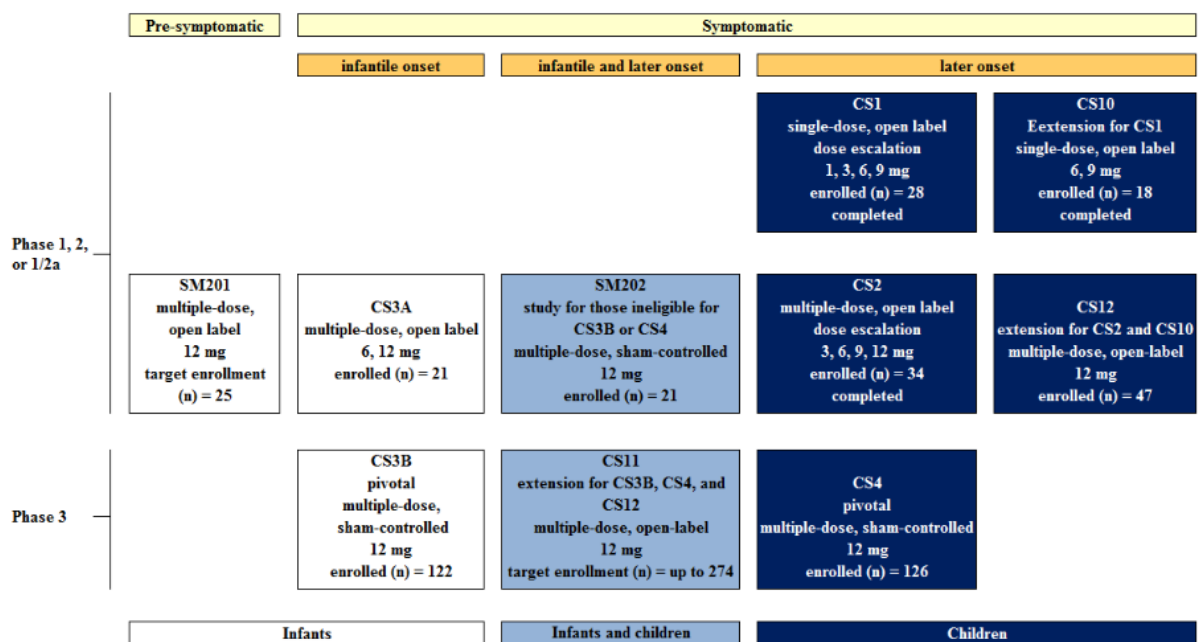


Abbildung 1: Klinischer Entwicklungsplan Nusinersen [39]

Tabelle 1: Übersicht supportiver Studien mit dem Wirkstoff Nusinersen¹⁾

Studie	Studiencharakteristika	Kommentar
Präsymptomatisch SMA		
SM201 (NURTURE)	<u>Design:</u> Offene, multizentrische, einarmige Phase-II-Präventionsstudie, 12 mg Nusinersen (altersadjustiert, intrathekal) <u>Einschluss:</u> N=25, Alter: ≤ 6 Wochen <u>Primärer Endpunkt:</u> Zeit bis zum Tod oder zur dauerhaften Beatmung (invasive oder nicht-invasive Ventilation ≥ 6 Stunden/Tag durchgehend für > 7 Tage oder Tracheotomie) <u>Studiendauer:</u> 5 Jahre <u>Status:</u> laufend seit 20. Mai 2015 bis 2022	<u>Limitation:</u> Präsymptomatische Patienten, noch keine ausreichende Laufzeit; nicht-vergleichende Studie. Der pU legt Daten von einer Zwischenanalyse mit Datenschnitt 31. Oktober 2016 vor. <u>EMA:</u> <u>Datenschnitt: 8. Juni 2017</u>
Frühkindliche symptomatische SMA		
CS3A	<u>Design:</u> Offene, zweiarmige Phase-II-Studie, 6 mg und 12 mg Nusinersen (altersadjustiert, intrathekal), dann 12 mg als Erhaltungstherapie <u>Einschluss:</u> N=20, Alter: ≥ 21 und ≤ 210 Tage <u>Primärer Endpunkt:</u> Anteil Patienten mit motorischen Meilenstein gemessen mittels Subskala 2 des HINE <u>Studiendauer:</u> 45 Monate <u>Status:</u> laufend seit 8. Mai 2013 (bis 2017)	<u>Limitation:</u> Laufende nicht aktiv-vergleichende; Dosisescalationsstudie <u>EMA:</u> <u>Datenschnitt 26. Januar 2016</u>
Frühkindliche und später einsetzende symptomatische SMA		
CS11 (SHINE)	<u>Design:</u> Offene, einarmige Phase-III-Extensionsstudie der Studien CS3A, ENDEAR (CS3B), CHERISH (CS4), CS12, Nusinersen 12 mg (altersadjustiert, intrathekal) <u>Einschluss:</u> N=274 (bis zu 289), Alter: alle Altersgruppen; <u>Primärer Endpunkt:</u> Verträglichkeit und Auftreten unerwünschter Ereignisse <u>Studiendauer:</u> 52 Monate <u>Status:</u> laufend (bis 2022)	<u>Limitation:</u> Laufende nicht aktiv-vergleichende Studie <u>EMA:</u> <u>Erwartet Langzeitdaten aus der Studie zu August 2023</u>
SM202 (EMBRACE)	<u>Design:</u> Randomisierte, doppelblinde, kontrollierte Phase-II-Studie, 12 mg Nusinersen, (intrathekal) oder Scheinbehandlung; 14 Monate mit anschließender offener Extensionsstudie über 30 Monate <u>Einschluss:</u> N=21, Alter: > 210 Tage, Einschlusskriterien der Studien CS3B und CS4 nicht erfüllt <u>Primärer Endpunkt:</u> Unerwünschte Ereignisse und schwerwiegende unerwünschte Ereignisse <u>Studiendauer:</u> 44 Monate <u>Status:</u> laufend (bis 2019)	<u>Limitation:</u> Laufender RCT <u>EMA:</u> <u>Erwartet Daten in 2019</u>
Später einsetzende symptomatische SMA		
CS1	<u>Design:</u> Offene, einarmige Phase-I-Studie, 1, 3, 6 oder 9 mg Nusinersen intrathekal <u>Einschluss:</u> N=28, Alter: 2–14 Jahre <u>Primärer Endpunkt:</u> Unerwünschte Ereignisse <u>Studiendauer:</u> 14 Monate <u>Status:</u> abgeschlossen (bis 2013)	<u>Limitation:</u> Nicht aktiv-vergleichende Studie, Dosierung nicht zulassungskonform <u>EMA:</u> =

Studie	Studiencharakteristika	Kommentar
CS2	<u>Design:</u> Offene, einarmige Phase-I/II-Studie, 3, 6, 9 oder 12 mg Nusinersen intrathekal <u>Einschluss:</u> N=34, Alter: 2–15 Jahre <u>Primärer Endpunkt:</u> Unerwünschte Ereignisse (UE), Schwerwiegende UE, Therapieabbruch aufgrund von UE, UE nach Schweregrad <u>Studiendauer:</u> 28 Monate <u>Status:</u> abgeschlossen (bis 2015)	<u>Limitation:</u> Nicht aktiv-vergleichende Dosis- eskalationsstudie; teilweise nicht zulassungskonform <u>EMA:</u> -
CS10	<u>Design:</u> Offene Phase-I-Studie, 6 oder 9 mg Nusinersen intrathekal, Extensionsstudie der Studie CS1 <u>Einschluss:</u> N=18, Alter: 2–15 Jahre <u>Primärer Endpunkt:</u> Unerwünschte Ereignisse <u>Studiendauer:</u> 13 Monate <u>Status:</u> abgeschlossen (bis 2014)	<u>Limitation:</u> Nicht aktiv-vergleichende Sicher- heitsstudie; Dosierung nicht zulas- sungskonform <u>EMA:</u> -
CS12	<u>Design:</u> Offene Phase-I-Studie, 12 mg Nusinersen intrathekal, Extensionsstudie der Studie CS2 <u>Einschluss:</u> N=47, Alter: alle Altersgruppen <u>Primärer Endpunkt:</u> Unerwünschte Ereignisse <u>Studiendauer:</u> 36 Monate <u>Status:</u> abgeschlossen (bis 2017)	<u>Limitation</u> Nicht aktiv-vergleichende Sicher- heitsstudie <u>EMA</u> -

¹⁾ Quelle: clinicaltrials.gov

2.3 Liste der verwendeten Quellen

Zur Bestimmung des Ausmaßes des Zusatznutzens für Nusinersen wurden folgende Studien und Daten herangezogen:

- Herstellerdossier des pU zu Nusinersen [1]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [16,17,18]
- Studienbericht inkl. Studienprotokoll und Statistischem Analyseplan der Studie CS3B (ENDEAR) [10] und der Studie CS4 (CHERISH) [11]

2.4 Studiencharakteristika

Die Unterlagen zum Nachweis des Ausmaßes des Zusatznutzens für Nusinersen basieren auf den Zulassungsstudien CS3B (nachfolgend ENDEAR) und CS4 (nachfolgend CHERISH). Die wesentlichen Studiencharakteristika sind in den Tabellen 2 bis 7 dargestellt.

Tabelle 2: Charakterisierung der Studie ENDEAR (Studienpopulation: Säuglinge ≤ 7 Monate)

Charakteristikum	Beschreibung
Design	Randomisierte, doppelblinde, kontrollierte, multizentrische Phase-III-Studie zur Untersuchung der Wirksamkeit, Sicherheit, Verträglichkeit und Pharmakokinetik von Nusinersen im Vergleich zu einer Scheinintervention bei Patienten mit infantiler spinaler Muskelatrophie (SMA). Die Randomisierung der Patienten auf die Vergleichsarme erfolgte im Verhältnis 2:1 und wurde anhand der Krankheitsdauer (Alter des Patienten bei Screening - Alter bei Symptombeginn: ≤ 12 vs. > 12 Wochen) stratifiziert. Für den primären Endpunkt waren eine Interimsanalyse (15. Juni 2016) und eine finale Analyse vorgesehen. Das Ende der Studie war für Juli 2017 geplant, jedoch wurde die

Dossierbewertung für Orphan Drugs

Nusinersen (Spinraza®)

Charakteristikum	Beschreibung
	Studie aufgrund eines erzielten Wirksamkeitsnachweises vorzeitig beendet (21. November 2016).
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • Jungen und Mädchen im Alter von ≤ 7 Monaten (210 Tagen) zum Zeitpunkt des Screenings. • Genetische Dokumentation der homozygoten Deletion, homozygoten Mutation oder Compound-Heterozygotie von 5q SMA. • Survival Motor Neuron (SMN)2-Kopienanzahl = 2. • Einsetzen der SMA-typischen Symptome im Alter von ≤ 6 Monaten (180 Tagen). • Bei Studieneintritt adäquate Nahrungs- und Flüssigkeitszufuhr (mit oder ohne Gastrostomie) nach Einschätzung des Prüfarztes. • Altersgerechtes Körpergewicht von ≥ 3. Perzentil auf Basis landesspezifischer Richtlinien. • Vorliegen entsprechender medizinischer Versorgung, wie z. B. Routineimpfungen (einschl. Grippe-, Pneumokokken- und Pneumovirus-Prophylaxe (Palivizumab), soweit verfügbar). Die Versorgung entspricht nach Einschätzung des Prüfarztes auch weiterhin den Richtlinien, die im Konsensus-Statement für die Best Supportive Care (BSC) für SMA aufgeführt werden. • Geboren zwischen 37. und 42. Schwangerschaftswoche. • Lebt für die Dauer der Studie ca. 9 Stunden vom teilnehmenden Studienzentrum entfernt; ein Wohnsitz von > 2 Stunden entfernt von einem Studienzentrum muss durch einen Prüfarzt und den medizinischen Monitor der Studie genehmigt werden. • Fähigkeit, alle Untersuchungen, Messungen und Studienbesuche vollständig zu absolvieren sowie Vorliegen einer angemessenen psychosozialen Unterstützung für Eltern oder Erziehungsberechtigte / den Patienten nach Einschätzung des Prüfarztes. <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Hypoxämie (O_2-Sättigung wach < 96 % oder O_2-Sättigung schlafend < 96 % ohne unterstützende Beatmung) während des Screenings. • Anzeichen oder Symptome von SMA ab Geburt oder innerhalb der ersten Woche nach Geburt. • Vorliegen einer unbehandelten oder unzureichend behandelten aktiven Infektion, die eine systemische antivirale oder antimikrobielle Therapie zu jeglichem Zeitpunkt während des Screenings erfordert. • Erkrankung des Gehirns oder Liquors, welche eine Lumbalpunktion (LP), die Zirkulation der Rückenmarksflüssigkeit oder die Sicherheitsbewertungen beeinträchtigen würde. • Vorliegen eines implantierten Shunts für die Drainage von Liquor oder eines implantierten ZNS-Katheters. • 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 Vorbehandlung mit Gentherapie, Antisense-Oligonukleotiden oder Zelltransplantation. • Anhaltender medizinischer Zustand, der die Durchführung und die Untersuchungen der Studie nach Einschätzung des Prüfarztes beeinträchtigen würde. Beispiele hierfür sind medizinische Behinderungen mit Ausnahme von SMA, die eine Untersuchung der Sicherheit beeinträchtigen oder die Fähigkeit des Patienten einschränken würde, sich den Untersuchungen zu unterziehen.

Charakteristikum	Beschreibung																								
Intervention und Zahl der Patienten	<p>Dosierung</p> <table border="1" data-bbox="453 427 1366 640"> <tr> <td data-bbox="453 427 906 591"> Nusinersen Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 15, 29, 64, 183 und 302. </td> <td data-bbox="906 427 1366 591"> Scheinintervention unter BSC Scheinintervention in Form eines Nadelstichs am unteren Rücken (keine LP) an den Studientagen 1, 15, 29, 64, 183 und 302. </td> </tr> <tr> <td data-bbox="453 591 906 640">N=80</td> <td data-bbox="906 591 1366 640">N=41</td> </tr> </table> <p>Eine Sedierung in Abhängigkeit des Zustandes des Patienten ist ggf. erforderlich. Zudem ist insbesondere bei jüngeren Patienten und Patienten mit Skoliose zur Überwachung der intrathekalen Applikation von Nusinersen ein Ultraschall (oder andere bildgebende Verfahren) angezeigt. Die Behandlung mit Nusinersen sollte nur durch Ärzte erfolgen, die Erfahrung in der Durchführung von LP vorweisen können.</p> <p>Zur Sicherstellung der Verblindung wurde Nusinersen bzw. die Scheinintervention in einem separaten Raum von qualifiziertem, nicht verblindeten Studienpersonal verabreicht. Weder Eltern, Prüfarzte, Studienkoordinatoren oder Endpunkterheber waren während der Behandlung anwesend. Nusinersen wurde in einer altersadjustierten Dosis verabreicht (siehe Tabelle unten). Ab einem Alter von 2 Jahren entspricht diese der empfohlenen Dosis von 12 mg (5 ml) pro Anwendung.</p> <table border="1" data-bbox="413 992 1386 1312"> <thead> <tr> <th>Alter</th> <th>Geschätztes Liquor-Volumen (ml)</th> <th>Injektionsvolumen (ml)</th> <th>Dosis (mg)</th> </tr> </thead> <tbody> <tr> <td>0–3 Monate (0–90 Tage)</td> <td>120</td> <td>4</td> <td>9,6</td> </tr> <tr> <td>3–6 Monate (91–180 Tage)</td> <td>130</td> <td>4,3</td> <td>10,3</td> </tr> <tr> <td>6–12 Monate (181–365 Tage)</td> <td>135</td> <td>4,5</td> <td>10,8</td> </tr> <tr> <td>12–24 Monate (366–730 Tage)</td> <td>140</td> <td>4,7</td> <td>11,3</td> </tr> </tbody> </table>	Nusinersen Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 15, 29, 64, 183 und 302.	Scheinintervention unter BSC Scheinintervention in Form eines Nadelstichs am unteren Rücken (keine LP) an den Studientagen 1, 15, 29, 64, 183 und 302.	N=80	N=41	Alter	Geschätztes Liquor-Volumen (ml)	Injektionsvolumen (ml)	Dosis (mg)	0–3 Monate (0–90 Tage)	120	4	9,6	3–6 Monate (91–180 Tage)	130	4,3	10,3	6–12 Monate (181–365 Tage)	135	4,5	10,8	12–24 Monate (366–730 Tage)	140	4,7	11,3
Nusinersen Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 15, 29, 64, 183 und 302.	Scheinintervention unter BSC Scheinintervention in Form eines Nadelstichs am unteren Rücken (keine LP) an den Studientagen 1, 15, 29, 64, 183 und 302.																								
N=80	N=41																								
Alter	Geschätztes Liquor-Volumen (ml)	Injektionsvolumen (ml)	Dosis (mg)																						
0–3 Monate (0–90 Tage)	120	4	9,6																						
3–6 Monate (91–180 Tage)	130	4,3	10,3																						
6–12 Monate (181–365 Tage)	135	4,5	10,8																						
12–24 Monate (366–730 Tage)	140	4,7	11,3																						
Ort und Zeitraum der Durchführung	<p>Studiendauer</p> <p>Einer 21-tägigen Screeningphase schlossen sich eine 10-monatige Behandlungsphase (302 Tage ± 7 Tage) und eine 3-monatige Follow-up-Phase (bis Tag 394 ± 7 Tage) an, in der die Behandlung aufgrund der langen Halbwertszeit von Nusinersen fortgesetzt wurde.</p> <ul style="list-style-type: none"> • Beginn der Studienbehandlung: 21. August 2014 • Interimsanalyse: 16. Juli 2016 • Ende der Studienbehandlung: 21. November 2016 • Finaler Datenschnitt: 16. Dezember 2016 • Veröffentlichung des Studienberichts: 10. Februar 2017 <p>Ort der Durchführung</p> <p>Die Studie ist an 31 Zentren in 13 Ländern durchgeführt worden: Australien, Belgien, Deutschland, Frankreich, Italien, Japan, Kanada, Korea, Schweden, Spanien, Türkei, Vereinigtes Königreich, Vereinigte Staaten von Amerika.</p>																								

Charakteristikum	Beschreibung
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>Primäre Endpunkte</p> <ul style="list-style-type: none"> • Anteil der Hammersmith Infant Neurological Examination (HINE-) Responder (Subskala 2) • Zeit bis zum Tod oder zur dauerhaften Beatmung (≥ 16 Stunden/Tag durchgehend für > 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses ODER Tracheotomie) <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Anteil der Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease (CHOP INTEND-) Responder • Überlebensrate • Prozentualer Anteil der Patienten, die keine dauerhafte Beatmung benötigen • Anteil der Compound Muscle Action Potential (CMAP-) Responder <p>Tertiäre Endpunkte</p> <ul style="list-style-type: none"> • Änderung der Wachstumsparameter im Vergleich zu Baseline (Gewicht gemäß Alter/Größe, Brustumfang, Verhältnis Kopf- zu Brustumfang, Armumfang) • Häufigkeit schwerwiegender respiratorischer Ereignisse • Anzahl der Stunden mit Beatmung • Häufigkeit und Dauer der Hospitalisierungen <p>Sicherheitsendpunkte</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse • Vitalzeichen und Gewicht • Neurologische Untersuchungen • Körperliche Untersuchungen • Klinische Laboruntersuchungen (chemische Serum- und Blutuntersuchungen, Urin-Analyse) • Elektrokardiogramm (EKG) • Begleitmedikation <p>Pharmakokinetikendpunkte</p> <ul style="list-style-type: none"> • Nusinersenspiegel im Liquor • Nusinersenspiegel im Plasma <p>Immunogenitätspunkt</p> <ul style="list-style-type: none"> • Nusinersenantikörper im Plasma
Subgruppenanalysen	<p>A priori geplant</p> <p>Subgruppenanalysen waren a priori für den primären Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung, motorische Meilensteine, CHOP INTEND und Gesamtüberleben geplant:</p> <ul style="list-style-type: none"> • Alter bei Symptombeginn (≤ 12 vs. > 12 Wochen) • Krankheitsdauer (≤ 12 vs. > 12 Wochen) • Geographische Region (Nordamerika vs. Europa vs. Asien-Pazifik-Raum) <p>Post hoc ausgewertete Subgruppen</p> <p>Es wurden zusätzlich post hoc definierte Subgruppenanalysen vorgenommen für:</p> <ul style="list-style-type: none"> • Alter bei der 1. Dosis (< 180 vs. ≥ 180 Tage) • Geschlecht

Abkürzungen: O₂: Sauerstoff; ZNS: Zentralnervensystem.

Das Original-Studienprotokoll [10] vom 25. November 2013 wurde vier Protokolländerungen unterzogen. Die wesentlichen für die Nutzenbewertung relevanten Protokolländerungen sind in der nachfolgenden Tabelle gelistet.

Tabelle 3: Protokolländerungen der Studie ENDEAR (Studienpopulation: Säuglinge ≤ 7 Monate)

Protokollversion	Wesentliche Protokolländerungen
Amendment 1 25. April 2014 (vor Studienbeginn)	Abschließende Aspekte des Studiendesigns.
Amendment 2 20. Juni 2014 (vor Beginn der Studienbehandlung)	<ul style="list-style-type: none"> • Ergänzung, dass alle Ereignisse des primären Endpunkts von einem verblinden, unabhängigen, zentralen Entscheidungskomitee (Endpoint Adjudication Committee) beurteilt werden mussten. • Ergänzung in Bezug auf Zuständigkeiten und Verblindung des Personals, welches Entscheidungen in Bezug auf die Notwendigkeit einer Beatmung und der Durchführung von Wirksamkeitsbewertungen traf (abgesehen Endpunkt HINE).
Amendment 3 26. September 2014 (nur Japan betreffend)	<ul style="list-style-type: none"> • Festlegung einer Nachbeobachtungsphase mittels geplanter Studienvisite zur Nachbeobachtung der Sicherheit bei Patienten, die die Behandlung abbrachen ohne ihr Einverständnis zurückgezogen zu haben. • Ergänzung, dass eine Vollnarkose oder eine Sedierung während der LP nicht erlaubt sind. Lokale Anästhesie (z. B. Lidocain) während der LP war zulässig. • HINE und motorische Meilensteine wurden als Anhänge hinzugefügt.
Amendment 3 22. April 2016	<ul style="list-style-type: none"> • Ergänzung, dass bei einer frühzeitigen Beendigung der Studie, basierend auf den Ergebnissen der Interimsanalyse, die Teilnahme an einer Langzeitstudie möglich war. • Ergänzung bezüglich der Entblindung bestimmter Vertreter des Studiensponsors während der Durchführung der Interimsanalyse. • Anpassung des Besuchsplans für Patienten, bei denen es aufgrund einer Erkrankung zu Behandlungsverzögerungen kam. • Die primären und sekundären Wirksamkeitsendpunkte wurden basierend auf neuen Informationen aus einer Phase-II-Studie und der natürlichen Verlaufsstudie geändert, um die Interpretierbarkeit der Endpunkte bei frühzeitiger Beendigung der Studie zu erhöhen. • Der Endpunkt motorische Meilensteine wurde vom sekundären zum primären Endpunkt erhoben. Zudem wurde die Analyse von „Änderung zu Baseline“ in „Anteile der Responder“ geändert. • Eine Begründung für die Stichprobengröße wurde basierend auf der statistischen Power des neu definierten Endpunktes Responder motorischer Meilensteine hinzugefügt. • Die Analyse zum CHOP INTEND wurde von „Änderung zu Baseline“ in „Anteile der Responder“ geändert • Eine Beschreibung der Endpunkte und der Zeitpunkt der Interimsanalyse wurden ergänzt.
Amendment 3 24. Mai 2016 (nur die Türkei betreffend)	<ul style="list-style-type: none"> • Ergänzung, dass bei einer frühzeitigen Beendigung der Studie, basierend auf den Ergebnissen der Interimsanalyse, die Teilnahme an einer Langzeitstudie oder einer offenen Studie möglich war. • Ergänzung bezüglich der Entblindung bestimmter Vertreter des Studiensponsors während der Durchführung der Interimsanalyse. • Anpassung des Besuchsplans für Patienten, bei denen es aufgrund einer Erkrankung zu Behandlungsverzögerungen kam. • Die primären und sekundären Wirksamkeitsendpunkte wurden basierend auf neuen Informationen aus einer Phase-II-Studie und der natürlichen Verlaufsstudie geändert, um die Interpretierbarkeit der Endpunkte bei frühzeitiger Beendigung der Studie zu erhöhen. • Eine Begründung für die Stichprobengröße wurde basierend auf der statistischen Power des neu definierten Endpunktes Responder motorischer Meilensteine hinzugefügt.

Protokollversion	Wesentliche Protokolländerungen
	<ul style="list-style-type: none"> • Eine Beschreibung der Endpunkte und der Zeitpunkt der Interimsanalyse wurden ergänzt.
Amendment 4 6. Mai 2016 (nur Japan betreffend)	<ul style="list-style-type: none"> • Ergänzung, dass bei einer frühzeitigen Beendigung der Studie, basierend auf den Ergebnissen der Interimsanalyse, die Teilnahme an einer Langzeitstudie möglich war. • Ergänzung bezüglich der Entblindung bestimmter Vertreter des Studiensponsors während der Durchführung der Interimsanalyse. • Meetings für den Review der Daten durch das Data Safety Monitoring Board fanden vierteljährlich statt. • Anpassung des Besuchsplans für Patienten, bei denen es aufgrund einer Erkrankung zu Behandlungsverzögerungen kam. • Die primären und sekundären Wirksamkeitsendpunkte wurden basierend auf neuen Informationen aus einer Phase-II-Studie und der natürlichen Verlaufsstudie geändert, um die Interpretierbarkeit der Endpunkte bei frühzeitiger Beendigung der Studie zu erhöhen. • Eine Begründung für die Stichprobengröße wurde basierend auf der statistischen Power des neu definierten Endpunktes Responder motorischer Meilensteine hinzugefügt. • Eine Beschreibung der Endpunkte und der Zeitpunkt der Interimsanalyse wurden ergänzt.

Abkürzungen: CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease; HINE: Hammersmith Infant Neurological Examination; LP: Lumbalpunktion(en).

Tabelle 4: Charakterisierung der Studie *CHERISH* (Studienpopulation: (Klein-)Kinder 2–12 Jahre alt)

Charakteristikum	Beschreibung
Design	<p>Randomisierte, doppelblinde, scheininterventionskontrollierte, multizentrische Phase-III-Studie zur Untersuchung der Wirksamkeit, Sicherheit, Verträglichkeit und Pharmakokinetik von Nusinersen im Vergleich zu einer Scheinintervention bei Patienten mit spätem Krankheitsbeginn der SMA. Die Randomisierung der Patienten auf die Vergleichsarme erfolgte im Verhältnis 2:1 und wurde anhand des Alters bei Screening (≤ 6 vs. > 6 Jahre) stratifiziert.</p> <p>Für den primären Endpunkt waren eine Interimsanalyse und eine finale Analyse vorgesehen. Die Studie wurde aufgrund eines erzielten Wirksamkeitsnachweises vorzeitig beendet.</p>
Population	<p>Wesentliche Einschlusskriterien</p> <ul style="list-style-type: none"> • Genetische Dokumentation einer homozygoten Deletion oder homozygoten Mutation oder Compound-Heterozygotie von 5q SMA. • Einsetzen der klinischen Anzeichen und Symptome von SMA im Alter von > 6 Monaten. • Männliche und weibliche Patienten im Alter von 2–12 Jahren. • Fähigkeit für selbstständiges Sitzen, aber nie fähig eigenständig zu gehen. • Motor Function Score (Hammersmith Functional Motor Scale Expanded (HFMSE)) ≥ 10 und ≤ 54 bei Screening. • Fähigkeit, alle Untersuchungen, Messungen und Studienbesuche vollständig zu absolvieren sowie Vorliegen einer angemessenen psychosozialen Unterstützung der Eltern oder des Erziehungsberechtigten / des Patienten nach Einschätzung des Prüfarztes. • Geschätzte Lebenserwartung > 2 Jahre ab Screening nach Auffassung des Prüfarztes.

Charakteristikum	Beschreibung				
	<ul style="list-style-type: none"> • Erfüllt altersgemäÙe Vorgaben des Prüfzentrums bezüglich der Anwendung einer Anästhesie und Sedierung, falls deren Verwendung im Rahmen der Studie erforderlich ist (nach Einschätzung des Prüfarztes, Anästhesisten oder Pneumologen). <p>Wesentliche Ausschlusskriterien</p> <ul style="list-style-type: none"> • Bei Screening Ateminsuffizienz, definiert als medizinische Notwendigkeit einer invasiven oder nicht-invasiven Beatmung für > 6 Stunden während eines 24-Stunden-Zeitraums. • Medizinische Notwendigkeit einer Magensonde, wenn der Großteil der Nahrungsaufnahme hierüber erfolgt (nach Einschätzung des Prüfarztes). • Schwere Kontrakturen oder schwere Skoliosen, nachweisbar im Rahmen der Röntgenuntersuchung beim Screening. • Hospitalisierung aufgrund von chirurgischen Eingriffen (d. h. Skoliose-Operation, andere chirurgische Eingriffe), pulmonalen Beschwerden oder Unterstützung der Nahrungszufuhr innerhalb von 2 Monaten des Screenings oder Planung desselben im Verlauf der Studie. • Vorliegen einer unbehandelten oder unzureichend behandelten aktiven Infektion, welche eine systemische antivirale oder antimikrobielle Therapie zu jeglichem Zeitpunkt während des Screenings erfordert. • Erkrankung des Gehirns oder Rückenmarks, einschließlich Tumore oder Anomalien in Magnetresonanz- oder Computertomographie, welche eine LP oder die Zirkulation der Rückenmarksflüssigkeit stören würde. • Vorliegen eines implantierten Shunts für die Drainage von Rückenmarksflüssigkeit oder eines implantierten ZNS-Katheters. • Vorgeschichte einer bakteriellen Meningitis. • Behandlung mit Nusinersen in einer vorangegangenen klinischen Studie. • Vorherige Verletzung (z. B. Fraktur der oberen oder unteren Extremitäten) oder chirurgische Eingriffe, von welchen der Patient sich noch nicht vollständig erholt hat bzw. noch instabil ist, welche die Fähigkeit des Patienten beeinflussen einen der für die Ergebnismessung verwendeten und im Protokoll vorgeschriebenen Tests durchzuführen. • Klinisch relevante Anomalien bei hämatologischen Parametern oder der klinischen Chemie oder EKG bei Screening, welche den Patienten nach Einschätzung des Prüfarztes ungeeignet für den Studieneinschluss machen würden. • Behandlung mit einem zur Behandlung von SMA nicht zugelassenen Wirkstoff (z. B. Salbutamol oder Salmeterol oral, Riluzol, Carnitin, Kreatin, Natriumphenylobutyrat etc.), einem biologischen Wirkstoff oder einem Medizinprodukt innerhalb von 1 Monat vor dem Screening oder 5 Halbwertszeiten des jeweiligen Wirkstoffs, je nachdem, welche Zeitspanne größer ist. Behandlung mit Valproat oder Hydroxyharnstoff innerhalb von 3 Monaten vor dem Screening. Jegliche Vorbehandlung mit Gentherapie, Antisense-Oligonukleotiden oder Zelltransplantation. • Anhaltender medizinischer Zustand, der die Durchführung und die Untersuchungen der Studie nach Einschätzung des Prüfarztes stören würde. Beispiele hierfür sind medizinische Behinderungen (z. B. Auszehrung, Kachexie oder schwere Anämie etc.), welche die Untersuchung der Sicherheit stören würde oder die Fähigkeit des Patienten einschränken würde, sich den Untersuchungen zu unterziehen. 				
Intervention und Zahl der Patienten	<p>Dosierung</p> <table border="1" data-bbox="456 1794 1366 2009"> <tbody> <tr> <td data-bbox="456 1794 906 1957"> <p>Nusinersen Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 29, 85 (Aufsättigung) und 274 (Erhaltung).</p> </td> <td data-bbox="906 1794 1366 1957"> <p>Scheinintervention unter BSC Scheinintervention in Form eines Nadelstichs am unteren Rücken (keine LP) an den Studientagen 1, 29, 85 und 274.</p> </td> </tr> <tr> <td data-bbox="456 1957 906 2009"> <p>N=84</p> </td> <td data-bbox="906 1957 1366 2009"> <p>N=42</p> </td> </tr> </tbody> </table>	<p>Nusinersen Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 29, 85 (Aufsättigung) und 274 (Erhaltung).</p>	<p>Scheinintervention unter BSC Scheinintervention in Form eines Nadelstichs am unteren Rücken (keine LP) an den Studientagen 1, 29, 85 und 274.</p>	<p>N=84</p>	<p>N=42</p>
<p>Nusinersen Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 29, 85 (Aufsättigung) und 274 (Erhaltung).</p>	<p>Scheinintervention unter BSC Scheinintervention in Form eines Nadelstichs am unteren Rücken (keine LP) an den Studientagen 1, 29, 85 und 274.</p>				
<p>N=84</p>	<p>N=42</p>				

Charakteristikum	Beschreibung
	<p>In Abhängigkeit von den institutionellen Bestimmungen konnte eine Anästhesie oder Sedierung der Patienten erfolgen. Die Behandlung mit Nusinersen sollte nur durch Ärzte erfolgen, die Erfahrung in der Durchführung von LP vorweisen können.</p> <p>Zur Sicherstellung der Verblindung wurde Nusinersen bzw. die Scheinintervention in einem separaten Raum von qualifiziertem, nicht verblindeten Studienpersonal verabreicht. Weder Eltern, Prüfarzte, Studienkoordinatoren oder Endpunkterheber waren während der Behandlung anwesend.</p>
Ort und Zeitraum der Durchführung	<p>Studiendauer Einer 28-tägigen Screeningphase schlossen sich eine 9-monatige Behandlungsphase (1, 29, 85, 274 ± 7 Tage) und eine 6-monatige Follow-up-Phase (bis Tag 456 ± 7 Tage) an, in der die Behandlung aufgrund der langen Halbwertszeit von Nusinersen fortgesetzt wurde.</p> <ul style="list-style-type: none"> • Beginn der Studienbehandlung: 24. November 2014 • Interimsanalyse: 31. August 2016 • Ende der Studie: 20. Februar 2017 • Finaler Datenschnitt: 3. März 2017 • Veröffentlichung des Studienberichts: 6. Juni 2017 <p>Ort der Durchführung Die Studie ist an 24 Zentren in 10 Ländern durchgeführt worden: Deutschland, Frankreich, Hongkong, Italien, Japan, Kanada, Schweden, Spanien, Südkorea und Vereinigte Staaten von Amerika.</p>
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkte</p> <ul style="list-style-type: none"> • Änderung in der motorischen Funktion im Vergleich zu Baseline mittels HFMSE. <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Anteil Patienten mit einer Verbesserung ≥ 3 Punkte im Vergleich zu Baseline im HFMSE nach 15 Monaten. • Patienten, die nach 15 Monaten einen neuen motorischen Meilenstein (gemäß WHO-Kriterien) erreicht haben. • Anzahl neuer motorischer Meilensteine (gemäß WHO-Kriterien) nach 15 Monaten. • Änderung nach 15 Monaten im Vergleich zu Baseline im Revised Upper Limb Module (RULM-) Test. • Anteil Patienten, die nach 15 Monaten freies Stehen erreicht haben. • Anteil Patienten, die nach 15 Monaten das Gehen mit Unterstützung erreicht haben. <p>Tertiäre Endpunkte</p> <ul style="list-style-type: none"> • Änderung der SMN-Proteinkonzentration im Liquor im Vergleich zu Baseline. • Veränderung des klinischen Gesamteindrucks (Clinical Global Impression of Change (CGI), Bewertung durch den Prüfarzt und die Eltern / den Erziehungsberechtigten). • Änderung der Lebensqualität (Pediatric Quality of Life Inventory (PedsQL)) im Vergleich zu Baseline. • Änderung der neuromuskulären Erkrankung aus Sicht der Eltern / des Erziehungsberechtigten (Assessment of Caregiver Experience with Neuromuscular Disease (ACEND)) im Vergleich zu Baseline • Krankheitsbedingte Hospitalisierungen und unerwünschte Ereignisse <p>Sicherheitsendpunkte</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse und schwerwiegende unerwünschte Ereignisse • Vitalzeichen • Neurologische Untersuchungen

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> • Körperliche Untersuchungen und Gewicht • Elektrokardiogramm • Begleitmedikation • Klinische Laboruntersuchungen (chemische Serum- und Blutuntersuchungen, Urin-Analyse) <p>Pharmakokinetikendpunkte</p> <ul style="list-style-type: none"> • Nusinersenspiegel im Liquor • Nusinersenspiegel im Plasma <p>Immunogenitätspunkt</p> <ul style="list-style-type: none"> • Nusinersenantikörper im Plasma
Subgruppenanalysen	<p>A priori geplant Folgende Subgruppenanalysen waren a priori für den primären Endpunkt und die sekundären Endpunkte geplant:</p> <ul style="list-style-type: none"> • Alter bei Screening (< 6 vs. ≥ 6 Jahre). • Geographische Region (Nordamerika vs. Europa vs. Asien-Pazifik-Raum). • Krankheitsdauer: Zeit von Diagnosestellung SMA bis Screening in Monaten nach Tertiles (< 25 vs. ≥ 25 Monate bis < 44 vs. ≥ 44 Monate). <p>Post hoc ausgewertete Subgruppen Es wurden zusätzlich post hoc definierte Subgruppenanalysen vorgenommen, für:</p> <ul style="list-style-type: none"> • Geschlecht.

Abkürzungen: BSC: Best Supportive Care; HFMSE: Hammersmith Functional Motor Scale Expanded; LP: Lumbalpunktion(en); SMA: Spinale Muskelatrophie; SMN: Survival Motor Neuron; WHO: World Health Organisation; ZNS: Zentralnervensystem

Das Original-Studienprotokoll [11] vom 17. Juli 2014 wurde drei Protokolländerungen unterzogen. Die wesentlichen für die Nutzenbewertung relevanten Protokolländerungen sind in der nachfolgenden Tabelle gelistet.

*Tabelle 5: Protokolländerungen der Studie CHERISH
(Studienpopulation: (Klein-)Kinder 2–12 Jahre alt)*

Protokollversion	Wesentliche Protokolländerungen
Amendment 1 26. September 2014	Überarbeitung der statistischen Methoden der Interimsanalyse.
Amendment 1 7. November 2014 (nur Japan betreffend)	<ul style="list-style-type: none"> • Festlegung einer Nachbeobachtungsphase mittels geplanter Studiervisite zur Nachbeobachtung der Sicherheit bei Patienten, die die Behandlung abbrechen ohne ihr Einverständnis zurückgezogen zu haben. • Ergänzung, dass eine Vollnarkose oder eine Sedierung während der LP nicht erlaubt sind. Lokale Anästhesie (z. B. Lidocain) während der LP war zulässig. • HFMSE und neurologische Beurteilung wurden als Anhänge hinzugefügt.
Amendment 1 16. Juli 2015 (nur Südkorea betreffend)	Ergänzung und Festlegen von Kriterien für die Durchführung der Interimsanalyse.
Amendment 2 22. Januar 2015 (nur Vereinigtes Königreich betreffend)	Streichen der Passage, die beschreibt, dass der Prüfarzt vor Entblindung der Behandlungszuteilung den klinischen Monitor kontaktieren sollte.

Protokollversion	Wesentliche Protokolländerungen
Amendment 2 3. Februar 2015 (nur Deutschland betreffend)	Anpassung des Ausschlusskriteriums, welches den Einsatz von Röntgen zur Erfassung einer Skoliose beschreibt. Röntgen wird nur angewendet, wenn es klinisch indiziert ist. Dies entspricht der BSC in Deutschland.
Amendment 2 30. Juni 2016	<ul style="list-style-type: none"> • Klarstellung hinsichtlich des Interimsanalyseplans, einschließlich Zeitpunkt, eingeschlossene Subpopulation, Endpunkte und Analysemethoden. • Klarstellung des Vorgehens im Falle einer frühzeitigen Beendigung der Studie basierend auf positiven Daten der Interimsanalyse. • Ergänzung über mögliche Veränderungen bzgl. der Durchführung der Studie im Falle eines positiven Nutzen-Risiko-Profiles und den Ergebnissen der Interimsanalyse. • Anpassung des Besuchsplanes für Patienten, bei denen es aufgrund einer Erkrankung zu Behandlungsverzögerungen kam. • Einzelheiten zu den Definitionen der Wirksamkeitsendpunkte wurden hinzugefügt.
Amendment 3 (nur Deutschland betreffend) und Amendment 2 (nur Japan betreffend) 14. u. 15. Juli 2016	<ul style="list-style-type: none"> • Vorgenommene Änderungen in Amendment 2 (global) sind ebenfalls in die länderspezifischen Amendments von Deutschland und Japan übernommen worden. • Korrektur des Besuchszeitraums an Tag 456 von + 7 Tage zu ± 7 Tagen.
Amendment 2 – Erratum 15. Juli 2016	Korrektur des Besuchszeitraums an Tag 456 von + 7 Tage zu ± 7 Tagen.
Amendment 2 (nur Südkorea betreffend) 25. Juli 2016	Update erfolgte auf Basis der Veränderungen in Amendment 2 (global).

Abkürzungen: BSC: Best Supportive Care; HFMSE: Hammersmith Functional Motor Scale Expanded; LP: Lumbalpunktion

*Tabelle 6: Charakterisierung der Intervention in der Studie ENDEAR
(Studienpopulation: Säuglinge ≤ 7 Monate)*

Intervention	Kontrolle
<p>Nusinersen Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 15, 29, 64, (Aufsättigung) 183 und 302 (Erhaltung), über eine 10-monatige Behandlungsphase und eine 3-monatige Follow-up-Phase (bis Tag 394 ± 7 Tage). Nusinersen wurde in einer altersadjustierten Dosis verabreicht (s. Tabelle 2). Ab einem Alter von 2 Jahren entspricht diese der empfohlenen Dosis von 12 mg (5 ml) pro Anwendung. (vgl. auch Tabelle 2)</p>	<p>Scheinintervention unter BSC Scheinintervention in Form eines Nadelstichs am unteren Rücken (keine LP) an den Studientagen 1, 15, 29, 64, 183 und 302, über eine 10-monatige Behandlungsphase und eine 3-monatige Follow-up-Phase (bis Tag 394 ± 7 Tage).</p>
<p><u>Erlaubte Begleitbehandlung</u> Während der Studie waren begleitende Medikamente oder Behandlungen erlaubt, die für unerwünschte Ereignisse als notwendig oder als angemessene unterstützende Maßnahme erachtet wurden.</p> <p><u>Unzulässige Begleitbehandlung</u> Untersagt waren andere experimentelle oder vermarktete Wirkstoffe, die für die Behandlung von SMA untersucht werden, wie z. B: orales Albuterol / Salbutamol, Riluzol, Carnitin, Kreatin, Natriumphenybutyrat, Valproat, Hydroxyharnstoff.</p>	

*Tabelle 7: Charakterisierung der Intervention in der Studie CHERISH
(Studienpopulation: (Klein-)Kinder 2–12 Jahre alt)*

Intervention	Kontrolle
<p>Nusinersen Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 29, 85 (Aufsättigung) und 274 (Erhaltung), über eine 9-monatige Behandlungsphase und eine 6-monatige Follow-up-Phase (bis Tag 456 ± 7 Tage).</p>	<p>Scheinintervention unter BSC Scheinintervention in Form eines Nadelstichs am unteren Rücken (keine LP) an den Studientagen 1, 29, 85 und 274, über eine 9-monatige Behandlungsphase und eine 6-monatige Follow-up-Phase (bis Tag 456 ± 7 Tage).</p>
<p><u>Erlaubte Begleittherapie</u> Während der Studie waren begleitend Medikamente oder Behandlungen erlaubt, die für unerwünschte Ereignisse als notwendig oder als angemessene unterstützende Maßnahme erachtet wurden.</p> <p><u>Unzulässige Begleittherapie</u> Untersagt waren andere experimentelle oder vermarktete Wirkstoffe, die für die Behandlung von SMA untersucht werden, wie z. B: orales Albuterol / Salbutamol, Riluzol, Carnitin, Kreatin, Natriumphenybutyrat, Valproat, Hydroxyharnstoff.</p>	

2.5 Endpunkte und statistische Methoden

2.5.1 In der Nutzenbewertung berücksichtigte Endpunkte

Für die vorliegende Nutzenbewertung wurden folgende vom pU als patientenrelevant eingestuft Endpunkte als bewertungsrelevant berücksichtigt:

Tabelle 8: In der Nutzenbewertung berücksichtigte Endpunkte der Studien ENDEAR und CHERISH

Endpunkt	Kategorie	Vom pU als patientenrelevant eingestuft	In der Nutzenbewertung berücksichtigt
Studie: ENDEAR (Studienpopulation: Säuglinge ≤ 7 Monate)			
Überlebensrate	Mortalität	•	•
Zeit bis zum Tod oder zur dauerhaften Beatmung ¹⁾	Mortalität/ Morbidity	•	•
Anteil der HINE-Responder (Subskala 2)	Morbidity	•	•
Anteil der CHOP INTEND-Responder		•	•
Unerwünschte Ereignisse	Sicherheit	•	•
Häufigkeit schwerwiegender respiratorischer Ereignisse		•	•
Häufigkeit und Dauer von Hospitalisierungen		•	•
Studie: CHERISH (Studienpopulation: (Klein-)Kinder 2–12 Jahre alt)			
Todesfälle ²⁾	Mortalität	•	•
Änderung in der motorischen Funktion im Vergleich zu Baseline mittels HFMSE	Morbidity	•	•
Anteil Patienten mit einer Verbesserung ≥ 3 Punkte im Vergleich zu Baseline im HFMSE nach 15 Monaten		•	•
Veränderung des PedsQL-Scores	Lebensqualität	•	•
Unerwünschte Ereignisse	Sicherheit	•	•
Häufigkeit von krankheitsbedingten Hospitalisierungen		•	•

- Endpunkt wird als bewertungsrelevant eingestuft und in der Nutzenbewertung berücksichtigt
- Endpunkt wird nicht als bewertungsrelevant eingestuft aber ergänzend in der Nutzenbewertung dargestellt
- Endpunkt wird nicht als bewertungsrelevant eingestuft

¹⁾ Die Teilkomponenten des kombinierten primären Endpunktes der ENDEAR-Studie Zeit bis zum Tod oder zur dauerhaften Beatmung werden ergänzend einzeln betrachtet.

²⁾ Todesfälle wurden im Rahmen der Sicherheitsbewertung erhoben und vom pU ergänzend ausgewertet.

Abkürzungen: CHOP INTEND: Children's Hospital of Philadelphia Infant Test for Neuromuscular Disease; HFMSE: Hammersmith Functional Motor Scale Expanded; HINE Hammersmith Infant Neurological Examination; PedsQL: Pediatric Quality of Life Inventory.

Dauerhafte Beatmung

Der kombinierte Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung (≥ 16 Stunden/Tag durchgehend für > 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses) ist einer von zwei primären Endpunkten der ENDEAR-Studie. Der kombinierte Endpunkt sowie die Teilkomponenten des Endpunktes werden nachfolgend dargestellt und als patientenrelevant angesehen. Auf die Darstellung weiterer Endpunkte zur Beatmung wird verzichtet (Anteil Patienten ohne dauerhafte Beatmung, Anzahl Stunden mit Beatmung, Zeit bis zur Beatmung – unabhängig von der Beatmungsdauer, Anzahl Patienten ohne Beatmung – unabhängig von der Beatmungsdauer), da diese bereits über die Ergebnisse des kombinierten Endpunktes sowie der Ergebnisdarstellung der Teilkomponente abgedeckt sind.

Hammersmith Infant Neurological Examination (HINE)

Der Endpunkt Anteil der HINE-Responder (Subskala 2) ist einer von zwei primären Endpunkten der ENDEAR-Studie. Darüber hinaus stellt der pU in seinem Nutzendossier den Endpunkt Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline dar. Beim HINE handelt es sich um ein Instrument zur Messung der motorischen Funktionsfähigkeit bei Säuglingen und Kleinkindern zwischen 2–24 Monaten. Es besteht aus 37 Items, verteilt auf drei Abschnitte (Subskalen): (1) einer neurologischen Untersuchung, (2) Beurteilung der motorischen Entwicklung und (3) Beurteilung des Verhaltens. In der Studie ENDEAR erfolgte die Beurteilung der Entwicklung motorischer Meilensteine anhand der Subskala 2 des HINE (Kopfkontrolle, Treten, Sitzen, bewusstes Greifen, Drehen, Krabbeln, Stehen und Gehen). Wie in Abschnitt 2.5.3 diskutiert, ist das Instrument bisher weder validiert noch ist klar, wie sich der Gesamtwert der Subskala 2 errechnet. Darüber hinaus fehlen Untersuchungen zur Reliabilität des Instrumentes und zur Validierung der gewählten Relevanzschwelle. Jedoch sind die abgebildeten, altersabhängigen Entwicklungsstufen der Subskala 2 mit der motorischen Entwicklung eines Säuglings vergleichbar. Außerdem wird die Veränderung der motorischen Funktionsfähigkeit in Form der Entwicklung motorischer Meilensteine als patientenrelevant angesehen. Infolgedessen, und weil es sich um den primären Endpunkt der ENDEAR-Studie handelt, werden die Ergebnisse des HINE (Subskala 2) nachfolgend berichtet und für die vorliegende Nutzenbewertung herangezogen. Auf die ergänzende Darstellung der Subskalen 1 und 3 des HINE wird aufgrund mangelnder Validierung des Instrumentes verzichtet.

Überlebensrate

Für die Bewertung des Gesamtüberlebens der Studienpopulation wurde der Endpunkt Zeit bis zum Tod herangezogen. Auf eine ergänzende Darstellung der Überlebensrate wird verzichtet.

Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease (CHOP INTEND)

Der CHOP INTEND ist eine Funktionsskala speziell für die Auswertung der Entwicklung motorischer Fähigkeiten von Säuglingen mit motorischer Schwäche. Der CHOP INTEND gilt als etabliertes und krankheitsspezifisches Instrument in der Patientenpopulation, da bereits in verschiedenen klinischen Studien eingesetzt, ohne jedoch final validiert zu sein. Die Ergebnisse des Endpunktes Anteil CHOP-INTEND-Responder in der ENDEAR-Studie werden unterstützend zum primären Endpunkt HINE-Responder (Subskala 2) berichtet.

Compound Muscle Action Potential (CMAP)

CMAP ist eine elektromyographische Untersuchung (elektrische Untersuchung der Muskelfunktion). Auf eine Ergebnisdarstellung dieses Endpunktes im Nutzendossier ist seitens des pU verzichtet worden. Der Einschätzung des pU wird für die vorliegende Nutzenbewertung gefolgt.

Motorische Meilensteine, Revised Upper Limb Module (RULM-) Test, freies Stehen und Gehen

In der CHERISH-Studie wurden neben dem Endpunkt Hammersmith Functional Motor Scale Expanded (HFMS-E) Responder folgende sekundäre Endpunkte definiert:

- Patienten, die nach 15 Monaten einen neuen motorischen Meilenstein erreicht haben;
- Anzahl neuer motorischer Meilensteine pro Patient nach 15 Monaten;

- Veränderung des RULM im Vergleich zu Baseline;
- Anzahl der Patienten, die freies Stehen erreicht haben;
- Anzahl der Patienten, die das Gehen mit Unterstützung erreicht haben.

Als primärer Endpunkt der CHERISH-Studie war die Veränderung des HFMSE im Vergleich zu Baseline definiert. Der HFMSE wurde entwickelt, um Veränderungen der motorischen Funktion bei Patienten mit SMA Typ II und III zu evaluieren. Das Instrument gilt als etabliert und validiert (siehe Abschnitt 2.5.3, Tabelle 12). Die motorische Entwicklung wird prinzipiell als patientenrelevant angesehen, die Studienergebnisse des HFMSE werden nachfolgend dargestellt und für die Nutzenbewertung von Nusinersen berücksichtigt. Der RULM ist eine Weiterentwicklung des ULM (Upper Limb Module) und dient der Beurteilung der Funktion der oberen Extremitäten bei nicht gehfähigen Patienten mit späterem SMA-Krankheitsbeginn. In einer Studie [19] wird auf eine Limitation des HFMSE hinsichtlich einer knappen Beurteilung der oberen Extremitäten hingewiesen. Untersuchungen zur Validität des Instruments wurden nicht identifiziert und auch im Nutzendossier finden sich diesbezüglich keine weiteren Informationen. Da sich die Krankheits-symptomatik eher proximal als distal und stärker in den Beinen als in den Armen manifestiert und der HFMSE als validiertes Instrument eingestuft wird, werden die Ergebnisse des RULM nicht dargestellt. Alle weiteren aufgelisteten Endpunkte der CHERISH-Studie sind über die Items des HFMSE abgedeckt und werden nachfolgend nicht adressiert: Patienten, die nach 15 Monaten einen neuen motorischen Meilenstein erreicht haben; Anzahl neuer motorischer Meilensteine pro Patient nach 15 Monaten; Anzahl der Patienten, die freies Stehen erreicht haben; Anzahl der Patienten, die das Gehen mit Unterstützung erreicht haben.

Assessment of Caregiver Experience with Neuromuscular Disease (ACEND)

Die Belastung der Eltern/Erziehungsberechtigten durch die Erkrankung des Kindes ist in der CHERISH-Studie als tertiärer Endpunkt anhand der Veränderung des ACEND im Vergleich zu Baseline bewertet worden. Das Instrument ist entwickelt worden, um Anstrengungen von Eltern bzw. betreuende Person durch schwerere neuromuskuläre Erkrankungen bei Kindern zu quantifizieren. Der Endpunkt wird nicht als patientenrelevant eingestuft. Auf eine Ergebnisdarstellung in der vorliegenden Nutzenbewertung wird entsprechend verzichtet.

Clinical Global Impression of Change (CGI)

Die Veränderung des klinischen Gesamteindrucks anhand CGI (Bewertung durch den Prüfarzt und die Eltern/Erziehungsberechtigten) war in der CHERISH-Studie als tertiärer Endpunkt definiert. Auf eine Ergebnisdarstellung des Endpunktes im Nutzendossier ist seitens pU verzichtet worden. Der Einschätzung des pU wird gefolgt.

2.5.2 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 9: Verzerrungspotential der Studien ENDEAR und CHERISH auf Studienebene

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patient	Verblindung Behandler	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studien-ebene
ENDEAR	ja	ja	ja	niedrig ¹⁾	nein	nein ²⁾	niedrig
CHERISH	ja	ja	ja	niedrig ¹⁾	nein	nein ²⁾	niedrig

¹⁾ Zur Sicherstellung der Verblindung wurde Nusinersen bzw. die Scheinintervention in einem separaten Raum von qualifiziertem, nicht verblindeten Studienpersonal verabreicht. Weder Eltern, Prüfarzte, Studienkoordinatoren oder Endpunkterheber waren während der Behandlung anwesend. Das für die Analyse von der Verblindung entbundene Studienpersonal war infolge nicht weiter an der Studie beteiligt.

²⁾ Durch den natürlichen Krankheitsverlauf bei SMA oder mögliche Wirksamkeitseffekte der Studienbehandlung erscheint trotz adäquater Verblindungsmaßnahmen in der Studienorganisation eine Aufrechterhaltung der Verblindung des untersuchenden Arztes, anderem beteiligten Studienpersonal und/oder der Eltern/Erziehungsberechtigten unter Umständen als nicht gegeben, was eine Verzerrung der Endpunkterhebung zur Folge haben kann.

ENDEAR (Studienpopulation: Säuglinge ≤ 7 Monate)

Die pivotale Studie ENDEAR (CS3B) ist eine multizentrische, randomisierte, kontrollierte, doppelblinde Phase-III-Studie. Die Studienteilnehmer wurden zufällig und zentral randomisiert und im Verhältnis von 2:1 den Vergleichsarmen Nusinersen und Scheinintervention zugeordnet. Die Gruppenzuteilung erfolgte verdeckt über ein Interactive Voice/Web Response System (IVRS). Die Randomisierung wurde stratifiziert für den Faktor Krankheitsdauer (Alter des Patienten bei Screening - Alter bei Symptombeginn: ≤ 12 vs. > 12 Wochen). Studienteilnehmer, Eltern/Erziehungsberechtigte, das Studienpersonal sowie der Sponsor waren für die Dauer der Studie verblindet. Zur Sicherstellung der Verblindung erfolgte die Verabreichung von Nusinersen in einem separaten Raum von ausgebildetem Studienpersonal, das nicht verblindet war. Weder die Eltern noch Prüfarzte, Studienkoordinatoren oder Endpunkterheber waren während der Behandlung anwesend. Entscheidungen in Bezug auf die Notwendigkeit von Beatmung und der Durchführung von Wirksamkeitsbewertungen wurden ebenfalls von verblindetem Studienpersonal getroffen. Für die Analyse der Zwischenergebnisse war eine Entblindung hinsichtlich der Behandlungszuteilung erforderlich. Das dafür von der Verblindung aufgehobene Studienpersonal war infolge nicht mehr an der Durchführung der Studie beteiligt. Das zusätzliche Data Safety Monitoring Board war gegenüber der Gruppenzuteilung verblindet. Eine Verzerrung seitens einer ergebnisgesteuerten Berichterstattung oder sonstiger beeinflussender Faktoren wurde nicht identifiziert, sodass das Verzerrungspotential auf Studienebene für die Studie ENDEAR insgesamt als niedrig eingestuft wird.

CHERISH (Studienpopulation: (Klein-)Kinder 2–12 Jahre alt)

In der multizentrischen, randomisierten, kontrollierten, doppelblinden Phase-III-Studie CHERISH (CS4) sind die Studienteilnehmer zufällig und zentral randomisiert und im Verhältnis von 2:1 den Vergleichsarmen Nusinersen und Scheinintervention zugeordnet worden. Die Gruppenzuteilung

erfolgte verdeckt über ein IVRS. Die Randomisierung wurde stratifiziert für das Alter des Patienten bei Screening (< 6 vs. ≥ 6 Jahre). Studienteilnehmer, Eltern/Erziehungsberechtigte, das Studienpersonal sowie der Sponsor waren für die Dauer der Studie verblindet. Zur Sicherstellung der Verblindung erfolgte die Verabreichung von Nusinersen in einem separaten Raum von ausgebildetem Studienpersonal, das nicht verblindet war. Weder die Eltern/Erziehungsberechtigten noch Prüfärzte, Studienkoordinatoren oder Endpunkterheber waren während der Behandlung anwesend. Entscheidungen in Bezug auf die Notwendigkeit von Beatmung und der Durchführung von Wirksamkeitsbewertungen wurden ebenfalls von verblindetem Studienpersonal getroffen. Für die Analyse der Zwischenergebnisse war eine Entblindung hinsichtlich der Behandlungszuteilung erforderlich. Das dafür von der Verblindung aufgehobene Studienpersonal war infolge nicht mehr an der Durchführung der Studie beteiligt. Das zusätzliche Data Safety Monitoring Board war gegenüber der Gruppenzuteilung verblindet. Eine Verzerrung seitens einer ergebnisgesteuerten Berichtserstattung oder sonstiger beeinflussender Faktoren wurde nicht identifiziert. Das Verzerrungspotential auf Studienebene für die Studie CHERISH wird insgesamt als niedrig eingestuft.

Tabelle 10: Verzerrungspotential der in der Nutzenbewertung berücksichtigten Endpunkte der Studien ENDEAR und CHERISH

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Mortalität					
<i>ENDEAR (Studienpopulation: Säuglinge ≤ 7 Monate)</i>					
Überlebensrate	ja	ja	nein	nein	niedrig
Morbidität					
<i>ENDEAR (Studienpopulation: Säuglinge ≤ 7 Monate)</i>					
Zeit bis zum Tod oder zur dauerhaften Beatmung ¹⁾	ja	ja	nein	unklar ²⁾³⁾	niedrig
Anteil der HINE-Responder (Subskala 2)	ja	ja	nein	unklar ³⁾⁴⁾	unklar
Anteil der CHOP INTEND-Responder	ja	ja	nein	unklar ³⁾⁴⁾	unklar
<i>CHERISH (Studienpopulation: (Klein-)Kinder 2–12 Jahre alt)</i>					
HFMSE-Responder ⁵⁾	ja	ja	nein	unklar ³⁾	niedrig
Lebensqualität					
<i>CHERISH (Studienpopulation: (Klein-)Kinder 2–12 Jahre alt)</i>					
Veränderung des PedsQL-Scores	ja	ja	nein	unklar ³⁾	niedrig

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Sicherheit					
<i>ENDEAR (Studienpopulation: Säuglinge ≤ 7 Monate)</i>					
Unerwünschte Ereignisse	ja	ja	ja	unklar ³⁾	niedrig
Schwerwiegende respiratorische Ereignisse	ja	ja	ja	unklar ³⁾	niedrig
Häufigkeit von Hospitalisierungen	ja	ja	ja	unklar ³⁾	niedrig
<i>CHERISH (Studienpopulation: (Klein-)Kinder 2–12 Jahre alt)</i>					
Unerwünschte Ereignisse	ja	ja	ja	unklar ³⁾	niedrig
Häufigkeit von krankheitsbedingten Hospitalisierungen	ja	ja	ja	unklar ³⁾	niedrig

¹⁾ Die Teilkomponenten des kombinierten primären Endpunktes der ENDEAR-Studie Zeit bis zum Tod oder zur dauerhaften Beatmung werden ergänzend einzeln betrachtet.

²⁾ Die tägliche Dokumentation der Beatmung erfolgte durch die Eltern/Erziehungsberechtigten der Studienteilnehmer. Die Informationen wurden bei jeder Visite und bei wöchentlichen Telefonkontakten eingeholt.

³⁾ Durch den natürlichen Krankheitsverlauf bei SMA oder mögliche Wirksamkeitseffekte der Studienbehandlung erscheint trotz adäquater Verblindungsmaßnahmen in der Studienorganisation eine Aufrechterhaltung der Verblindung des untersuchenden Arztes, anderem beteiligten Studienpersonal und/oder der Eltern/Erziehungsberechtigten unter Umständen als nicht gegeben, was eine Verzerrung der Endpunkterhebung zur Folge haben kann.

⁴⁾ Das Erhebungsinstrument ist unzureichend validiert (vgl. 2.5.3).

⁵⁾ In der CHERISH-Studie wurden mittels HFMSE folgende Morbiditätsendpunkte erhoben: Änderung in der motorischen Funktion im Vergleich zu Baseline mittels HFMSE und Anteil Patienten mit einer Verbesserung ≥ 3 Punkte im Vergleich zu Baseline im HFMSE nach 15 Monaten.

Abkürzungen: CHOP INTEND: Children's Hospital of Philadelphia Infant Test for Neuromuscular Disease; ITT: Intention-to-Treat; HFMSE: Hammersmith Functional Motor Scale Expanded; HINE Hammersmith Infant Neurological Examination; PedsQL: Pediatric Quality of Life Inventory.

2.5.3 Charakterisierung der berücksichtigten Endpunkte und Einschätzung der Patientenrelevanz

Laut SGB V § 35a Abs. 1 Satz 10 gilt der medizinische Zusatznutzen durch die Zulassung als belegt. Nachweise gemäß § 5 Abs. 1 bis 6 sind nicht verpflichtend vorzulegen. Zur Beurteilung der Patientenrelevanz der vorgelegten Endpunkte wurden das Nutzendossier, der zugehörige Studienbericht, das Studienprotokoll, der statistische Studienplan und die in diesen Quellen zitierte Literatur herangezogen. Das Ergebnis der Betrachtung der Endpunkte findet sich aufgeteilt nach Endpunktkategorien in den Tabellen 11 bis 14.

Tabelle 11: Beschreibung der Endpunkte der Kategorie Mortalität in der Studie ENDEAR

Mortalität	
Endpunkt	Studie ENDEAR: Überleben
Operationalisierung	<p>Es wurde die Zeit bis zum Tod von der Verabreichung der ersten Dosis der Studienmedikation oder Randomisierung bis zum Eintritt des Todes erfasst.</p> <p>Erhebungszeitpunkte: Die Erhebung des Endpunktes schloss den gesamten Studienzeitraum ein.</p> <p>Zensurierung: Eine Zensurierung erfolgte am Tag des letzten Follow-up-Termins für Studienteilnehmer, die am Leben sind. Beginn der Berechnung der Zeit bis zum Tod oder des Tages der Zensur ist das Datum der ersten Dosis der Studienmedikation oder, wenn unvollständig, das Datum der Randomisierung. Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunktes ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz: Mortalität gilt als valider Endpunkt. Gemäß § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (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.“ Die Endpunktkomponenten des primären Endpunktes Zeit bis zum Tod und Zeit bis zur dauerhaften Beatmung werden als patientenrelevant eingestuft und für die Bewertung des Zusatznutzens von Nusinersen berücksichtigt. Aufgrund von qualitativer und quantitativer Heterogenität sowie möglicher Verzerrungseffekte des kombinierten Endpunktes werden die Endpunktkomponenten zusätzlich einzeln betrachtet.</p>

Tabelle 12: Beschreibung der Endpunkte der Kategorie Morbidität in den Studien ENDEAR und CHERISH

Morbidität	
Endpunkt	Studie ENDEAR: Zeit bis zum Tod oder zur dauerhaften Beatmung
Operationalisierung	<p>Der primäre Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung ist ein kombinierter Endpunkt. Die dauerhafte Beatmung ist definiert als Beatmung von ≥ 16 Stunden/Tag durchgehend für > 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses oder Tracheotomie.</p> <p><u>Akutes, reversibles Ereignis</u> Sofern sich eines oder mehrere der folgenden Kriterien für ein aktuelles, reversibles Ereignis zwischen 7 Tage vor und 7 Tage nach dem Beginn des Schwellenwertes von ≥ 16 Stunden dauerhafter Beatmung pro Tag ereignete, wurde eine 14-tägige Schonfrist eingeleitet, um dem Patient Zeit zu geben sich von dem akuten reversiblen Ereignis zu erholen. Die 14-tägige Schonfrist verlängerte sich nicht durch Eintritt weiterer Kriterien. Sobald der Schwellenwert von > 21 aufeinanderfolgenden Tagen erreicht wird, ist der Bedarf an einer Beatmung aufgrund der Progression der SMA-Erkrankung gegeben und der Endpunkt ist erfüllt. Bei Eintritt und Dokumentation eines der folgenden Kriterien gilt der Endpunkt als nicht erfüllt. Ein akutes, reversibles Ereignis war 7 Tage vor und 7 Tage nach Erreichen einer dauerhaften Beatmung (≥ 16 Stunden pro Tag) durch folgende Kriterien definiert:</p>

Morbidität	
	<ul style="list-style-type: none"> • Fieber von 38,9° C (Messung im Ohr, rektal, axillär, auf der Haut, sublingual). • Infektion <ul style="list-style-type: none"> ○ Blut-, Sputum-, Rachen- oder CSF-Kultur positiv für Viren, Bakterien oder Pilze; ○ Blut-, Rachen-, Sputum- oder CSF-Virus-PCR-positiv; ○ Infektiöse Antigen-Diagnostik (z. B. strep + oder HepB Ag +) im Blut, Rachen, Sputum oder CSF; ○ Positive mikroskopische Visualisierung (z. B. bronchoalveoläre Spülung oder Gramfärbung in Gewebebiopsie) im Blut, Rachen, Sputum oder CSF. • Chirurgische Eingriffe <ul style="list-style-type: none"> ○ Operation (d. h. Magensonde, orthopädische Eingriffe); ○ Jeder Eingriff, der eine lokale Anästhesie oder Vollnarkose erfordert. <p>Beurteilt wurde der Endpunkt durch ein verblindetes, zentrales und unabhängiges Entscheidungskomitee (Endpoint Adjudication Committee).</p> <p>Erhebungszeitpunkte: Die Beatmungsdauer wurde täglich mittels Tagebuch erhoben. Fehlende Daten wurden mit denjenigen Daten eines zeitlich nahen Tages mit der längsten Beatmungsdauer imputiert. Mit Beginn des Tagebuchs musste dieses bis Studienende täglich geführt werden.</p> <p>Zensierung: Patienten wurden zum Zeitpunkt des letzten Eintrages ins Tagebuch zensiert. Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunkts ist nachvollziehbar beschrieben. Es handelt sich um einen kombinierten Endpunkt, der sich aus Mortalität und verschiedenen Endpunkten der Kategorie Morbidität zusammensetzt.</p> <p>Validität und Patientenrelevanz: Die Beatmungspflichtigkeit wird als patientenrelevant angesehen, die Zeit bis zur dauerhaften Beatmung gilt entsprechend als valider Endpunkt. Mortalität gilt 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.“ Der primäre Endpunkt wird als patientenrelevant eingestuft und für die Bewertung des Zusatznutzens von Nusinersen berücksichtigt. Aufgrund von qualitativer und quantitativer Heterogenität sowie möglicher Verzerrungseffekte des kombinierten Endpunktes werden die Teilkomponenten des Endpunktes zusätzlich einzeln betrachtet.</p>
Endpunkt	Studie ENDEAR: Hammersmith Infant Neurological Examination (HINE) (Subskala 2)
Operationalisierung	<p>Die Entwicklung motorischer Meilensteine wurde mittels Subskala 2 des dreiteiligen Messinstruments HINE erhoben. Die Subskala 2 besteht aus 8 Kategorien, die verschiedene motorische Meilensteine abbilden: Bewusstes Greifen, Fähigkeit in Rückenlage zu treten, Kopfkontrolle, Drehen, Sitzen, Krabbeln, Stehen und Gehen. Innerhalb jedes motorischen Meilensteins lassen sich von links nach rechts 3–5 Entwicklungsstufen abbilden. In einzelnen Spalten ist auch das Alter beschrieben, in dem dieser besondere Meilenstein bei normalen Säuglingen erreicht wird. Die 8 Kategorien wurden bei jeder Untersuchung gemessen.</p>

Morbidity					
HINE (Subskala 2) – Motorische Meilensteine [24]					
Motorischer Meilenstein	Verbesserung ----->				
Bewusstes Greifen	Kein Greifen	Verwendet die ganze Hand	Benutzt Zeigefinger und Daumen, unvollständiges Greifen	Pinzettengriff	
Fähigkeit zu treten (in Supination)	Kein Treten	Horizontales Treten, Beine werden nicht gehoben	Treten nach oben (vertikal) (3 Monate)	Berührung der Beine (4–5 Monate)	Berührung der Zehen (5–6 Monate)
Kopfkontrolle	Keine aufrechte Haltung (< 3 Monate)	Kopf wackelt (4 Monate)	Dauerhaft aufrecht (5 Monate)		
Drehen	Kein Drehen	Drehen zur Seite (4 Monate)	Vom Bauch auf den Rücken (6 Monate)	Vom Rücken auf den Bauch (7 Monate)	
Sitzen	Kein Sitzen	Sitzen mit Unterstützung an den Hüften (4 Monate)	Sitzen mit Unterstützung der Arme (6 Monate)	Stabiles Sitzen (7 Monate)	Dreht sich (Rotation) (10 Monate)
Krabbeln	Kopf wird nicht angehoben	Auf den Ellbogen (3 Monate)	Auf ausgestreckten Händen (4–5 Monate)	Krabbelt flach auf dem Bauch (8 Monate)	Auf Händen und Knien (10 Monate)
Stehen	Kann das eigene Gewicht nicht halten	Kann das eigene Gewicht halten (4–5 Monate)	Stehen mit Hilfe (8 Monate)	Stehen ohne Hilfe (12 Monate)	
Gehen	Kein Gehen	Federnd gehen (6 Monate)	Herumgehen (mit Festhalten) (11 Monate)	Unabhängiges Gehen (15 Monate)	
<p>Jedes Item wird individuell bewertet und ein Gesamtscore (maximal 26 Punkte) aus der Summe der individuellen Werte zusammenaddiert.</p> <p>Erhebungszeitpunkte: Die Erhebung der motorischen Meilensteine mittels HINE erfolgte zum Zeitpunkt des Screenings (Baseline) sowie an den Tagen 64, 183, 302, 394. Fehlende Daten wurden imputiert.</p> <p>Responsekriterium: Ein Responder war wie folgt definiert: a) Der Patient zeigte eine Verbesserung um mindestens 2 Punkte in der Kategorie „Fähigkeit zu treten“ oder erreichte in dieser Kategorie den Höchstwert (Zehen berühren) oder der Patient zeigte eine Verbesserung um einen Punkt in der Kategorie Kopfkontrolle, Drehen, Sitzen, Krabbeln, Stehen oder Gehen UND b) Der Patient zeigte in mehr Items (Ausnahme: bewusstes Greifen) eine Verbesserung (wie in a) definiert) als eine Verschlechterung. Für die Kategorie „Fähigkeit zu treten“ galt hierbei analog zu a), dass eine Abnahme um mindestens 2 Punkte oder auf den Wert 0 erfolgen musste, damit dies als Verschlechterung gilt.</p> <p>Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>					

Morbidität	
Bewertung	<p>Operationalisierung: Beim HINE handelt es sich um ein Instrument zur Messung der motorischen Funktionsfähigkeit bei Säuglingen und Kleinkindern zwischen 2–24 Monaten. Es besteht aus 37 Items, verteilt auf drei Abschnitte: (1) einer neurologischen Untersuchung, (2) Beurteilung der motorischen Entwicklung und (3) Beurteilung des Verhaltens. In der Studie ENDEAR erfolgte die Beurteilung der Entwicklung motorischer Meilensteine anhand der Subskala 2 des HINE (Kopfkontrolle, Treten, Sitzen, bewusstes Greifen, Drehen, Krabbeln, Stehen und Gehen).</p> <p>Validität und Patientenrelevanz: Die Inhalte des HINE basieren auf empirischen Kriterien in einer Population von Ein- bis Eineinhalbjährigen ohne neurologische Auffälligkeiten im Säuglingsalter. Nähere Informationen sowohl zum Erstellungsprozess als auch zum Aufbau des Instrumentes fehlen, erscheinen aber auf die normale motorische Entwicklung eines Säuglings übertragbar [24,36,37]. Mit der Subskala 2 des HINE wird die neurologische Untersuchung eines Säuglings normiert und dokumentiert, es können der Status Quo sowie eine motorische Entwicklung erfasst werden. Bislang liegen keine Ergebnisse zur Reliabilität des HINE in einer SMA-Population vor [24,33]. Die Items sind altersabhängig, sodass von einer patientenindividuellen Variabilität hinsichtlich der Funktionsfähigkeit auszugehen ist. Eine Bewertung mittels Punkteskala ist nicht nachvollziehbar, wobei die Zusammensetzung des Gesamtscores von maximal 26 Punkten sowie dessen Interpretation nicht detailliert dargelegt werden. Ergebnisse zur Validierung des Instrumentes liegen bislang nicht vor [24,33]. In einer retrospektiven, multizentrischen Studie wurde die Entwicklung von motorischen Meilensteinen bei SMA Typ I mittels HINE untersucht [14]. Die Neugeborenen erreichten nur in drei Items mehr als 0 Punkte – Kopfkontrolle, Fähigkeit zu treten und Greifen –, aber nie mehr als 2 Punkte. Unter Berücksichtigung des natürlichen Krankheitsverlaufs bei bestätigter SMA Typ I ist jegliche motorische Veränderung im Sinne einer verbesserten Mobilität als Meilenstein und somit als relevant anzusehen [14]. Das vom pU definierte Responsekriterium erscheint nachvollziehbar, jedoch fehlen unterstützende Untersuchungen zur Minimal Clinical Important Difference (MCID). Unklar bleibt, ob eine dauerhafte Response erforderlich war oder ein einmaliges Erreichen des Responsekriteriums ausreichend war, um als Responder gezählt zu werden.</p> <p>Das Erreichen eines motorischen Meilensteins im Sinne einer verbesserten Mobilität kann insgesamt als patientenrelevant erachtet werden. Darüber hinaus ist das definierte Responsekriterium unter Berücksichtigung fehlender Daten zur Validität des Instruments und zur MCID zu betrachten.</p> <p>Der primäre Endpunkt wird nachfolgend dargestellt und für die Bewertung des Zusatznutzens von Nusinersen berücksichtigt.</p>
Endpunkt	Studie ENDEAR: Children’s Hospital of Philadelphia Infant Test of Neuromuscular Disease (CHOP INTEND)
Operationalisierung	Motorische Fähigkeiten wurden mit der Funktionsskala CHOP INTEND erhoben. Der CHOP INTEND wurde speziell für die Auswertung der Entwicklung der Motorik von Säuglingen mit motorischer Schwäche, einschließlich Säuglingen mit SMA, entwickelt. 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 erreichbar ist. Beurteilt werden entweder spontane Bewegungen (durch Beobachtung) oder aktive Bewegungen: (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.

Morbidität	
Morbidität	<p>Erhebungszeitpunkte: Die Erhebung des CHOP INTEND erfolgte zum Zeitpunkt des Screenings (2 Messungen zu Baseline) sowie an den Tagen 64, 183, 302, 394 jeweils vor Dosierung (z. B. durch einen Physiotherapeuten). Fehlende Daten wurden imputiert.</p> <p>Responsekriterium: Ein Responder ist definiert als eine Verbesserung im Vergleich zu Baseline um mindestens 4 Punkte. Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunkts ist nachvollziehbar. Der CHOP INTEND ist speziell zur Beurteilung motorischer Fähigkeiten bei Säuglingen mit SMA Typ I entwickelt worden.</p> <p>Validität und Patientenrelevanz: Der CHOP INTEND ist ein von Experten entwickeltes, etabliertes und krankheitsspezifisches Instrument. In mehreren Studien konnten gute Testgütekriterien nachgewiesen werden, wobei die Validität weiter zu etablieren gilt. Untersuchungen zur MCID des Instrumentes existieren bislang nicht [10,19,20,22,23]. Unter Berücksichtigung des natürlichen Krankheitsverlaufs bei Patienten mit bestätigter SMA Typ I, bei denen eine Verbesserung des aktuellen Gesundheitszustandes nicht zu erwarten ist, erscheint jeglicher Fortschritt motorischer Fähigkeiten als erstrebenswert und patientenrelevant. Seitens des pU wird eine Responseschwelle von ≥ 4 Punkten als klinisch relevante Veränderung angegeben und damit begründet, dass dieser Wert außerhalb der Testvariabilität liege. Die Herleitung dieser Schwelle ist jedoch nicht vollständig nachvollziehbar beschrieben [20,22,27]. Der CHOP INTEND gilt als etabliertes Instrument in der Patientenpopulation, da bereits in verschiedenen klinischen Studien eingesetzt, ohne jedoch final validiert zu sein. Die Ergebnisse des CHOP INTEND werden unterstützend zum primären Endpunkt berichtet.</p>
Endpunkt	Studie CHERISH: Hammersmith Functional Motor Scale Expanded (HF MSE)
Operationalisierung	<p>Die Änderung der motorischen Funktion im Vergleich zu Baseline wurde mittels HF MSE erhoben.</p> <p>Der HF MSE ist ein reliables und validiertes Instrument und wurde speziell für die Evaluierung von Veränderungen der motorischen Funktion und klinischer Progression bei Kindern mit SMA Typ II und III entwickelt. Die erweiterte Skala HF MSE setzt sich aus der Hammersmith Function Motor Scale (HFMS) – bestehend aus 20 Items – und 13 zusätzlichen Items zusammen (insgesamt 33 Items), um die Bewertung von ambulanten SMA-Patienten zu ermöglichen.</p> <p>Jedes Item wird auf einer 3-Punkte-Skala bewertet: 0 (unfähig), 1 (ist mit Hilfestellung fähig) und 2 (fähig). Eine maximale Gesamtpunktzahl von 66 Punkten ist möglich, wobei eine höhere Punktzahl einen besseren motorischen Funktionsstatus beschreibt.</p> <p>Die 33 Items umfassen (1) Sitzen auf Sockel/Stuhl; (2) Langsitz; (3) Eine Hand zum Kopf im Sitzen; (4) Zwei Hände zum Kopf im Sitzen; (5) Von Rückenlage in Seitenlage; (6) Bauchlage in Rückenlage, Rolle über rechts; (7) Von Bauchlage in Rückenlage, Drehung nach links; (8) Von Rückenlage in Bauchlage, Drehung nach rechts; (9) Von Rückenlage in Bauchlage, Drehung nach links; (10) Vom Sitzen ins Liegen; (11) Aufstützen auf Unterarmen; (12) Hebt Kopf aus Bauchlage; (13) Aufstützen auf ausgestreckten Armen; (14) Vom Liegen ins Sitzen; (15) Vierfüßlerstand; (16) Krabbeln; (17) Hebt Kopf aus Rückenlage; (18) Stehen mit Unterstützung; (19) Freies Stehen; (20) Gehen; (21) Hüftflexion rechts in Rückenlage; (22) Hüftflexion links in Rückenlage; (23) Vom Kniestand in halben Kniestand rechts; (24) Vom Kniestand in halben Kniestand links; (25) Vom Kniestand in den Stand mit linkem Bein zuerst (über halben Kniestand); (26) Vom Kniestand in den Stand mit rechtem Bein zuerst (über halben Knie-</p>

Morbidität	<p>stand links); (27) Vom Stand in sitzende Position auf dem Fußboden; (28) Hockstand; (29) Springt 30 cm nach vorne; (30) 4 Stufen steigen mit Handlauf; (31) 4 Stufen hinuntergehen mit Handlauf; (32) Steigt 4 Stufen ohne Armunterstützung hoch; (33) Geht 4 Stufen ohne Armunterstützung hinunter.</p> <p>Erhebungszeitpunkte: Die Erhebung des HFMSE erfolgte zum Zeitpunkt des Screenings (2 Messungen zu Baseline) sowie an den Tagen 92, 169, 274, 365, 456. Fehlende Daten wurden imputiert.</p> <p>Responsekriterium: Sekundärer Endpunkt der Studie CHERISH ist eine Responderanalyse des HFMSE. Ein Responder ist dabei definiert als eine Veränderung des HFMSE-Baselinewertes bis Monat 15 um ≥ 3 Punkte. Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunkts ist nachvollziehbar. Der HFMSE dient der Evaluierung von Veränderungen der motorischen Funktion bei SMA Typ II und III.</p> <p>Validität und Patientenrelevanz: In der Entwicklung des krankheitsspezifischen Instrumentes HFMSE fanden sich mehrfach Expertengruppen (bestehend aus Medizinerinnen und Physiotherapeuten) zusammen und es wurden insgesamt drei Pilotierungsphasen durchgeführt [28,32]. In verschiedenen Untersuchungen zu psychometrischen Eigenschaften des HFMSE zeigte sich das Instrument reliabel und valide. Die Patienten waren im Durchschnitt älter als 2 Jahre mit spätem SMA-Krankheitsbeginn in ambulanter oder stationärer Betreuung [23,31]. Untersuchungen zur MCID des HFMSE fehlen bisher [21]. Der pU gibt eine Veränderung des HFMSE von ≥ 3 Punkten im Vergleich zu Baseline als klinisch relevant an. Er zitiert einen klinischen RCT zur stationären Behandlung von Kindern mit SMA (2–8 Jahre) mit Valproinsäure. Untersucht wurde eine mögliche Verbesserung der motorischen Funktion nach 6 Monaten mittels modifizierter HFMS. Ein statistisch signifikantes Ergebnis und Begründung für eine MCID von ≥ 3 Punkten lassen sich dort nicht identifizieren [38]. In einer weiteren Studie zur Krankheitsprogression bei ambulanten und nicht-ambulanten SMA-Patienten (Typ II und III) erreichten die meisten Patienten (> 75 %) nach 12 Monaten eine Veränderung der HFMSE-Punktzahl zu Baseline um ± 2 Punkte. Weniger als 10 % der Patienten zeigten eine Verbesserung von mehr als 2 Punkten. Weitere Untersuchungen sind notwendig um ± 2 Punkte als klinisch relevante Veränderung des HFMSE zu bestätigen [19,30]. Die Verbesserung/Verschlechterung der körperlichen Funktionalität, (z. B. die motorische Entwicklung bei Kindern) ist patientenrelevant. Mittels HFMSE sind der motorische Funktionsstatus und Fortschritt bei späten Störungen der Muskelfunktion, nach zunächst mehr oder weniger normaler Entwicklung, dokumentier- und abbildbar. Das Instrument ist für Patienten mit SMA Typ II und III validiert. Eine MCID ist bisher nicht bestätigt. Der Endpunkt wird gemäß § 2 Satz 3 AM-NutzenV als patientenrelevant eingestuft und in für die Nutzenbewertung von Nusinersen herangezogen.</p>
Endpunkt	Studie ENDEAR: Häufigkeit schwerwiegender respiratorischer Ereignisse
Operationalisierung	<p>Unerwünschte Ereignisse, die den Systemorganklassen (SOC) Erkrankungen der Atemwege, des Brustraums und Mediastinums als primäre oder sekundäre SOC zugeordnet und als schwerwiegend eingestuft wurden, werden als Ereignisse gewertet. Seltene Ereignisse der SOC Infektionen und parasitäre Erkrankungen sowie Untersuchungen (falls vorhanden) werden anhand dieses Vorgehens nicht berücksichtigt. Um diese Ereignisse zu identifizieren, ist eine detaillierte Bewertung der unerwünschten Ereignisse durchgeführt worden.</p>

Morbidität	
	<p>Erhebungszeitpunkte: Schwerwiegende respiratorische Ereignisse wurden bei jeder Studienvisite (vor Dosierung, während Dosierung und nach Dosierung) sowie während des gesamten Studienzeitraums erhoben. Fehlende Werte wurden imputiert. Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunktes ist nachvollziehbar. Es bleibt jedoch offen, ob die Ereignisse mit dem Fortschreiten der Grunderkrankung einhergehen oder weil eine gesonderte Analyse nicht beschrieben wird.</p> <p>Validität und Patientenrelevanz: Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von <u>Nebenwirkungen</u> oder einer Verbesserung der Lebensqualität.“ Der Endpunkt Schwerwiegende respiratorische Ereignisse wird im Sinne der Krankheits-symptomatik und -schwere als patientenrelevant angesehen.</p>
Endpunkt	
	<p>Studie ENDEAR: Häufigkeit von Hospitalisierungen</p>
Operationalisierung	<p>Hospitalisierung ist definiert als die Aufnahme in eine medizinische Einrichtung von mehr als 24 Stunden; ein unerwünschtes Ereignis nicht zwingend vorausgesetzt. Der Endpunkt erfasst die Anzahl und Dauer von Krankenhausaufenthalten je nach Auftreten des Ereignisses. Dabei wird die aggregierte Hospitalisierungsrate zugrunde gelegt. Es wird die Gesamtdauer im Krankenhaus während der Studie pro Studienteilnehmer berechnet und anteilmäßig auf die Studientage verteilt. Die Bewertung der Ereignisse erfolgte mittels medizinischer Untersuchung durch ein verblindetes Komitee.</p> <p>Erhebungszeitpunkte: Der Endpunkt wurde über den gesamten Studienzeitraum erfasst. Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunktes ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz: Der Endpunkt wurde mittels medizinischer Untersuchung durch ein verblindetes Komitee beurteilt. Darüber hinaus führt der pU aus, dass regionale Unterschiede ausgeschlossen werden können und somit kein Hinweis auf eine Effektmodifikation bestehe. Gemäß § 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.“ Entsprechend ist der Endpunkt als patientenrelevant einzustufen und wird für die Bewertung des Zusatznutzens von Nusinersen berücksichtigt.</p>

Morbidität	
Endpunkt	Studie CHERISH: Krankheitsbedingte Hospitalisierungen und unerwünschte Ereignisse
Operationalisierung	<p>Hospitalisierung ist definiert als die Aufnahme in eine medizinische Einrichtung von mehr als 24 Stunden; ein unerwünschtes Ereignis nicht zwingend vorausgesetzt. Um unerwünschte Ereignisse als krankheitsbedingt zu definieren erfolgte eine medizinische Bewertung durch ein verblindetes Komitee. Dabei wird erfasst, ob ein Ereignis schwerwiegend einzuordnen ist, denn als schwerwiegend gelten Ereignisse, die einen stationären Krankenhausaufenthalt erfordern oder den Krankenhausaufenthalt verlängern.</p> <p>Erhebungszeitpunkte: Das Auftreten von Hospitalisierungen wurde im Rahmen des Safety-Monitorings über die gesamte Studiendauer erfasst. Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunktes ist nachvollziehbar beschrieben. Dargestellt wird der Endpunkt Häufigkeit von krankheitsbedingten Hospitalisierungen.</p> <p>Validität und Patientenrelevanz: Der Endpunkt wurde im Rahmen des Sicherheits-Monitorings durch ein verblindetes Komitee beurteilt. Darüber hinaus führt der pU aus, dass regionale Unterschiede ausgeschlossen werden können und somit kein Hinweis auf eine Effektmodifikation bestehe. 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 Häufigkeit von krankheitsbedingte Hospitalisierungen als patientenrelevant einzustufen und wird für die Bewertung des Zusatznutzens von Nusinersen berücksichtigt.</p>

Tabelle 13: Beschreibung der Endpunkte der Kategorie Lebensqualität in der Studie CHERISH

Mortalität	
Endpunkt	Studie CHERISH: Veränderung des Pediatric Quality of Life Inventory (PedsQL-) Scores
Operationalisierung	<p>Die gesundheitsbezogene Lebensqualität wurde mittels der Generic-Core-Skala und dem Neuromuskulären Modul 3.0 des PedsQL (Version 4.0) erhoben.</p> <p>Der PedsQL besteht aus einer kurzen, generischen Hauptskala und konditionsspezifischen Modulen, die für den Einsatz in bestimmten klinischen Populationen entwickelt wurden. Der pädiatrische Fragebogen ist altersspezifisch und wird bei Kindern und Jugendlichen im Alter von 5 bis 18 Jahren eingesetzt, ein Elternfragebogen zur Beurteilung der gesundheitsbezogenen Lebensqualität des Kindes für Kinder und Jugendliche im Alter von 2 bis 18 Jahren. Die generische Skala des PedsQL 4.0 beinhaltet die Beurteilung der physischen Funktion (8 Items), von Emotionen (5 Items), der sozialen Funktion (5 Items) und der Schulfähigkeit (5 Items). Das Neuromuskuläre Modul 3.0 des PedsQL wurde entwickelt, um Dimensionen der Lebensqualität speziell für Kinder mit neuromuskulären Erkrankungen im Alter von 2 bis 18 Jahren, einschließlich Kinder mit SMA, abzubilden. Das Modul beinhaltet 3 Dimensionen die sich über 25 Items verteilen, in der Altersklasse 5–7 Jahre nur 17 Items. Beurteilt werden die neuromuskuläre Erkrankung des Kindes (17 Items zu Krankheitsfortschritt und assoziierten Symptomen), Kommunikation (3 Items hinsichtlich Kommunikationsfähigkeit der Patienten mit Ärzten und anderen zur Erkrankung) und Möglichkeiten der Familie (5 Items zu finanzieller und sozialer Unterstützung).</p>

Mortalität	
	<p>In der CHERISH-Studie wurden Fragebögen für folgende Altersgruppen erhoben: 2–4 Jahre (für diese Altersklasse ist ausschließlich ein Elternfragebogen vorhanden), 5–7 Jahre, 8–12 Jahre und 13–18 Jahre.</p> <p>Die Bewertung erfolgt auf einer 5-Punkte-Likert-Skala: 0 = nie ein Problem, 1 = fast nie ein Problem, 2 = manchmal ein Problem, 3 = häufig ein Problem, 4 = fast immer ein Problem. Die Items werden entgegengesetzt bewertet und linear auf einer Skala von 0 bis 100 transformiert, hierbei entspricht 0 = 100, 1 = 75, 2 = 50, 3 = 25, 4 = 0. Der Gesamtscore ergibt sich aus der Summe der Items geteilt durch die Anzahl der beantworteten Items. Ein höherer Gesamtscore weist auf eine bessere Lebensqualität hin. Sofern mehr als 50 % der Items einer Dimension nicht bewertet wurden, erfolgt keine Berechnung. Andernfalls wird der Mittelwert einer Dimension aus der Summe aller bewerteten Items ermittelt.</p> <p>Erhebungszeitpunkte: Die Erhebung des PedsQL erfolgte zum Zeitpunkt des Screenings (2 Messungen, Baseline) sowie an den Tagen 92, 169, 274, 365, 456. Fehlende Daten wurden imputiert.</p> <p>Die Auswertung des Endpunktes erfolgte auf Basis der ITT-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunktes ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz: Der PedsQL (http://www.pedsq.org/about_pedsq.html) ist ein etabliertes Instrument und wurde in den 1990er Jahren als generisches Instrument zur Erfassung der Lebensqualität bei pädiatrischen Populationen mit chronischen Erkrankungen entwickelt. Seither wurden für zahlreiche Erkrankungen spezifische Module entwickelt. Der Fragebogen wird sowohl von Eltern/Erziehungsberechtigten als auch von Kindern ausgefüllt. Mit der Generic-Core-Skala des PedsQL 4.0 werden bezugnehmend auf die letzten 4 Wochen die Dimensionen physische Funktion, Emotionen, soziale Funktion und Schulfähigkeit erfasst. Die interne Validität wurde an einer Stichprobe mit ca. 5.500 Kindern (ca. 570 davon chronisch krank) untersucht. Es konnte gezeigt werden, dass Kinder mit chronischen Erkrankungen konsistent eine niedrigere Lebensqualität berichten als gesunde Kinder. Die interne Konsistenz wurde an einer umfangreichen Stichprobe mit 8.836 gesunden und 847 chronisch kranken Kindern untersucht (Kinderfragebogen Cronbachs Alpha 0,89; Elternfragebogen Cronbachs Alpha 0,92). Die Test-Retest-Reliabilität bei Kindern (7–12 Jahre) wird mit 0,86 angegeben. Die Skala wurde zudem von der American Spinal Muscular Atrophy Randomized Trials (AmSMART-) Gruppe validiert und ist veränderungssensitiv. In einer Patientenpopulation mit SMA wurde für die Generic-Core-Skala des PedsQL 4.0 eine moderate bis gute Test-Retest-Reliabilität mit einem Interkorelationskoeffizienten (ICC) von 0,60 (Elternfragebogen) und einem ICC von 0,84 (Kinderfragebogen, selbstberichtet) erreicht [25]. Die MCID wird mit 4,4 (Änderung im Gesamtscore für selbstberichtete Ergebnisse) bzw. mit 4,5 (Elternfragebogen) angegeben [40,41]. Das Instrument ist in verschiedene Sprachen übersetzt.</p> <p>Die Generic-Core-Skala des PedsQL (Version 4.0) ist jedoch kein krankheitsspezifisches Instrument und kann keine detaillierten Informationen über die spezifischen Faktoren liefern, die die Lebensqualität bei pädiatrischen Patienten mit neuromuskulären Erkrankungen beeinflussen. Daher wurde für den PedsQL das Neuromuskuläre Modul 3.0 entwickelt, um verschiedene Dimensionen der Lebensqualität für Kinder im Alter von 2 bis 18 Jahren mit neuromuskulären Störungen, insbesondere SMA, zu messen. Bei der internen Konsistenzprüfung wurde das erforderliche Mindest-Alpha von 0,7 sowohl für den Kinderfragebogen als auch den Elternfragebogen überschritten. Die Test-Retest-Reliabilität lag für den Kinderfragebogen bei ICC = 0,81 und für den Elternfragebogen bei ICC = 0,89 [25,40]. Für das Neuromuskuläre Modul 3.0 liegen</p>

Mortalität	
	<p>bisher keine Informationen zur Änderungssensitivität und zur Festlegung von Relevanzschwellen vor.</p> <p>Insgesamt handelt es sich bei der Generic-Core-Skala und dem Neuromuskulären Modul 3.0 des PedsQL (Version 4.0) um gut untersuchte Instrumente zur Messung der Lebensqualität bei Kindern und Jugendlichen mit chronischen Erkrankungen mit ausreichender Testgüte. Eine entsprechende krankheitsspezifische Validierung in einer Patientenpopulation mit SMA ist erfolgt.</p> <p>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.“</p>

Tabelle 14: Beschreibung der Endpunkte der Kategorie Sicherheit in den Studien ENDEAR und CHERISH

Sicherheit	
Endpunkt	Studie ENDEAR: Unerwünschte Ereignisse (UE)
Operationalisierung	<p>Ein <u>unerwünschtes Ereignis (UE)</u> ist definiert als jegliches unerwünschte medizinische Ereignis, z. B. ein unangenehmes und ungewolltes körperliches Anzeichen, Symptom oder ein Laborparameter, welches während der klinischen Studie auftrat oder sich verschlechterte.</p> <p>Drei mögliche Szenarien sind zu berücksichtigen:</p> <ul style="list-style-type: none"> • Das 1. UE trat vor der ersten Dosis auf, das 2. UE trat bei oder nach der 1. Dosis auf und der Schweregrad stieg: nur das 2. UE wurde gezählt. • Das 1. UE trat vor der ersten Dosis auf, das 2. UE trat bei oder nach der 1. Dosis auf und der Schweregrad sank: keines der UE wurde gezählt. • Beide UE traten bei oder nach der 1. Dosis auf: war die Schwere des 2. UE höher als die Schwere des 1. UE, so zählten beide UE. Verringerte sich die Schwere, wurde nur das 1. UE gewertet. <p>Eine Klassifizierung der UE erfolgte nach Systemorganklasse gemäß Medical Dictionary for Regulatory Activities (MedDRA), Version 16.1.</p> <p><u>Schwerwiegende unerwünschte Ereignisse (SUE)</u> sind definiert als UE,</p> <ul style="list-style-type: none"> • das den Tod eines Patienten zur Folge hat; • das ein lebensbedrohendes UE darstellt – das schließt keine Ereignisse ein, die bei einem schwereren Verlauf tödlich enden können; • das einen Krankenhausaufenthalt oder die Verlängerung eines Krankenhausaufenthalts erfordert; • das eine bleibende oder bedeutsame Behinderung oder Handlungsunfähigkeit herbeiführt; • das eine kongenitale Anomalie oder einen Geburtsfehler nach sich zieht; • das ein bedeutendes medizinisches Ereignis darstellt, welches weder zum Tod führt noch lebensbedrohlich ist oder eine Hospitalisierung erfordert, aber den Patienten gefährdet oder eine medizinische Behandlung nach sich zieht, um einen der oben genannten Punkte zu verhindern. Jegliche mutmaßliche Übertragung eines infektiösen Erregers durch die Studienmedikation, wird als bedeutendes medizinisches Ereignis angesehen. <p>Therapieabbrüche aufgrund von UE sind definiert als ein UE, das einen permanenten Abbruch der Studienbehandlung erfordert.</p>

Sicherheit	
	<p>Erhebungszeitpunkte: Unerwünschte Ereignisse wurden bei jeder Studienvsiste (vor Dosierung, während Dosierung und nach Dosierung) sowie über die gesamte Studie erhoben. Fehlende Werte wurden imputiert. Die Auswertung des Endpunktes erfolgte auf Basis der Safety-Population.</p>
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunktes ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz: Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patienten-relevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von <u>Nebenwirkungen</u> oder einer Verbesserung der Lebensqualität.“</p>
Endpunkt	Studie CHERISH: Unerwünschte Ereignisse (UE)
Operationalisierung	<p>Ein <u>unerwünschtes Ereignis (UE)</u> ist definiert als jegliches unerwünschte medizinische Ereignis, z. B. ein unangenehmes und ungewolltes körperliches Anzeichen, Symptom oder ein Laborparameter, welches während der klinischen Studie auftrat oder sich verschlechterte. Ein kausaler Zusammenhang zur Studienmedikation musste nicht zwingend bestehen.</p> <p>Drei mögliche Szenarien sind zu berücksichtigen:</p> <ul style="list-style-type: none"> • Das 1. UE trat vor der ersten Dosis auf, das 2. UE trat bei oder nach der 1. Dosis auf und der Schweregrad stieg: nur das 2. UE wurde gezählt. • Das 1. UE trat vor der ersten Dosis auf, das 2. UE trat bei oder nach der 1. Dosis auf und der Schweregrad sank: keines der UE wurde gezählt. • Beide UE traten bei oder nach der 1. Dosis auf: war die Schwere des 2. UE höher als die Schwere des 1. UE, so zählten beide UE. Verringerte sich die Schwere, wurde nur das 1. UE gewertet. <p>Eine Klassifizierung der UE erfolgte nach Systemorganklasse (MedDRA, Version 17.1).</p> <p><u>Schwerwiegende unerwünschte Ereignisse (SUE)</u> sind definiert als UE,</p> <ul style="list-style-type: none"> • das den Tod eines Patienten zur Folge hat; • das ein lebensbedrohendes UE darstellt – das schließt keine Ereignisse ein, die bei einem schwereren Verlauf tödlich enden können; • das einen Krankenhausaufenthalt oder die Verlängerung eines Krankenhausaufenthalts erfordert; • das eine bleibende oder bedeutsame Behinderung oder Handlungsunfähigkeit herbeiführt; • das eine kongenitale Anomalie oder einen Geburtsfehler nach sich zieht; • das ein bedeutendes medizinisches Ereignis darstellt, welches weder zum Tod führt noch lebensbedrohlich ist oder eine Hospitalisierung erfordert, aber den Patienten gefährdet oder eine medizinische Behandlung nach sich zieht, um einen der oben genannten Punkte zu verhindern. Jegliche mutmaßliche Übertragung eines infektiösen Erregers durch die Studienmedikation, wird als bedeutendes medizinisches Ereignis angesehen. <p><u>Therapieabbrüche aufgrund von UE</u> sind definiert als ein UE, dass einen permanenten Abbruch der Studienbehandlung erfordert.</p> <p>Erhebungszeitpunkte: Unerwünschte Ereignisse wurden bei jeder Studienvsiste (vor Dosierung, während Dosierung und nach Dosierung) sowie über die gesamte Studie erhoben. Fehlende Werte wurden imputiert. Die Auswertung des Endpunktes erfolgte auf Basis der Safety-Population.</p>

Sicherheit	
Bewertung	<p>Operationalisierung: Die Operationalisierung des Endpunkts ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz: Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patienten-relevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von <u>Nebenwirkungen</u> oder einer Verbesserung der Lebensqualität.“</p>

2.5.4 Statistische Methoden

ENDEAR

Eine statistische Hypothese ist im Statistischen Analyseplan der ENDEAR-Studie nicht beschrieben.

Primäre Endpunkte

Die Studie hat zwei primäre Endpunkte:

- Erster primärer Endpunkt: Anteil Responder mit Motorischem Meilenstein mittels HINE (Subskala 2) zum Zeitpunkt Tag 183, 302 oder 394.
- Zweiter primärer Endpunkt: Zeit bis zum Tod oder zur dauerhaften Beatmung (≥ 16 Stunden/Tag durchgehend für > 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses oder Tracheotomie)

Für die primären Endpunkte HINE-Responder (Subskala 2) waren sechs Sensitivitätsanalysen geplant:

1. Anteil der HINE-Responder (Subskala 2) – Auswertung wie Hauptanalyse, ausgenommen Patienten des Efficacy Sets ohne Untersuchungen an Tag 183, 302 oder 394).
2. Anteil der HINE-Responder (Subskala 2) – Auswertung wie Hauptanalyse, jedoch nur Patienten mit einer Verbesserung.
3. Anteil der HINE-Responder (Subskala 2) – Patienten mit einer Verbesserung um mindestens zwei Punkte im Gesamtscore, ohne bewusstes Greifen.
4. Anteil der HINE-Responder (Subskala 2) – Safety-Analysis-Set.
5. Anteil der HINE-Responder (Subskala 2) – Per-Protokoll-Set.
6. Anteil der HINE-Responder (Subskala 2) – ausgenommen Patienten ohne Baseline-Bewertung.

Für den zweiten primären Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung waren sieben Sensitivitätsanalysen mit Adaption des statistischen Modells und/oder Analysepopulation vorgesehen.

Interimsanalyse und finale Wirksamkeitsanalyse

Eine Interimsanalyse war für den 15. Juni 2016 geplant. Die finale Analyse sollte mit Bewertung des letzten Patienten nach 13 Monaten erfolgen oder früher, wenn die Studie vorzeitig beendet wird, weil die Durchführung einer Scheinintervention ethisch nicht länger vertretbar oder machbar erscheint. Eine Interimsanalyse war lediglich für den ersten primären Endpunkt vorgesehen, da bis

zum Zeitpunkt der Analyse eine geringe Ereignisrate erwartet wurde. Sowohl der Interims- als auch der finalen Analyse liegen Werte der Erhebungszeitpunkte Tag 183, 302 oder 394 zugrunde.

Vor dem Hintergrund, dass der zweite primäre Endpunkt sowie alle sekundären Wirksamkeitsendpunkte keiner Interimsanalyse unterzogen wurden, ist ein Alpha von 0,05 für die abschließende Auswertung dieser Endpunkte verwendet worden. Der erste primäre Wirksamkeitsendpunkt wurde basierend auf dem linearen Lan-DeMets-Ansatz unter Annahme einer Informationsfraktion von 64 % bei einem Alpha von 0,032 getestet. Es wurde erwartet, dass diese $N \sim 80$ Patienten (einschließlich Patienten die verstorben sind oder die Studie vorzeitig beendet haben) zum Zeitpunkt der Interimsanalyse den Erhebungszeitpunkt Tag 183 erreicht haben. Da die Scheinintervention ethisch nicht länger vertretbar war, entschied sich der Sponsor die Studie vorzeitig zu beenden. Der erste primäre Endpunkt zeigte zum Zeitpunkt der Interimsanalyse ein statistisch signifikantes Ergebnis. Basierend auf demselben statistischen Modell bei einem Alpha von 0,05 erfolgte eine finale Analyse des zweiten primären Endpunktes sowie aller anderen sekundären Wirksamkeitsendpunkte.

Begründung der Fallzahlen

Die Stichprobengröße für diese Studie basiert auf einer Verdopplung der medianen Zeit bis zum Tod oder zur dauerhaften Beatmung im Nusinersen-Arm im Vergleich zum Kontrollarm. Basierend auf Daten einer natürlichen Verlaufsstudie für die Zielpopulation sei die mediane Zeit bis zum Ereigniszeitpunkt für den Scheininterventionsarm mit 5 bis 6 Monaten vom Zeitpunkt der Verabreichung der 1. Dosis der Studienmedikation zu kalkulieren. Bei einer Randomisierung im Verhältnis 2:1 und 13 Monaten Nachbeobachtung, würde mit 111 Patienten eine Power von 80 % erreicht, um eine Verdopplung der medianen Zeit bis zum Tod oder zur dauerhaften Beatmung für die Nusinersen-Gruppe im Vergleich zur Kontrollgruppe bei einem zweiseitigen Signifikanzniveau von 5 % zu erzielen. Für den primären Endpunkt wird die Power auf ca. 60 % geschätzt, um einen statistisch signifikanten Behandlungsunterschied zwischen den Gruppen zum Zeitpunkt der Interimsanalyse ($N \sim 80$) zu zeigen. Dies sei unter einer Voraussetzung von drei Respondern in der Kontrollgruppe ($3/26 = 11,5\%$) sowie 20 Respondern im Nusinersen-Arm ($20/52 = 38,5\%$) und einem Alpha von 0,035 erreicht. In der finalen Analyse mit einem Alpha von 0,03 sei bei 111 Patienten etwa bei 78 % der Patienten eine Ansprechrate von 38,5 % im Interventions- vs. 11,5 % im Kontrollarm erreicht.

Analysepopulationen

Die finale Analyse erfolgte für folgende Populationen:

- **Intention-to-Treat (ITT):** Alle randomisierten Patienten, die mindestens eine Dosis des Studienmedikamentes / der Scheinintervention erhielten. Analysen erfolgen für Studienteilnehmer nach zugewiesener Behandlung.
- **Efficacy-Set (ES):** Subpopulation der ITT-Population, beinhaltet alle Patienten mit Studienvisiten an Tag 183, 302 oder 394 und alle Patienten mit einem Zeitfenster von mindestens 190 Tagen ($183 + 7$ Tage) zwischen dem Datum der ersten Einnahme der Studienmedikation und dem geplanten Datenschnitt für die finale Analyse. Patienten, die verstarben oder die Studie abbrachen, mussten ein Zeitfenster von mindestens 176 Tagen ($183 - 7$ Tage) zwischen der ersten Einnahme der Studienmedikation und dem finalen geplanten Datenschnitt vorweisen.

- **Per-Protokoll (PP):** Subpopulation der ITT Population, die mindestens alle vier Aufsättigungsdosen/Scheinverfahren erhalten haben, eine Bewertung der Wirksamkeitsparameter zu Baseline und Tag 183 sowie keine relevante Protokollverletzung aufweisen.
- **Safety-Analysis-Set (SAS):** Alle Patienten, die mindestens eine Dosis des Studienmedikamentes / der Scheinintervention erhielten. Alle Sicherheitsanalysen erfolgen auf Basis dieser Sicherheitspopulation. Patienten, die versehentlich Nusinersen erhielten, obwohl sie der Kontrollgruppe zugeordnet waren, wurden im Nusinersen-Arm analysiert.
- **Pharmakokinetik (PK):** Alle randomisierten Patienten mit mindestens einer auswertbaren Pharmakokinetik nach Baseline.

Die Studienpopulation (N=121) wurde im Verhältnis von 2:1 auf die Behandlungsarme randomisiert (Nusinersen: N=80, Scheinintervention: N=41) und hinsichtlich Krankheitsdauer (Alter des Patienten bei Screening - Alter bei Symptombeginn: ≤ 12 vs. > 12 Wochen) stratifiziert.

CHERISH

Eine statistische Hypothese ist im Statistischen Analyseplan der CHERISH-Studie nicht beschrieben.

Primärer Endpunkt

Der primäre Endpunkt der Studie ist die Veränderung des HFMSE-Wertes nach 15 Monaten im Vergleich zu Baseline.

Für den primären Endpunkt waren fünf Sensitivitätsanalysen mit Adaption des statistischen Modells und/oder der Analysepopulation vorgesehen.

Interimsanalyse und finale Wirksamkeitsanalyse

Bei der Interimsanalyse wurde nur der primäre Wirksamkeitsendpunkt (Änderung des HFMSE-Wertes nach 15 Monaten im Vergleich zu Baseline) getestet. Da die sekundären Wirksamkeitsendpunkte nicht in der Interimsanalyse untersucht wurden (d. h. keine Alpha-Ausgaben), ist in der finalen Analyse dieser Endpunkte ein Alphalevel von 0,05 verwendet worden. Die Interimsanalyse des primären Wirksamkeitsendpunkts erfolgte unter Verwendung eines Alphalevels von 0,025. Die Ergebnisse der Interimsanalyse des primären Wirksamkeitsendpunktes zeigten ein statistisch signifikantes Ergebnis. Infolgedessen ist die Studie frühzeitig beendet worden, weil die Durchführung einer scheininterventionskontrollierten Studie nicht mehr als ethisch gerechtfertigt erachtet wurde. Die Durchführung einer Interimsanalyse war möglich, falls alle 6-Monats-Bewertungen abgeschlossen waren und mindestens 39 Patienten eine vollständige 15-Monats-Bewertung durchlaufen haben. Da für einige Patienten die 15-Monats-Daten nicht vorliegen werden, werden die fehlenden Werte gemäß der multiplen Imputationsmethode imputiert [34].

Die primäre Wirksamkeitsanalyse erfolgte auf Basis des ITT-Sets. Fehlende HFMSE-Werte wurden unter Verwendung der Mehrfach-Imputationsmethode [34,35] imputiert. Jeder Datensatz wurde mittels Kovarianzanalyse (ANCOVA-Modell) analysiert, mit Behandlungsgruppe als Faktor und Alter bei Screening sowie Baseline-HFMSE-Wert als Kovariaten. Die Schätzungen aus den 100 Datensätzen wurden kombiniert, um eine Gesamtschätzung mit entsprechenden Konfidenzintervallen (KI) und p-Wert [29] zu liefern.

Begründung der Fallzahlen

Die Stichprobengröße wurde auf Grundlage verfügbarer Daten zum natürlichen Krankheitsverlauf sowie den klinischen Studien CS1 und CS2 kalkuliert. Eine statistische Power von mindestens 90 % würde mit 70 Patienten im Nusinersen-Arm und 35 Patienten im Scheininterventions-Arm erreicht, um einen 3-Punkte-Unterschied in der Veränderung des HFMSE zu erzielen (mit einer Standardabweichung von 4,4 unter Verwendung eines zweiseitigen t-Tests mit einem Alphalevel von 0,05). Der Einschluss von 117 Patienten stellt sicher, dass eine geringe Drop-out-Rate die Power der primären Wirksamkeitsanalyse nicht beeinflusst.

Analysepopulationen

Die finale Analyse erfolgte für folgende Populationen:

- **Intention-to-Treat:** Alle randomisierten Patienten, die mindestens eine Dosis des Studienmedikamentes / der Scheinintervention erhielten. Analysen erfolgen für Studienteilnehmer nach zugewiesener Behandlung.
- **Efficacy Set:** Subpopulation der ITT-Population, beinhaltet alle Patienten mit Studienvisiten an Tag 456 (z. B. Monat 15) und alle Patienten mit einem Zeitfenster von mindestens 463 Tagen (456 + 7 Tage) zwischen dem Datum der ersten Einnahme der Studienmedikation und dem geplanten Datenschnitt für die finale Analyse.
- **Per-Protokoll:** Alle Patienten der ITT-Population, die mindestens die 3 Aufsättigungsdosen der Studienmedikation/Scheinintervention erhielten, eine Baseline- sowie eine Untersuchung der Wirksamkeitsparameter an Tag 169 und keine relevanten Protokollverletzungen aufweisen.
- **Safety-Analysis-Set:** Alle Patienten, die mindestens eine Dosis des Studienmedikamentes / der Scheinintervention erhielten. Alle Sicherheitsanalysen erfolgen auf Basis dieser Sicherheitspopulation.
- **Pharmakokinetik:** Alle randomisierten Patienten mit mindestens einer auswertbaren Pharmakokinetik nach Baseline.

Die Studienpopulation (N=126) wurde im Verhältnis von 2:1 auf die Behandlungsarme randomisiert (Nusinersen: N=84, Scheinintervention: N=42) und hinsichtlich des Alters bei Screening (≤ 6 vs. > 6 Jahre) stratifiziert.

2.6 Ergebnisse zum Zusatznutzen

Als Grundlage für die Nutzenbewertung von Nusinersen im Vergleich zu BSC gelten die vom pU eingereichten zwei pivotalen Studien ENDEAR (CS3B) und CHERISH (CS4). In der ENDEAR-Studie werden Ergebnisse für Patienten mit infantiler SMA (Typ I) dargestellt, in der CHERISH-Studie entsprechend Ergebnisse für Patienten mit einem späteren Krankheitsbeginn der SMA. Beide Studien sind aufgrund eines erzielten Wirksamkeitsnachweises vorzeitig beendet worden. Die endgültigen Analysen wurden mit Fertigstellung des Studienberichtes zur ENDEAR-Studie bis zum 10. Februar 2017, basierend auf dem finalen Datenschnitt vom 16. Dezember 2016, durchgeführt. In der CHERISH-Studie wurde der finale Datenschnitt am 3. März 2017 gesetzt, die anschließenden statistischen Endpunktauswertungen erfolgten bis zur Fertigstellung der Ergebnisse am 6. Juni 2017.

Seitens des pU wurden mit Dossiereinreichung entsprechend aufbereitete Analysen für die Studien ENDEAR und CHERISH, basierend auf den jeweils finalen Datenschnitten, beigefügt. Ergebnisse zu Mortalität, Morbidität, Lebensqualität und Sicherheit basieren entsprechend auf diesen Analysen. Diskrepanzen zu den in den Studienberichten dargestellten Ergebnissen werden adressiert. Allgemeine Angaben zur Charakterisierung der Studienpopulation wurden den Studienberichten zu ENDEAR und CHERISH entnommen.

2.6.1 Charakterisierung der Studienpopulation

In die pivotalen Studien ENDEAR und CHERISH sind annähernd gleich viele Patienten randomisiert und im Verhältnis 2:1 auf die Studienarme aufgeteilt worden. Die Interventionsarme umfassten 80 bzw. 84 Probanden und die Kontrollarme 41 bzw. 42 Probanden. Alle Patienten erhielten entweder Nusinersen oder eine Scheinintervention. Dabei bestand ein Unterschied der Verabreichungszeiträume der Studienmedikation zwischen den Studien. In der ENDEAR-Studie ist die Studienmedikation gemäß aktueller Fachinformation (Stand: Mai 2017) als intrathekale Bolusinjektion über 1–3 Minuten an den Studientagen 1, 15, 29, 64 (Aufsättigung), 183 und 302 (Erhaltung), über eine 10-monatige Behandlungsphase und eine 3-monatige Follow-up-Phase (bis Tag 394 ± 7 Tage) verabreicht worden. Wegen der langen Halbwertszeit von Nusinersen ist die Behandlung während des Follow-ups¹ fortgeführt worden. In der CHERISH-Studie ist Nusinersen respektive die Scheinintervention an den Studientagen 1, 29, 85 (Aufsättigung) und 274 (Erhaltung), über eine 9-monatige Behandlungsphase und eine 6-monatige Follow-up-Phase² (bis Tag 456 ± 7 Tage), in der die Studienbehandlung fortgeführt wurde, verabreicht worden. Die Follow-up-Phase unterscheidet sich aufgrund der fortgesetzten Exposition somit nicht von der Behandlungsphase. Die Diskrepanz in der Dosierung der beiden Studien wird seitens pU mit einer Vereinbarung über die Dosierungsschemata mit der European Medicines Agency (EMA) begründet. Dabei sei die Dosierung in der ENDEAR-Studie für die Fachinformation in Abstimmung mit der EMA als geeignet erachtet worden. Die Zweckmäßigkeit der Dosierung würde in laufenden Studien und Monitoringprozessen weiter bewertet [1]. In den beigefügten Zulassungsunterlagen finden sich keine ergänzenden Informationen zur Festlegung der Dosierung [18]. In der ENDEAR-Studie schlossen im Nusinersen-Arm 49 % und im Kontrollarm 32 % der Probanden vorzeitig die Studie ab, in der CHERISH-Studie waren es entsprechend 21 vs. 19 % der Probanden. Zu einem Behandlungsabbruch in der ENDEAR-Studie kam es bei 59 % der Studienteilnehmer unter Nusinersen und bei 63 % der Studienteilnehmer unter BSC. Darunter waren 16 vs. 39 % (Nusinersen vs. BSC), die aufgrund von unerwünschten Ereignissen abbrachen. In der CHERISH-Studie brach insgesamt nur ein Patient im Nusinersen-Arm die Behandlung ab. Die Studienteilnehmer beider Studien hatten die Möglichkeit die Behandlung nach Studienende in einer offenen Extensionsstudie weiterzuführen.

¹ Nach Abschluss der Studiervisite an Tag 303 begann die 3-monatige Follow-up-Phase. Die Phase bestand aus einem Besuch an Tag 394 und wöchentlichen Telefonabfragen. Patienten, die in den letzten drei Wochen vor der Studiervisite an Tag 394 ≥ 16 Stunden/Tag Beatmung aufwiesen, wurden bis zur Bestätigung oder Widerlegung des Erreichens des primären Endpunktes weiterhin mittels Telefon kontaktiert.

² Nach Abschluss der Studiervisite an Tag 275 begann die 6-monatige Follow-up-Phase. Dieser Zeitraum bestand aus Studiervisiten an den Tagen 365 und 456 und monatlichen Telefonabfragen.

Angesichts der vermehrten Anzahl Todesfälle in der Kontrollgruppe, war die Studiendauer in der ENDEAR-Studie pro Patient in der Nusinersen-Gruppe im Median länger als in der Kontrollgruppe (Intervention 280,0 Tage vs. Kontrolle 187,0 Tage). Die mediane Studiendauer war in der CHERISH-Studie in beiden Behandlungsgruppen ähnlich (Intervention 450,5 Tage vs. Kontrolle 451,0 Tage). Weitere Angaben sind Tabelle 15 zu entnehmen.

Tabelle 15: Allgemeine Angaben der Studien ENDEAR und CHERISH

Studie	ENDEAR		CHERISH	
	BSC (N=41)	Nusinersen (N=80)	BSC (N=42)	Nusinersen (N=84)
Allgemeine Angaben				
Randomisierte Patienten, n (%)	41 (100)	81 (101)	42 (100)	84 (100)
Patienten ohne Behandlung, n (%)	0	1 (1)	0	0
Patienten mit Behandlung, n (%)	41 (100)	80 (100)	42 (100)	84 (100)
Studie abgeschlossen, n (%)	24 (59)	65 (81)	42 (100)	84 (100)
vorzeitig abgeschlossen	13 (32)	39 (49)	8 (19)	18 (21)
abgeschlossen inkl. Follow-up	11 (27)	26 (33)	34 (81)	66 (79)
Behandlungsabbruch, n (%)	26 (63)	47 (59)	0	1 (1)
vorzeitiges Studienende	10 (24)	32 (40)	0	1 (1)
unerwünschte Ereignisse	16 (39)	13 (16)	0	0
Einwilligung zurückgezogen	0	2 (3)	0	0
andere	0	0	0	0
Studienabbruch, n (%)	17 (41)	15 (19)	0	0
unerwünschte Ereignisse	16 (39)	13 (16)	0	0
Einwilligung zurückgezogen	1 (2)	2 (3)	0	0
andere			0	0
Studiendauer in Tagen pro Patient				
Mittelwert (SD)	226,2 (140,1)	265,0 (123,2)	443,5 (24,0)	440,6 (30,0)
Median (min;max)	187,0 (20;423)	280,0 (6;442)	450,0 (364;482)	450,5 (324;470)

Abkürzungen: BSC: Best Supportive Care; SD: Standardabweichung

Eine Zusammenfassung der Studiencharakteristika der pivotalen Studien ENDEAR und CHERISH ist in Tabelle 16 dargestellt. In der ENDEAR-Studie hatten zum Zeitpunkt des Datenschnitts am 16. Dezember 2016 73 Probanden (91 %) mindestens 4 Dosen Nusinersen erhalten (d. h. die Aufsättigungsphase beendet), wovon 32 Probanden (40 %) alle 6 Dosen erhielten. Von den 41 Probanden in der Kontrollgruppe hatten sich 34 Probanden (83 %) 4 Scheinverfahren unterzogen, wobei 14 Probanden (34 %) alle 6 Verfahren erhielten. Die kumulative Dosis der Studienbehandlung betrug im Median 53,5 mg. In der CHERISH-Studie hatten mit dem Datenschnitt am 3. März 2017 alle 84 Probanden mindestens 3 Dosen Nusinersen erhalten, 83 Probanden (99 %) erhielten 4 Dosen der Studienmedikation. In der Kontrollgruppe haben sich alle 42 Probanden 4 Scheinverfahren unterzogen. Die kumulative Dosis der Studienbehandlung betrug im Median 48 mg.

Beide Studien weisen eine ausgeglichene Geschlechterverteilung auf, d. h. der Anteil Mädchen und Jungen innerhalb der Studienpopulationen war annähernd identisch. Das Alter bei Screening war mit einem mittleren Alter von 147,2 vs. 164,7 Tagen im Nusinersen-Arm der ENDEAR-Studie geringer als im BSC-Arm. Bei Verabreichung der 1. Dosis der Studienmedikation waren die Probanden in der Interventionsgruppe zwischen 3 und 6 Monaten alt (54 %), in der Kontrollgruppe älter als 6 Monate (61 %). Die Mehrzahl der Patienten wurde in Nordamerika und Europa

eingeschlossen und war zudem kaukasischer Abstammung. In der Nusinersen-Gruppe waren mehr Patienten aus dem Asien-Pazifik-Raum eingeschlossen als in der Kontrollgruppe (15 vs. 5 %). 90 % der Probanden unter Nusinersen vs. 78 % unter BSC waren bei Symptombeginn jünger als 12 Wochen (Mittelwert (MW): 7,9 vs. 9,6 Wochen). Das Alter bei SMA-Diagnose lag entsprechend im Mittel bei 12,6 vs. 17,5 Wochen (Intervention vs. Kontrolle). Im Interventionsarm wiesen zu Studienbeginn mehr Probanden eine paradoxe Atmung (89 vs. 66 %), Pneumonie / respiratorische Symptome (35 vs. 22 %) und Schluckbeschwerden / Probleme bei der Nahrungsaufnahme (51 vs. 29 %) auf als im Kontrollarm. In der CHERISH-Studie waren 83 % und 86 % (Nusinersen vs. BSC) der Probanden jünger als 6 Jahre, davon waren im Kontrollarm die meisten Probanden 2 Jahre alt. Das mittlere Alter bei Symptombeginn war vergleichbar in beiden Studienarmen und lag bei 11,1 Monaten im Nusinersen-Arm vs. 11,3 Monaten im BSC-Arm. Eine SMA-Diagnose wurde bei den Probanden im Interventionsarm im Mittel mit 19,9 Monaten gestellt im Vergleich zu 18,3 Monaten im Kontrollarm. Zu Studienbeginn konnten in der Kontrollgruppe mehr Patienten frei stehen als im Interventionsarm (29 vs. 13 %). Die Mehrzahl der Patienten wurde auch hier in Nordamerika und Europa eingeschlossen und war zu 75 % kaukasischer Abstammung. Während in der ENDEAR-Studie, abgesehen von einem Probanden in der Kontrollgruppe, alle Studienteilnehmer 2 SMN2-Kopien aufwiesen, waren es in der CHERISH-Studie jeweils 88 %, die 3 SMN2-Kopien hatten.

Tabelle 16: Charakterisierung der Studienpopulation der Studien ENDEAR und CHERISH

Merkmal	ENDEAR		CHERISH	
	BSC (N=41)	Nusinersen (N=80)	BSC (N=42)	Nusinersen (N=84)
<i>Geschlecht, n (%)</i>				
männlich	17 (41)	37 (46)	21 (50)	38 (45)
weiblich	24 (59)	43 (54)	21 (50)	46 (55)
<i>Alter bei Screening (Tage)</i>				
< 30 Tage (jünger als 1 M.)	1 (2)	0 (0)		
≥ 30 Tage / < 90 Tage (1–3 M.)	3 (7)	10 (3)	-	-
≥ 90 Tage / < 180 Tage (3–6 M.)	15 (37)	47 (59)		
≥ 180 Tage (6 M. und älter)	22 (54)	23 (29)		
<i>Alter bei Screening (Tage)</i>				
Mittelwert (SD)	164,7 (48,5)	147,2 (46,9)	-	-
Median (min;max)	190,0 (20;211)	151,5 (32;210)		
<i>Alter bei 1. Dosis (Tage)</i>				
< 30 Tage (jünger als 1 M.)	0	0		
≥ 30 Tage / < 90 Tage (1–3 M.)	2 (5)	6 (8)	-	-
≥ 90 Tage / < 180 Tage (3–6 M.)	14 (34)	43 (54)		
≥ 180 Tage (6 M. und älter)	25 (61)	31 (39)		
<i>Alter bei 1. Dosis (Tage)</i>				
Mittelwert (SD)	180,5 (50,9)	163,4 (49,6)	-	-
Median (min;max)	205,0 (30;262)	164,5 (52;242)		
<i>Alter (Jahre)</i>				
2			19 (45)	22 (26)
3			6 (14)	18 (21)
4	-	-	9 (21)	23 (27)
5			2 (5)	7 (8)
< 6			36 (86)	70 (83)
≥ 6			6 (14)	14 (17)

Merkmal	ENDEAR		CHERISH	
	BSC (N=41)	Nusinersen (N=80)	BSC (N=42)	Nusinersen (N=84)
<i>Alter (Jahre)</i>				
Mittelwert (SD)	-	-	3,4 (1,61)	3,8 (1,63)
Median (min;max)			3,0 (2;7)	4,0 (2;9)
<i>Abstammung, n (%)</i>				
asiatisch	1 (2)	5 (6)	7 (17)	16 (19)
dunkelhäutig	0 (0)	3 (4)	1 (2)	1 (1)
kaukasisch	36 (88)	68 (85)	30 (71)	64 (76)
multipel ¹⁾	2 (5)	1 (1)	4 (10)	3 (4)
andere	1 (2)	3 (4)	0	0
nicht berichtet	1 (2)	0 (0)	-	-
<i>Geografische Region, n (%)</i>				
Nordamerika	22 (54)	38 (48)	23 (55)	47 (56)
Europa	17 (41)	30 (38)	14 (33)	28 (33)
Asien-Pazifik-Raum	2 (5)	12 (15)	5 (12)	9 (11)
<i>Schwangerschaftsdauer (Wochen)</i>				
Mittelwert (SD)	39,3 (1,33)	38,9 (1,30)	-	-
Median (min;max)	40,0 (37;42)	39,0 (36;41)		
<i>Geburtsgewicht (kg)</i>				
Mittelwert (SD)	3,5 (0,74)	3,3 (0,44)	-	-
Median (min;max)	3,4 (2,13;6,13)	3,3 (2,41;4,74)		
<i>Alter bei Symptombeginn (ENDEAR: Wochen / CHERISH: Monate)</i>				
≤ 12 Wochen	32 (78)	72 (90)	-	-
> 12 Wochen	9 (22)	8 (10)	-	-
Mittelwert (SD)	9,6 (4,66)	7,9 (3,95)	11,3 (3,45)	11,1 (3,32)
Median (min;max)	8,0 (1;20)	6,5 (2;18)	11,0 (6;20)	10,0 (6;20)
<i>Krankheitsdauer (Wochen)</i>				
≤ 12 Wochen	18 (44)	34 (43)	-	-
> 12 Wochen	23 (56)	46 (58)	-	-
Mittelwert (SD)	13,9 (5,66)	13,2 (5,46)		
Median (min;max)	12,7 (0;23,1)	13,1 (0;25,9)		
<i>Alter bei SMA-Diagnose (ENDEAR: Wochen / CHERISH: Monate)</i>				
Mittelwert (SD)	17,5 (7,47)	12,6 (6,63)	18,3 (7,56)	19,9 (7,94)
Median (min;max)	20,0 (2;30)	11,0 (0;29)	18,0 (0;46)	18,0 (0;48)
<i>Zeit von Krankheitsbeginn bis Studienregistrierung (Monate)</i>				
Mittelwert (SD)	-	-	34,8 (18,74)	39,9 (20,16)
Median (min;max)			30,2 (10;80)	39,3 (8;94)
<i>Zeit von Diagnose bis Studien- registrierung (Monate)</i>				
Mittelwert (SD)	-	-	27,8 (18,46)	31,1 (20,05)
Median (min;max)			26,0 (2;72)	27,8 (2;86)
<i>SMN2-Kopien²⁾, n (%)</i>				
2 Kopien	40 (98)	80 (100)	4 (10)	6 (7)
anders als 2 Kopien	1 (2)	0 (0)	-	-
3 Kopien	-	-	37 (88)	74 (88)
4 Kopien	-	-	1 (2)	2 (2)

Merkmal	ENDEAR		CHERISH	
	BSC (N=41)	Nusinersen (N=80)	BSC (N=42)	Nusinersen (N=84)
unbekannt	0 (0)	0 (0)	0 (0)	2 (2)
<i>Personen mit..., n (%)</i>				
Hypotonie	41 (100)	80 (100)		
Motorische Entwicklungsverzögerung	39 (95)	71 (89)		
Paradoxe Atmung	27 (66)	71 (89)		
Pneumonie / respiratorische Symptome	9 (22)	28 (35)	-	-
Schwäche der Extremitäten	41 (100)	79 (99)		
Schluckbeschwerden / Probleme bei der Nahrungsaufnahme	12 (29)	41 (51)		
anderes	14 (34)	20 (25)		
<i>Personen mit motorischem Meilenstein, n (%)</i>				
selbstständiges Sitzen	-	-	42 (100)	84 (100)
freies Stehen	-	-	12 (29)	11 (13)
Gehen mit Unterstützung	-	-	14 (33)	20 (34)
selbstständiges Gehen (min. 50 m)	-	-	0 (0)	0 (0)
<i>Personen mit Rollstuhl, n (%)</i>				
ja	-	-	29 (69)	64 (76)
nein	-	-	13 (31)	20 (24)
<i>Personen in Physiotherapie, n (%)</i>				
ja	-	-	38 (90)	78 (93)
nein	-	-	4 (10)	5 (6)
unbekannt	-	-	0 (0)	1 (1)
<i>Verabreichte Dosen der Studienmedikation, n (%)</i>				
1	2 (5)	1 (1)	0 (0)	0 (0)
2	1 (2)	2 (3)	0 (0)	0 (0)
3	4 (10)	4 (5)	0 (0)	1 (1)
4	15 (37)	20 (25)	42 (100)	83 (99)
5	5 (12)	21 (26)	-	-
6	14 (34)	32 (40)	-	-
<i>Verabreichte Dosen der Studienmedikation, n (%)</i>				
Mittelwert (SD)	4,5 (1,38)	4,9 (1,13)	4,0 (0,0)	4,0 (0,11)
Median (min;max)	4,0 (1;6)	5,0 (1;6)	4,0 (4;4)	4,0 (3;4)

¹⁾ „Multipel“ bezogen auf die Abstammung der Studienteilnehmer wird im CSR der Studie ENDEAR nicht genauer spezifiziert.

²⁾ Die SMN2-Genkopienzahl wurde mittels Test eines Zentrallabors ermittelt. Der Patient mit 3 Kopien, basierend auf den Daten des zentralen Labors, wies bei Screening, basierend auf dem Ergebnis eines lokalen Labors, eine Kopienzahl von 2 auf. Bei fehlenden Ergebnissen des Zentrallabors, ist die Kopienzahl bei Screening, basierend auf den Daten des lokalen Labors, für alle Patienten 2.

Abkürzungen: BSC: Best Supportive Care; M: Monate; MW: Mittelwert; SD: Standardabweichung; SMA: Spinale Muskelatrophie; SMN: Survival of Motor Neuron;

2.6.2 Mortalität

ENDEAR

Zeit bis zum Tod

In der Nusinersen-Gruppe waren 13 Patienten (16 %), in der Vergleichsgruppe 16 Patienten (39 %) zum Zeitpunkt der Datenanalyse verstorben. Das Risiko zu versterben verringerte sich für Patienten unter der Behandlung mit Nusinersen signifikant um 63,3 % (Hazard Ratio (HR): 0,367; 95%-KI: [0,18;0,76]; p=0,0074).

Tabelle 17: Ergebnisse zum Endpunkt Zeit bis zum Tod in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾	Patienten mit Ereignis n (%)	Median ²⁾ [95%-KI] (Wochen)	N ¹⁾	Patienten mit Ereignis n (%)	Median ²⁾ [95%-KI] (Wochen)	HR ³⁾ [95%-KI] p-Wert
Zeit bis zum Tod ⁴⁾	41	16 (39)	n. a. [23,1;n.a.]	80	13 (16)	n. a. [n.a.;n.a.]	0,367 [0,18;0,76] 0,0074

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ 50. Perzentil Median in Wochen; basierend auf Kaplan-Meier-Schätzer.

³⁾ Basierend auf einem Cox-Proportional-Hazard-Modell mit Behandlung (BSC, Nusinersen) als kategoriale Variable.

⁴⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC. [1]

Abkürzungen: BSC: Best Supportive Care; HR: Hazard Ratio; ITT: Intention-to-Treat; KI: Konfidenzintervall; n.a.: not applicable/ unzutreffend; pU: pharmazeutischer Unternehmer.

CHERISH

Todesfälle

Todesfälle wurden in der CHERISH-Studie im Rahmen der Sicherheitsbewertung erhoben. Während der Studie ist weder unter Nusinersen noch unter BSC ein Patient verstorben.

2.6.3 Morbidität

ENDEAR

Zeit bis zum Tod oder zur dauerhaften Beatmung

In der Interventionsgruppe war das Risiko zu versterben oder das Erfordernis dauerhaft beatmet zu werden um 53 % statistisch signifikant verringert im Vergleich zur Kontrollgruppe (HR: 0,47; 95%-KI: [0,28;0,78]; p=0,0037). Die mediane Zeit bis zum Tod oder zur dauerhaften Beatmung betrug unter der Scheinbehandlung 22,6 Wochen (95%-KI [13,6;31,3]). Unter Nusinersen trat bei 39 % der Patienten ein Ereignis auf, wobei die mediane Zeit bis zum Ereignis nicht ermittelt werden konnte (n.a.; 95%-KI: [36,3;n.a.]).

Es bleibt unklar, warum sich die Ereignisse der Teilendpunkte Zeit bis zum Tod (n=16) und Zeit bis zur dauerhaften Beatmung (n=13) in der Kontrollgruppe nicht mit der Anzahl Ereignisse des kombinierten Endpunktes (n=28) decken.

Tabelle 18: Ergebnisse zum Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾	Patienten mit Ereignis n (%)	Median ²⁾ [95%-KI] (Wochen)	N ¹⁾	Patienten mit Ereignis n (%)	Median ²⁾ [95%-KI] (Wochen)	HR ³⁾ [95%-KI] p-Wert
Zeit bis zum Tod oder zur dauerhaften Beatmung ⁴⁾	41	28 (68)	22,6 [13,6;31,3]	80	31 (39)	n. a. [36,3;n.a.]	0,47 [0,28;0,78] 0,0037

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ 50. Perzentil Median in Wochen; basierend auf Kaplan-Meier-Schätzer.

³⁾ Basierend auf einem Cox-Proportional-Hazard-Modell mit Behandlung (BSC, Nusinersen) als kategoriale Variable.

⁴⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; HR: Hazard Ratio; ITT: Intention-to-Treat; KI: Konfidenzintervall; n.a.: not applicable/unzutreffend; pU: pharmazeutischer Unternehmer.

Zeit bis zur dauerhaften Beatmung

Die Teilkomponente Zeit bis zur dauerhaften Beatmung wurde ergänzend zum zweiten primären Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung ausgewertet. Als dauerhafte Beatmung galt dabei eine Beatmung von ≥ 16 Stunden/Tag durchgehend für mehr als 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses oder Tracheotomie.

Obwohl zu Studienbeginn ein größerer Anteil Studienteilnehmer in der Interventionsgruppe eine Pneumonie bzw. respiratorische Symptome aufwiesen (siehe Tabelle 16), zeigt sich in der Analyse des Endpunktes, dass unter Nusinersen anteilig weniger Patienten einer dauerhaften Beatmung bedurften (Nusinersen 23 % vs. BSC 32 %). Zu Studienbeginn war für keinen Studienteilnehmer eine dauerhafte Beatmung erforderlich. Der Unterschied zwischen den Vergleichsarmen ist nicht statistisch signifikant.

Tabelle 19: Ergebnisse zum Endpunkt Zeit bis zur dauerhaften Beatmung in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾	Patienten mit Ereignis n (%)	Median ²⁾ [95%-KI] (Wochen)	N ¹⁾	Patienten mit Ereignis n (%)	Median ²⁾ [95%-KI] (Wochen)	HR ³⁾ [95%-KI] p-Wert
Zeit bis zur dauerhaften Beatmung ⁴⁾	41	13 (32)	n.a. [22,6; n.a.]	80	18 (23)	n. a. [n.a.; n.a.]	0,56 [0,27;1,15] 0,1142

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ 50. Perzentil Median in Wochen; basierend auf Kaplan-Meier-Schätzer.

³⁾ Basierend auf einem Cox-Proportional-Hazard-Modell mit Behandlung (BSC, Nusinersen) als kategoriale Variable. Tod wurde nicht als Ereignis gezählt und Patienten die verstarben wurden zum Zeitpunkt des Todes zensiert.

⁴⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; HR: Hazard Ratio; ITT: Intention-to-Treat; KI: Konfidenzintervall; n.a.: not applicable/unzutreffend; pU: pharmazeutischer Unternehmer.

Anteil der HINE-Responder (Subskala 2)

Die Analyse des primären Endpunktes Responder mit Erreichen eines motorischen Meilensteins mittels HINE (nach Definition, siehe Abschnitt 2.5.3) basieren auf den Studienvisiten an Tag 183, 302 und 394 im Vergleich zu Baseline für jeden Studienteilnehmer. Fehlende Daten wurden imputiert. Zum Zeitpunkt der Analyse erreichten 51 % der Patienten unter einer Behandlung mit Nusinersen einen motorischen Meilenstein im Vergleich zu keinem Patienten unter einer Behandlung mit BSC. Die Responserate war in der Interventionsgruppe signifikant erhöht (Relatives Risiko (RR): 43,04; 95%-KI: [2,72;682,32]; $p=0,0076$). Die Ergebnisse sind konsistent mit den Ergebnissen der Interimsanalyse.

Es gilt zu berücksichtigen, dass keine Untersuchungen zur Validität des Erhebungsinstrumentes und klinischen Relevanzschwelle vorliegen (siehe Abschnitt 2.5.3). Offen bleibt, ob ein Responder eine dauerhafte Response zeigten musste oder dies einmalig ausreichend war, um in die Responderanalyse aufgenommen zu werden. Eine ergänzende Analyse erfolgte für die ES-Population (Tabelle 21). Die Ergebnisse sind konsistent mit den Ergebnissen der ITT-Population (Tabelle 20). Die durchgeführten Sensitivitätsanalysen stützen ebenfalls die Robustheit der Ergebnisse des primären Endpunktes (siehe auch 2.5.4).

Tabelle 20: Ergebnisse zum Endpunkt Anteil der HINE-Responder (Subskala 2) in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC (N=41) ¹⁾²⁾	Nusinersen (N=80) ¹⁾²⁾	Nusinersen vs. BSC
ENDEAR	Patienten mit Ereignis n (%)	Patienten mit Ereignis n (%)	RR ³⁾ [95%-KI] p-Wert
HINE-Responder (Subskala 2) ³⁾	0 (0)	41 (51)	43,04 [2,72;682,32] 0,0076

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Die Analyse umfasst alle Patienten der ITT (N=121): 110 Patienten des ES plus 11 Patienten, die aufgrund von Tod oder frühzeitiger Beendigung nicht Teil des ES sind. Im Median waren die 11 Patienten 140 Tage (105–148) in der Studie.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; HINE: Hammersmith Infant Neurological Examination; ITT: Intention-to-Treat; KI: Konfidenzintervall; n.a.: not applicable/unzutreffend; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko.

Tabelle 21: Ergebnisse zum Endpunkt Anteil der HINE-Responder (Subskala 2) in der Studie ENDEAR – ES-Population

Studie Endpunkt	BSC (N=37) ¹⁾²⁾	Nusinersen (N=73) ¹⁾²⁾	Nusinersen vs. BSC
ENDEAR	Patienten mit Ereignis n (%)	Patienten mit Ereignis n (%)	RR ³⁾ [95%-KI] p-Wert
HINE-Responder (Subskala 2) ⁴⁾	0 (0)	37 (51)	38,51 [2,43; 610,14] 0,0096

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Die Ergebnisse basieren auf allen verfügbaren Daten des ES (Patienten mit einer Verbesserung gemäß definierten Responsekriterium zum Zeitpunkt Tag 183, 302 oder 394). Das ES ist definiert als Teilmenge der ITT mit mindestens einer Tag-183-Bewertung. Aufgrund des frühzeitigen Studienabschlusses sind 11 Patienten nicht Teil des ES, da sie nicht lange genug in der Studie waren (keine Tag-183-Bewertung). Die ITT-Population besteht im Nusinersen-Arm aus 80 Patienten und im BSC-Arm aus 41 Patienten.

³⁾ Das RR wurde auf Basis der Vierfeldertafel berechnet. Falls in einem Studienarm kein Ereignis beobachtet wurde (Nullzellen), wurde entsprechend der methodischen Empfehlungen des IQWiG eine Kontinuitätskorrektur mit einem Korrekturwert von 0,5 für jede der Zellen angewendet.

⁴⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; ES: Efficacy Set; HINE: Hammersmith Infant Neurological Examination; ITT: Intention-to-Treat; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko.

Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline

Ergänzend legt der pU eine Auswertung des HINE-Scores (Subskala 2) im Vergleich zu Baseline vor. Die Auswertung basiert auf den Erhebungszeitpunkten Tag 183, 302 und 394 im Vergleich zu Baseline für jeden Studienteilnehmer. Fehlende Daten wurden imputiert. Eine Auswertung erfolgte sowohl für die ITT- als auch für die ES-Population der ENDEAR-Studie.

Die Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline war im Nusinersen-Behandlungsarm (MW (SD): 3,43 (4,12)) deutlich höher als im Vergleichsarm (MW (SD): -0,17 (1,26)). Der Unterschied zwischen den Vergleichsarmen war statistisch signifikant zugunsten von Nusinersen (Differenz MW: 3,60; 95%-KI: [2,29;4,90]; $p < 0,0001$). Die Effektstärke mit einem Hedges' g von 1,05 (95%-KI: [0,64;1,44]; $p < 0,0001$) wird als relevant interpretiert. Die Auswertung der ES-Population zeigt vergleichbare Ergebnisse (Tabelle 23).

Tabelle 22: Ergebnisse zum Endpunkt Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾²⁾	n	MW (SD)	N ¹⁾²⁾	n	MW (SD)	
ENDEAR							Differenz MW [95%-KI]³⁾ p-Wert Hedges' g [95%-KI]³⁾ p-Wert Hedges' g
Veränderung des HINE-Scores (Subskala 2) von Baseline zu Tag 183, 302 oder 394 ⁴⁾	41	41	-0,17 (1,26)	80	80	3,43 (4,12)	3,60 [2,29;4,90] <0,0001 1,05 [0,64;1,44] <0,0001

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Die Analyse umfasst alle Patienten der ITT (N=121): 110 Patienten des ES plus 11 Patienten, die aufgrund von Tod oder frühzeitiger Beendigung nicht Teil des ES sind. Im Median waren die 11 Patienten 140 Tage (105–148) in der Studie.

³⁾ Für Patienten, die keine Bewertung an Tag 183 und später aufwiesen, wurde die letzte verfügbare Bewertung verwendet.

⁴⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; ES: Efficacy Set; HINE: Hammersmith Infant Neurological Examination; ITT: Intention-to-Treat; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

Tabelle 23: Ergebnisse zum Endpunkt Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline in der Studie ENDEAR – ES-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾²⁾	n	MW (SD)	N ¹⁾²⁾	n	MW (SD)	
ENDEAR							Differenz MW [95%-KI]³⁾ p-Wert Hedges' g [95%-KI]³⁾ p-Wert Hedges' g
Veränderung des HINE-Scores (Subskala 2) von Baseline zu Tag 183, 302 oder 394 ⁴⁾	37	20	-0,20 (1,44)	73	58	4,38 (4,35)	4,58 [2,60;6,56] <0,0001 1,19 [0,65;1,73] <0,0001

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Von den 110 Patienten im ES verstarben 29 Patienten (13 Patienten (18 %) in der Nusinersen-Gruppe, 16 Patienten (43 %) unter BSC). Drei Patienten schieden aus der Studie aus anderen Gründen aus (2 Patienten (3 %) der Nusinersen-Gruppe, 1 Patient (3 %) unter BSC); diese wurden nicht in der Analyse berücksichtigt.

³⁾ Für Patienten, die keine Bewertung an Tag 183 und später aufwiesen, wurde die letzte verfügbare Bewertung verwendet.

⁴⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; ES: Efficacy Set; HINE: Hammersmith Infant Neurological Examination; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

Anteil der CHOP-INTEND-Responder

Die Ergebnisse zum CHOP INTEND werden unterstützend zum primären Endpunkt HINE berichtet (vgl. 2.5.3). Ein Responder war definiert als eine Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline von ≥ 4 Punkten, wobei die Responseschwelle als fraglich erachtet wird. Insgesamt erreichten statistisch signifikant deutlich mehr Probanden unter Nusinersen (73 %) die Responseschwelle als im Vergleichsarm (2 %) (RR: 29,73; 95%-KI: [4,27;206,98]; $p=0,0006$). Die Auswertungen der ES-Population sind konsistent mit den Ergebnissen der ITT-Population.

Die Sensitivitätsanalysen bestätigen die Robustheit der Ergebnisse der Hauptanalyse (vgl. 2.5.4).

Tabelle 24: Ergebnisse zum Endpunkt Anteil der CHOP-INTEND-Responder in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC (N=41) ¹⁾²⁾	Nusinersen (N=80) ¹⁾²⁾	Nusinersen vs. BSC
ENDEAR	Patienten mit Ereignis n (%)	Patienten mit Ereignis n (%)	RR [95%-KI] p-Wert
CHOP-INTEND-Responder ³⁾	1 (2)	58 (73)	29,73 [4,27;206,98] 0,0006

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Die Analyse umfasst alle Patienten der ITT (N=121): 110 Patienten des ES plus 11 Patienten, die aufgrund von Tod oder frühzeitiger Beendigung nicht Teil des ES sind. Im Median waren die 11 Patienten 140 Tage (105–148) in der Studie.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease; ES: Efficacy Set; ITT: Intention-to-Treat; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko.

Tabelle 25: Ergebnisse zum Endpunkt Anteil der CHOP-INTEND-Responder in der Studie ENDEAR – ES-Population

Studie Endpunkt	BSC (N=37) ¹⁾²⁾	Nusinersen (N=73) ¹⁾²⁾	Nusinersen vs. BSC
ENDEAR	Patienten mit Ereignis n (%)	Patienten mit Ereignis n (%)	RR [95%-KI] p-Wert
CHOP-INTEND-Responder ³⁾	1 (3)	52 (71)	26,36 [3,79;183,18] 0,0009

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Die Ergebnisse basieren auf allen verfügbaren Daten des ES (Patienten mit einer Verbesserung gemäß definierten Responsekriterium zum Zeitpunkt Tag 183, Tag 302 oder Tag 394). Das ES ist definiert als Teilmenge der ITT mit mindestens einer Tag-183-Bewertung. Aufgrund des frühzeitigen Studienabschlusses sind 11 Patienten nicht Teil des ES, da sie nicht lange genug in der Studie waren (keine Tag-183-Bewertung). Die ITT besteht im Nusinersen-Arm aus 80 und im BSC-Arm aus 41 Patienten.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease; ES: Efficacy Set; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko.

Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline

Ergänzend legt der pU eine Auswertung des CHOP INTEND im Vergleich zu Baseline vor. Zu Studienbeginn waren die CHOP-INTEND-Gesamtwerte in beiden Studienarmen vergleichbar: Nusinersen 8,0 bis 48,5 Punkte (Median: 27,50) und BSC 10,5 bis 50,5 Punkte (Median: 28,50). Eine Auswertung erfolgte sowohl für die ITT- als auch für die ES-Population der ENDEAR-Studie.

Die Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline fiel im Nusinersen-Behandlungsarm (MW (SD): 9,46 (8,41)) deutlich höher aus als im Vergleichsarm (MW (SD): -9,72 (8,73)). Der Unterschied zwischen den Vergleichsarmen ist statistisch signifikant (Differenz MW: 18,66; 95%-KI: [15,42;21,90]; $p < 0,0001$) bei nachweislich großer Effektstärke (Hedges' g : 2,19; (95%-KI: [1,72;2,66]; $p < 0,0001$) zugunsten von Nusinersen. Es ergeben sich keine abweichenden Ergebnisse in der ES-Population (Tabelle 27).

Tabelle 26: Ergebnisse zum Endpunkt Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾²⁾	n	MW (SD)	N ¹⁾²⁾	n	MW (SD)	
ENDEAR							Differenz MW [95%-KI] ³⁾ p-Wert Hedges' g [95%-KI] ³⁾ p-Wert Hedges' g
Veränderung des CHOP-INTEND-Gesamtwertes von Baseline zu Tag 183, 302, 394 ⁴⁾	41	41	-9,20 (8,73)	80	80	9,46 (8,41)	18,66 [15,42;21,90] <0,0001 2,19 [1,72;2,66] <0,0001

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Die Analyse umfasst alle Patienten der ITT (N=121): 110 Patienten des ES plus 11 Patienten, die aufgrund von Tod oder frühzeitiger Beendigung nicht Teil des ES sind. Im Median waren die 11 Patienten 140 Tage (105–148) in der Studie.

³⁾ Für Patienten, die keine Bewertung an Tag 183 und später aufwiesen, wurde die letzte verfügbare Bewertung verwendet.

⁴⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease; ITT: Intention-to-treat; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

Tabelle 27: Ergebnisse zum Endpunkt Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline in der Studie ENDEAR – ES-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾²⁾	n	MW (SD)	N ¹⁾²⁾	n	MW (SD)	
ENDEAR							Differenz MW [95%-KI] p-Wert Hedges' g [95%-KI] p-Wert Hedges' g
Veränderung des CHOP-INTEND-Gesamtwertes von Baseline zu Tag 183, 302, 394 ³⁾	37	20	-12,85 (8,88)	73	58	12,24 (7,89)	25,09 [20,88;29,30] <0,0001 3,08 [2,37;3,78] <0,0001

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Von den 110 Patienten im ES verstarben 29 Patienten (13 Patienten (18 %) in der Nusinersen-Gruppe, 16 Patienten (43 %) unter BSC). Drei Patienten schieden aus der Studie aus anderen Gründen aus (2 Patienten (3%) der Nusinersen-Gruppe, 1 Patient (3 %) unter BSC); diese wurden nicht in der Analyse berücksichtigt.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease; ES: Efficacy Set; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

CHERISH

Veränderung des HFMSE im Vergleich zu Baseline

Mittels HFMSE wurde der primäre Endpunkt der CHERISH-Studie, die Änderung der motorischen Funktion nach 15 Monaten im Vergleich zu Baseline, untersucht. Im Vergleich zu Baseline zeigt sich eine signifikante Veränderung des HFMSE-Scores nach 15 Monaten zugunsten einer Behandlung mit Nusinersen (Differenz LS MW: 4,42; 95%-KI: [2,46;6,38]; p<0,0001). Das 95%-KI

des Hedges' g liegt dabei vollständig oberhalb der Irrelevanzschwelle von 0,2, womit sich ein relevanter Effekt ableiten lässt.

Eine im Studienbericht dargestellte, ergänzende Auswertung in der ES-Population (Differenz LS MW: 5,2; 95%-KI: [3,4;7,1]) sowie die durchgeführten Sensitivitätsanalysen (vgl. 2.5.4) zum primären Endpunkt unterstützen die Ergebnisse der Hauptanalyse in der ITT-Population.

Tabelle 28: Ergebnisse zum Endpunkt Veränderung des HFMSE-Scores im Vergleich zu Baseline in der Studie CHERISH – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾	n (%)	MW (SD)	N ¹⁾	n (%)	MW (SD)	Differenz LS MW [95%-KI] ²⁾ p-Wert Hedges' g [95%-KI] ²⁾ p-Wert Hedges' g
Veränderung des HFMSE-Scores von Baseline zu Monat 15 ³⁾	42	34 (81)	-0,76 (4,7)	84	66 (79)	3,97 (4,86)	4,42 [2,46;6,38] <0,0001 0,99 [0,55;1,42] <0,0001

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Basierend auf einem ANCOVA-Modell adjustiert für Baseline. Die Schätzer werden aus der Anpassung des ANCOVA-Modells für jeden der imputierten Datensätze bestimmt.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; HFMSE: Hammersmith Functional Motor Scale Expanded; ITT: Intention-to-Treat; KI: Konfidenzintervall; LS: Least Square; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

Anteil HFMSE-Responder im Vergleich zu Baseline

Untersucht wurde eine Veränderung um ≥ 3 Punkte nach 15 Monaten im Vergleich zu Baseline. Das Erreichen der Responseschwelle war in der Nusinersen-Gruppe mehr als zweimal höher als in der Scheinkontrollgruppe. Das Ergebnis ist mit einem RR von 2,17 (95%-KI: [1,23;3,81]) und einem p-Wert von 0,0073 statistisch signifikant.

Die Sensitivitätsanalysen des Endpunktes (inklusive einer Auswertung der ES-Population) zeigen gleichgerichtete Effekte und bestätigen die Ergebnisse der Hauptanalyse.

Tabelle 29: Ergebnisse zum Endpunkt Anteil der HFMSE-Responder in der Studie CHERISH – ITT-Population

Studie Endpunkt	BSC (N=42) ¹⁾	Nusinersen (N=84) ¹⁾	Nusinersen vs. BSC
CHERISH	Patienten mit Ereignis n (%)	Patienten mit Ereignis n (%)	RR ²⁾ [95%-KI] p-Wert
HFMSE-Responder nach 15 Monaten im Vergleich zu Baseline ³⁾	34 (81)	66 (79)	2,17 [1,23;3,81] 0,0073

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Bei fehlenden Werten basiert die Berechnung des RR und der zugehörigen Statistik in dieser Tabelle auf imputierten Daten.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; HFMSE: Hammersmith Functional Motor Scale Expanded; ITT: Intention-to-Treat; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko.

Häufigkeit schwerwiegender respiratorischer Ereignisse

ENDEAR

Für den Endpunkt Häufigkeit schwerwiegender respiratorischer Ereignisse zeigt sich kein Unterschied zwischen den Behandlungsarmen (Tabelle 30).

Tabelle 30: Ergebnisse zum Endpunkt Häufigkeit schwerwiegender respiratorischer Ereignisse in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾	Gesamtzahl Ereignisse	MW (SD)	N ¹⁾	Gesamtzahl Ereignisse	MW (SD)	
ENDEAR							Rate Ratio²⁾ [95%-KI] p-Wert
Häufigkeit schwerwiegender respiratorischer Ereignisse ³⁾	41	117	2,9 (2,93)	80	238	3,0 (3,28)	0,81 [0,53;1,25] 0,35

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Auswertung basierend auf einer negativen Binomialregression mit der Behandlung (BSC, Nusinersen) als kategoriale Kovariate.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

CHERISH

Der Endpunkt Häufigkeit schwerwiegender respiratorischer Ereignisse ist post hoc ausgewertet worden. Es zeigt sich kein Unterschied zwischen den Behandlungsarmen (Tabelle 31).

Tabelle 31: Ergebnisse zum Endpunkt Häufigkeit schwerwiegender respiratorischer Ereignisse in der Studie CHERISH – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC
	N ¹⁾	Gesamtzahl Ereignisse	MW (SD)	N ¹⁾	Gesamtzahl Ereignisse	MW (SD)	
CHERISH							Rate Ratio²⁾ [95%-KI] p-Wert
Häufigkeit schwerwiegender respiratorischer Ereignisse ³⁾	42	14	0,3 (0,82)	84	11	0,1 (0,46)	0,40 [0,14;1,14] 0,09

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Auswertung basierend auf einer negativen Binomialregression mit der Behandlung (BSC, Nusinersen) als kategoriale Kovariate.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

Häufigkeit von Hospitalisierungen

ENDEAR

Für den Endpunkt Häufigkeit von Hospitalisierungen zeigt sich kein Unterschied zwischen den Behandlungsarmen (Tabelle 32).

Tabelle 32: Ergebnisse zum Endpunkt Häufigkeit von Hospitalisierungen in der Studie ENDEAR – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC Rate Ratio ²⁾ [95%-KI] p-Wert
	N ¹⁾	Gesamtzahl Ereignisse	MW (SD)	N ¹⁾	Gesamtzahl Ereignisse	MW (SD)	
ENDEAR							
Häufigkeit von Hospitalisierungen ³⁾	41	119	2,9 (2,29)	80	264	3,3 (2,75)	0,91 [0,66;1,26] 0,57

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Auswertung basierend auf einer negativen Binomialregression mit der Behandlung (BSC, Nusinersen) als kategoriale Kovariate.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

Häufigkeit von krankheitsbedingten Hospitalisierungen

CHERISH

Für den Endpunkt Häufigkeit von krankheitsbedingten Hospitalisierungen zeigt sich ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen zugunsten von Nusinersen (Rate Ratio: 0,35; 95%-KI: [0,14;0,87]; p=0,02; vgl. Tabelle 33).

Tabelle 33: Ergebnisse zum Endpunkt Häufigkeit von krankheitsbedingten Hospitalisierungen in der Studie CHERISH – ITT-Population

Studie Endpunkt	BSC			Nusinersen			Nusinersen vs. BSC Rate Ratio ²⁾ [95%-KI] p-Wert
	N ¹⁾	Gesamtzahl Ereignisse	MW (SD)	N ¹⁾	Gesamtzahl Ereignisse	MW (SD)	
CHERISH							
Häufigkeit von krankheitsbedingten Hospitalisierungen ³⁾	42	16	0,4 (0,82)	84	11	0,1 (0,40)	0,35 [0,14; 0,87] 0,02

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Auswertung basierend auf einer negativen Binomialregression mit der Behandlung (BSC, Nusinersen) als kategoriale Kovariate.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1].

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; KI: Konfidenzintervall; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

2.6.4 Lebensqualität

Veränderung des PedsQL im Vergleich zu Baseline

In der CHERISH-Studie ist die gesundheitsbezogene Lebensqualität anhand der Generic-Core-Skala und dem Neuromuskulären Modul 3.0 des PedsQL (Version 4.0) untersucht worden. Die Erhebung des PedsQL erfolgte zum Zeitpunkt des Screenings (zwei Baseline-Messungen) sowie nach 3, 6, 9, 12 und 15 Monaten.

Aufgrund der geringen Rücklaufquoten (< 70 %) in beiden Studienarmen, maximal 25 % im Nusinersen-Arm und maximal 19 % im Kontrollarm, ist seitens des pU auf eine ausführliche Ergebnisdarstellung in Modul 4 des Herstellerdossiers verzichtet worden.

Die Rücklaufquoten sind bereits nach 3 Monaten so gering (24 % Nusinersen-Arm vs. 19 % BSC-Arm) – was sich bis zum letzten Erhebungszeitpunkt nach 15 Monaten (14 % Nusinersen-Arm vs. 12 % BSC-Arm) fortsetzt –, dass die Aussagekraft der Ergebnisse als unzureichend eingeschätzt wird. Nach abschließender Sichtung der Studienergebnisse wird der Einschätzung des pU gefolgt.

2.6.5 Sicherheit

Von den 122 Probanden der ENDEAR-Studie, die in diese Studie randomisiert wurden, erhielten 121 eine Studienbehandlung (N=80 Nusinersen und N=41 BSC) und entsprechen der Sicherheitspopulation. Für die CHERISH-Studie wurden insgesamt 179 Probanden gescreent, von denen 126 im Verhältnis 2:1 randomisiert wurden (N=84 Nusinersen und N=42 BSC) und die mindestens eine Dosis der Studienbehandlung erhielten. Da alle Probanden jeweils die Studienbehandlung erhielten der sie zugewiesen wurden, sind die Sicherheits- und die ITT-Population identisch.

Die Zahl der schweren Atemwegereignisse und die Anzahl und Dauer der Hospitalisierungen sind in der Studie ENDEAR sowie die Häufigkeit von krankheitsbedingten Hospitalisierungen in der Studie CHERISH als Wirksamkeitsendpunkte beschrieben und entsprechend unter 2.6.3 verortet.

Unerwünschte Ereignisse

ENDEAR

Unter der Studienbehandlung trat bei insgesamt 96 % der Probanden im Nusinersen-Arm und bei 98 % im BSC-Arm ein unerwünschtes Ereignis auf. Dabe war die Studiendauer und somit auch Expositionsdauer für die Patienten im Interventionsarm wesentlich länger als im Kontrollarm. Für die Endpunkte schwere UE, SUE und Therapieabbruch aufgrund von UE zeigt sich ein statistisch signifikanter Behandlungsunterschied zugunsten von Nusinersen. Das Relative Risiko war in der Interventionsgruppe entsprechend um 30 %, 20 % und 58 % geringer als in der Kontrollgruppe (schwere UE: RR=0,70; 95%-KI: [0,55;0,89]; p=0,004 / SUE: RR=0,80; 95%-KI: [0,70;0,92]; p=0,002 / Therapieabbruch aufgrund von UE: RR=0,42; 95%-KI [0,22;0,78]; p=0,006 / Tabelle 34).

Der Prozentsatz der Betroffenen von schweren Ereignissen war in der Interventionsgruppe niedriger (Nusinersen vs. BSC: 56 vs. 80 %). Viele der am häufigsten gemeldeten UE wurden als schwer bewertet: Atemstillstand (18 vs. 37 %), akuter Atemstillstand (14 vs. 22 %), Atemnot (15 vs. 20 %), Pneumonie (13 vs. 12 %) und Atelektase (8 vs. 10 %).

Ein geringerer Anteil Patienten in der Interventionsgruppe hatte ein SUE (Nusinersen vs. Kontrolle: 76 vs. 95 %). Die am häufigsten (≥ 10 %) berichteten SUE in der Nusinersen- oder Kontrollgruppe waren:

- Atemnot (26 vs. 20 %);
- Atemversagen (25 vs. 39 %);
- Pneumonie (24 vs. 12 %);
- Akuter Atemstillstand (14 vs. 22 %);
- Atelektase (18 vs. 10 %);
- Aspirationspneumonie (10 vs. 12 %);
- Atemstillstand (6 vs. 10 %);
- Virale Infektion der oberen Atemwege (4 vs. 15 %);
- Retention des Bronchialsekrets (1 vs. 12 %);
- Herz- und Atemstillstand (6 vs. 12 %);

Neben den 29 Probanden, die ein SUE mit tödlichem Ausgang hatten, gab es keine weiteren Mitglieder in der Sicherheitspopulation, die wegen eines UE die Behandlung eingestellt oder sich aus der Studie zurückgezogen haben.

Die vorliegenden Daten zur Sicherheit sind nicht ausschließlich als Nebenwirkungen anzusehen, da diese auch Ereignisse der Grunderkrankung umfassen.

Tabelle 34: Ergebnisse zum Endpunkt Unerwünschte Ereignisse in der Studie ENDEAR – SAS-Population

Studie Endpunkt	BSC (N=41) ¹⁾²⁾	Nusinersen (N=80) ¹⁾²⁾	Nusinersen vs. BSC
ENDEAR	Patienten mit Ereignis n (%)	Patienten mit Ereignis n (%)	RR ³⁾ [95%-KI] p-Wert
Unerwünschte Ereignisse ³⁾	40 (98)	77 (96)	-
- mild	1 (2)	7 (9)	3,59 [0,46;28,18] 0,22
- moderat oder schwer	39 (95)	70 (88)	0,92 [0,83;1,03] 0,13
- schwer	33 (80)	45 (56)	0,70 [0,55;0,89] 0,004
Schwerwiegende unerwünschte Ereignisse	39 (95)	61 (76)	0,80 [0,70;0,92] 0,002
Therapieabbruch aufgrund unerwünschter Ereignisse	16 (39)	13 (16)	0,42 [0,22;0,78] 0,006

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Die Analyse umfasst alle Patienten des SAS (N=121); SAS- und ITT-Population sind identisch.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SAS: Safety-Analysis-Set.

CHERISH

Für die Sicherheitsendpunkte UE, milde, moderate und schwere UE, SUE und Therapieabbrüche aufgrund von UE zeigt sich jeweils kein statistisch signifikanter Unterschied zwischen den Vergleichsarmen (Tabelle 35).

Die häufigsten SUE waren Pneumonie (Nusinersen vs. BSC: 2 vs. 14 %), virale Pneumonie (4 vs. 0 %) und Atemnot (2 vs. 5 %). Virale Pneumonie war das einzige SUE, dass von mehr als 2 Probanden in der Nusinersen-Gruppe gemeldet wurde.

Die vorliegenden Daten zur Sicherheit sind nicht ausschließlich als Nebenwirkungen anzusehen, da diese auch Ereignisse der Grunderkrankung umfassen.

Tabelle 35: Ergebnisse zum Endpunkt Unerwünschte Ereignisse in der Studie CHERISH – SAS-Population

Studie Endpunkt	BSC (N=42) ¹⁾²⁾	Nusinersen (N=84) ¹⁾²⁾	Nusinersen vs. BSC
CHERISH	Patienten mit Ereignis n (%)	Patienten mit Ereignis n (%)	RR³⁾ [95%-KI] p-Wert
Unerwünschte Ereignisse ³⁾	42 (100)	78 (93)	-
- mild	19 (45)	39 (46)	1,03 [0,69;1,54] 0,90
- moderat oder schwer	23 (55)	39 (46)	0,85 [0,59;1,21] 0,37
- schwer	3 (7)	4 (5)	0,67 [0,16;2,84] 0,58
Schwerwiegende unerwünschte Ereignisse	12 (29)	14 (17)	0,58 [0,30;1,15] 0,12
Therapieabbruch aufgrund unerwünschter Ereignisse	0 (0)	0 (0)	-

¹⁾ Anzahl Patienten in der Auswertung.

²⁾ Die Analyse umfasst alle Patienten des SAS (N=126); SAS- und ITT-Population sind identisch.

³⁾ Quelle: Zusätzliche Analysen des pU zu Nusinersen vs. BSC [1], ohne Angabe des verwendeten statistischen Modells.

Abkürzungen: BSC: Best Supportive Care; ITT: Intention-to-Treat; KI: Konfidenzintervall; pU: pharmazeutischer Unternehmer; RR: Relatives Risiko; SAS: Safety-Analysis-Set.

Häufige UE nach Systemorganklassen und Preferred Term und UE von besonderem Interesse

ENDEAR

Die am häufigsten aufgetretenen unerwünschten Ereignisse nach Systemorganklassen und Preferred Term waren Atemwegserkrankungen und/oder Infektionen: Infektion der oberen Atemwege (Nusinersen vs. BSC: 30 vs. 22 %), Nasopharyngitis (19 vs. 10 %), Pneumonie (29 vs. 17 %), Bronchitis (8 vs. 2 %), Stauung der oberen Atemwege (8 vs. 2 %), Bronchitis viral (6 vs. 0 %), Grippe (6 vs. 0 %), Bronchiolitis und virale Infektionen (jeweils 10 vs. 7 %), virale Infektion der oberen Atemwege (10 vs. 17 %), Infektion der Atemwege (11 vs. 5 %), Rhinovirusinfektionen (13 vs. 15 %), Atemnot (26 vs. 29 %), Dyspnoe (9 vs. 15 %), Atemstillstand (25 vs. 39 %), akutes Atem-versagen (14 vs. 24 %), Atelektase (23 vs. 29 %), Husten (11 vs. 20 %). Andere häufig gemeldete Ereignisse beinhalten Verstopfung (35 vs. 22 %), gastroösophageale Refluxkrankheit (13 vs. 20 %), Erbrechen (18 vs. 20 %), Zahnen (18 vs. 7 %), Diarrhö (14 vs. 17 %), Verstopfung (35 vs. 22 %), Dysphagie (11 vs. 20 %), Pyrexie (56 vs. 59 %), Abnahme der Sauerstoffsättigung (13 vs. 24 %) und Aspirationspneumonie (11 vs. 17 %).

Als UE von besonderem Interesse gilt, basierend auf präklinischen Befunden, Epilepsie. Es gab keinen Fall von Epilepsie, aber einen Fall von Krampfanfällen im Nusinersen-Arm im Zusammenhang mit einer Hirnverletzung, eine bekannte Ursache von provozierten Krampfanfällen.

CHERISH

Auch in der CHERISH-Studie waren die am häufigsten gemeldeten unerwünschten Ereignisse nach Systemorganklassen und Preferred Term respiratorische Erkrankungen oder Infektionen: Infektion der oberen Atemwege (Nusinersen vs. BSC: 30 vs. 45 %), Nasopharyngitis (24 vs. 36 %), Bronchitis (10 vs. 10 %) und Husten (25 vs. 21 %). Andere häufig aufgetretene Ereignisse waren: Gastroenteritis (12 vs. 19 %), virale Gastroenteritis (10 vs. 10 %), Influenza (10 vs. 7 %), Ohreninfektion (8 vs. 12 %), Pyrexie (43 vs. 36 %), Erbrechen (29 vs. 12 %), Diarrhö (10 vs. 7 %), Verstopfung (6 vs. 12 %), Kopfschmerzen (29 vs. 7 %) und Rückenschmerzen (25 vs. 0 %).

Es wurde kein UE von besonderem Interesse, d. h. keine Epilepsie oder Krampfanfälle, berichtet. Es gab ein Ereignis von Myoklonus des linken Ringfingers unter Nusinersen. Das Ereignis wurde behoben.

2.6.6 Subgruppenanalysen

Berücksichtigt werden Subgruppen, die eine Interaktion mit einem Interaktionstest $p=0,05$ aufweisen [26]. Aufgrund kleiner Subgruppengrößen sowie eines Ungleichgewichts zwischen den Subgruppen in den vorliegenden Studienpopulationen der Studien ENDEAR und CHERISH kann das Vorliegen einer Effektmodifikation nur mit Unsicherheit beurteilt werden kann.

ENDEAR

In der ENDEAR-Studie waren a priori Subgruppenanalysen für den primären Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung, motorische Meilensteine (mittels HINE), CHOP INTEND und Gesamtüberleben geplant:

1. Alter bei Symptombeginn (≤ 12 vs. > 12 Wochen),
2. Krankheitsdauer (≤ 12 vs. > 12 Wochen),
3. Geographische Region (Nordamerika vs. Europa vs. Asien-Pazifik-Raum).

Post hoc wurden für die Faktoren Alter bei der 1. Dosis (< 180 vs. ≥ 180 Tage) und Geschlecht für alle Endpunkte zusätzlich Subgruppenanalysen ausgewertet.

Für den kombinierten Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung zeigt sich eine Subgruppeninteraktion für die Faktoren Krankheitsdauer (≤ 12 Wochen), Geographische Region, Alter bei der 1. Dosis (< 180 Tage) und Geschlecht (Mädchen). Für den Teilendpunkt Zeit bis zur dauerhaften Beatmung sind die Ergebnisse mit den Subgruppen Krankheitsdauer und Geschlecht konsistent. Darüber hinaus zeigt sich eine Interaktion für die Subgruppe Region, wobei sich kein signikantes Ergebnis für eine bestimmte Subgruppe abzeichnet. Eine Interaktion zeigt sich ebenfalls beim Endpunkt Zeit bis zum Tod für die Sugruppe Alter bei Symptombeginn (≤ 12 Wochen zugunsten von Nusinersen). In der Analyse der CHOP-INTEND-Veränderung zu Baseline ist eine Interaktion für die Subgruppe Krankheitsdauer nachweisbar. In der Subgruppe Alter bei Symptombeginn (≤ 12 Wochen) ergibt sich ebenfalls ein signifikanter Vorteil für Nusinersen für den Endpunkt Therapieabbrüche aufgrund von UE.

Für keinen Faktor zeigte sich konstant über alle Endpunkte eine Subgruppeninteraktion.

CHERISH

A priori geplant waren folgende Subgruppenanalysen für den primären Endpunkt und die sekundären Endpunkte:

1. Krankheitsdauer: Zeit von Diagnosestellung SMA bis Screening in Monaten nach Terzilen (< 25 vs. \geq 25 Monate bis < 44 vs. \geq 44 Monate),
2. Alter bei Screening (< 6 vs. \geq 6 Jahre),
3. Geographische Region (Nordamerika vs. Europa vs. Asien-Pazifik-Raum).

Es wurden zusätzlich post hoc definierte Subgruppenanalysen vorgenommen, für den Faktor Geschlecht.

Es wurden keine Subgruppeninteraktionen für die in der Nutzenbewertung dargelegten Endpunkte identifiziert.

3 Methodische Anmerkungen und Bewertung der Unterlagen

3.1 Zulassungsstatus und Zulassungspopulation von Nusinersen (Spinraza®)

Gemäß Zulassung wird Nusinersen (Spinraza®) angewendet zur Behandlung der 5q-assoziierten spinalen Muskelatrophie (SMA) [12]. Nusinersen 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 besteht seit dem 30. Mai 2017.

Die vorliegende Nutzenbewertung zu Nusinersen basiert auf den Zulassungsstudien ENDEAR (CS3B) und CHERISH (CS4) [10,11]. Von einer Übertragbarkeit der Zulassungspopulation auf den deutschen Versorgungskontext ist auszugehen. In die zulassungsbegründeten Studien wurden Patienten mit infantiler SMA Typ I (ENDEAR) oder späterem SMA-Krankheitsbeginn (CHERISH) eingeschlossen. Untersucht wurde die Wirksamkeit und Sicherheit von Nusinersen 12 mg (5 ml) als intrathekale Bolusinjektion im Vergleich zu einer Scheinintervention. Sowohl die ENDEAR- als auch die CHERISH-Studie sind aufgrund des Erreichens vordefinierter Wirksamkeitskriterien vorzeitig beendet worden. Patienten konnten jeweils nach dem letzten Studienbesuch an einer offenen Langzeitstudie teilnehmen.

3.2 Design und Methodik der Studien ENDEAR und CHERISH

Bei der Zulassungsstudie ENDEAR (CS3B) handelt es sich um eine randomisierte, doppelblinde, scheininterventionskontrollierte, multizentrische Phase-III-Studie zur Beurteilung der Wirksamkeit und Sicherheit von Nusinersen im Vergleich zu einer Scheinintervention unter BSC bei Patienten mit infantiler SMA (Typ I, Säuglinge < 7 Monate alt). Eine Stratifizierung der Studienpopulation wurde hinsichtlich des Faktors Krankheitsdauer (≤ 12 vs. > 12 Wochen) vorgenommen. Es war eine Interimsanalyse (Zeitpunkt 15. Juni 2016) für den primären Endpunkt vorgesehen. Voraussetzung war, dass bis dahin 78 Patienten die Studienvisite an Tag 183 absolviert haben. Insgesamt wurden 121 Probanden im Verhältnis 2:1 randomisiert und den Behandlungsarmen zugeordnet (Nusinersen N=80 vs. Kontrolle N=41). Nusinersen wurde in einer altersadjustierten Dosierung (ab einem Alter von 24 Monaten 12 mg (5 ml)) gegenüber der Scheinintervention als intrathekale Bolusinjektion an den Studientagen 1, 15, 29, 64 (Aufsättigung), 183 und 302 (Erhaltung), über eine 10-monatige Behandlungsphase und eine 3-monatige Follow-up-Phase (bis Tag 394 \pm 7 Tage) verabreicht. Aufgrund der langen Halbwertszeit von Nusinersen wurde die Behandlung während des Follow-ups fortgesetzt. Fraglich bleibt, warum im Studienbericht trotz weiterführender Behandlung von einer post-treatment Follow-up-Phase gesprochen wird. Die Follow-up-Phase unterscheidet sich aufgrund der fortgesetzten Exposition nicht von der Behandlungsphase und beinhaltete nach der letzten Studienvisite an Tag 394 wöchentliche Telefonabfragen. Angesichts der vermehrten Anzahl Todesfälle in der Kontrollgruppe war die Studiendauer und somit auch die Expositionsdauer in der Nusinersen-Gruppe im Median länger als in der Kontrollgruppe (Intervention 280,0 Tage vs. Kontrolle 187,0 Tage). Zum Zeitpunkt des Datenschnitts am 16. Dezember 2016 hatten 73 Probanden (91 %) mindestens 4 Dosen und 32 Probanden (40 %) alle 6 Dosen Nusinersen erhalten. In der Kontrollgruppe hatten sich 34 Probanden (83 %) 4 und

14 Probanden (34 %) allen 6 Scheinverfahren unterzogen. Die kumulative Dosis der Studienbehandlung betrug im Median 53,5 mg.

Aufgrund eines erzielten Wirksamkeitsnachweises ist die Studie am 21. November 2016 vorzeitig beendet worden. Der Datenschnitt erfolgte am 16. Dezember 2016. In der ENDEAR-Studie schlossen unter Nusinersen 49 % der Probanden und im Kontrollarm 32 % der Probanden vorzeitig die Studie ab. Demgegenüber beendeten 33 vs. 27 % (Nusinersen vs. BSC) die Studie inklusive Follow-up. Dabei kam es in der ENDEAR-Studie bei 59 % der Studienteilnehmer unter Nusinersen und bei 63 % der Studienteilnehmer unter BSC zu einem Behandlungsabbruch, 16 vs. 39 % (Nusinersen vs. BSC) brachen die Behandlung aufgrund von unerwünschten Ereignissen ab. (vgl. Tabelle 2 und Tabelle 15).

Bei der Zulassungsstudie CHERISH handelt es sich ebenfalls um eine randomisierte, doppelblinde, scheininterventionskontrollierte, multizentrische Phase-III-Studie. Untersucht wurde die Wirksamkeit, Sicherheit, Verträglichkeit und Pharmakokinetik von Nusinersen im Vergleich zu einer Scheinintervention bei Patienten mit spätem Krankheitsbeginn der SMA ((Klein-) Kinder 2–12 Jahre alt). Eine Stratifizierung wurde anhand des Alters bei Screening (≤ 6 vs. > 6 Jahre) vorgenommen. Eine Interimsanalyse für den primären Endpunkt war für den 31. August 2016 vorgesehen. Es sind 126 Probanden im Verhältnis 2:1 randomisiert und auf die Studienarme Nusinersen (N=84) und BSC (N=42) verteilt worden. Nusinersen wurde in der CHERISH-Studie in einer Dosierung von 12 mg (5 ml) als intrathekale Bolusinjektion an den Studientagen 1, 29, 85 (Aufsättigung) und 274 (Erhaltung), über eine 9-monatige Behandlungsphase und eine 6-monatige Follow-up-Phase (bis Tag 456 ± 7 Tage) verabreicht. Aufgrund der langen Halbwertszeit von Nusinersen wurde die Behandlung in der Follow-up-Phase fortgesetzt, welche sich daher nicht von der Behandlungsphase unterscheidet. Während des Follow-ups gab es zwei Studienvisiten (Tag 365 und 456) sowie monatliche Telefonabfragen. Fraglich bleibt, warum im Studienbericht trotz weiterführender Behandlung von einer post-treatment Follow-up-Phase gesprochen wird. Die verabreichte Dosierung entspricht nicht den Angaben der Fachinformation von Nusinersen. Die unterschiedlichen Dosierungsschemata der beiden Zulassungsstudien werden seitens pU jedoch mit einer Vereinbarung mit der EMA begründet. Dabei sei die Verabreichung einer fixen Dosis von 12 mg Nusinersen mit 4 Aufsättigungsdosen an Tag 0, 14, 28 und 63 sowie anschließenden Erhaltungsdosen alle 4 Monate als geeignet erachtet worden. Die Zweckmäßigkeit des Dosierungsschemas gemäß Fachinformation würde in laufenden Studien und üblichen Monitoringprozessen weiter evaluiert [1,12]. Im beigefügten EPAR finden sich keine ergänzenden Informationen zur Festlegung der Dosierung [18]. Wie die ENDEAR-Studie wurde die CHERISH-Studie aufgrund eines erzielten Wirksamkeitsnachweises vorzeitig beendet. Unter Nusinersen schlossen 21 % der Probanden und im Kontrollarm 19 % der Probanden vorzeitig die Studie ab, inklusive Follow-up waren es 79 vs. 81 % (Nusinersen vs. BSC). Die mediane Studiendauer war in beiden Behandlungsgruppen ähnlich (Intervention 450,5 Tage vs. Kontrolle 450,0 Tage). Mit dem Datenschnitt am 3. März 2017 hatten alle 84 Probanden mindestens 3 Dosen und 83 Probanden (99 %) 4 Dosen, Nusinersen erhalten. In der Kontrollgruppe haben sich alle 42 Probanden 4 Scheinverfahren unterzogen. Die kumulative Dosis der Studienbehandlung betrug im Median 48 mg (vgl. Tabelle 4 und Tabelle 15).

Die Studienpopulation der ENDEAR-Studie bestand aus Patienten mit infantiler SMA Typ I im Alter von < 7 Monaten. Alle Studienteilnehmer, abgesehen von einem Probanden der Kontrollgruppe, wiesen 2 SMN2-Kopien auf. Unterschiede in der Charakteristik der Studienpopulation zeigten sich insbesondere für die Merkmale Alter bei Screening (Nusinersen 147,2 Tage vs. BSC 164,7 Tage), Alter bei 1. Dosis der Studienmedikation (Nusinersen zwischen 3 und 6 Monate alt (54 %),

Kontrollgruppe älter als 6 Monate (61 %)), Alter bei Symptombeginn (Nusinersen-MW: 7,9 Wochen vs. BSC-MW: 9,6 Wochen) und Alter bei SMA-Diagnose (Nusinersen-MW: 12,6 Wochen vs. BSC-MW: 17,5 Wochen). Zudem wiesen zu Studienbeginn mehr Probanden im Interventionsarm eine paradoxe Atmung, Pneumonie / respiratorische Symptome und Schluckbeschwerden / Probleme bei der Nahrungsaufnahme auf als im Kontrollarm. In die CHERISH-Studie waren Patienten mit späterem SMA-Krankheitsbeginn im Alter von 2–12 Jahren eingeschlossen. Die Mehrzahl der Probanden waren jünger als 6 Jahre (Nusinersen 83 % vs. BSC 86 %), davon die meisten 2 Jahre alt. 88 % der Studienpopulation hatten 3 SMN2-Kopien. Zu Studienbeginn konnten in der Kontrollgruppe mehr Patienten frei stehen als im Interventionsarm (29 vs. 13 %), ansonsten waren die Behandlungsarme hinsichtlich der Baseline-Charakteristika vergleichbar (vgl. Tabelle 16).

Sowohl die ENDEAR- als auch die CHERISH-Studie waren insgesamt über alle Merkmale adäquat verblindet. Zur Sicherstellung der Verblindung erfolgte die Verabreichung von Nusinersen in einem separaten Raum durch ausgebildetes Studienpersonal, das nicht verblindet war. Weder die Eltern noch Prüfärzte, Studienkoordinatoren oder Endpunkterheber waren während der Behandlung anwesend. Für die Analyse der Zwischenergebnisse war eine Entblindung hinsichtlich der Behandlungszuteilung erforderlich. Das dafür von der Verblindung aufgehobene Studienpersonal war infolge nicht mehr an der Durchführung der Studien beteiligt. Durch den natürlichen Krankheitsverlauf bei SMA oder mögliche Wirksamkeitseffekte der Studienbehandlung, erscheint – trotz adäquater Verblindungsmaßnahmen in der Studienorganisation – eine Aufrechterhaltung der Verblindung des untersuchenden Arztes, des beteiligten Studienpersonals und/oder der Eltern/Erziehungsberechtigten unter Umständen als nicht gegeben und infolgedessen eine Verzerrung der Endpunkterhebung als möglich. Aufgrund des Studiendesigns und der dargelegten Studiencharakteristika ist für die Studien ENDEAR und CHERISH jeweils von einem niedrigen Verzerrungspotential auszugehen.

Die Fallzahlberechnung der ENDEAR- und der CHERISH-Studie erfolgte jeweils für den primären Endpunkt. In beiden Studien war eine Interimsanalyse für den primären Endpunkt vorgesehen, am 15. Juni 2016 (ENDEAR) bzw. 31. August 2016 (CHERISH). Die abschließende Datenanalyse erfolgte nach 13 bzw. 15 Monaten. Mit Einreichung des Nutzendossiers zu Nusinersen legt der pU über den Studienbericht hinaus zusätzliche, für das Dossier angefertigte, Datenanalysen für den finalen Datenschnitt vor. Teilweise bleibt unklar, welches statistische Modell für die ergänzenden Berechnungen verwendet wurde. Dies ist von besonderem Interesse, da die Effektschätzer teilweise leicht von denen im Studienbericht abweichen. Die Ergebnisse beider Quellen sind gleichgerichtet und es ist wahrscheinlich, dass der Unterschied auf verschiedene statistische Modelle zurückzuführen ist. Dies wird jedoch nicht ausgewiesen.

3.3 Wirksamkeit

Für den Nachweis der Wirksamkeit von Nusinersen bei Patienten mit SMA reicht der pU Daten zur Mortalität sowie zur Beatmung (Zeit bis zum Tod oder zur dauerhaften Beatmung, Zeit bis zur dauerhaften Beatmung), zur motorischen Entwicklung (HINE, CHOP INTEND und HFMSE) und Krankheitslast (Häufigkeit schwerwiegender respiratorischer Ereignisse, Hospitalisierungen) ein. Alle Endpunkte werden als patientenrelevant eingestuft und für die vorliegende Nutzenbewertung berücksichtigt.

Mortalität

In der ENDEAR-Studie verstarben insgesamt 16 % der Patienten unter Nusinersen und 39 % unter der Scheinintervention. Es zeigte sich ein statistisch signifikanter Unterschied zugunsten von Nusinersen (HR: 0,367; 95%-KI: [0,18;0,76]; $p=0,0074$). Das Risiko zu versterben verringerte sich für Patienten unter der Behandlung mit Nusinersen signifikant um 63,3 %. Das mediane Überleben wurde in beiden Studienarmen nicht erreicht. Aufgrund des insgesamt geringen Anteils erreichter Ereignisse von 24 % (29/121) und des Nichterreichens des medianen Überlebens ist die Aussagesicherheit der Ergebnisse eingeschränkt.

In der CHERISH-Studie ereigneten sich weder unter Nusinersen noch unter BSC Todesfälle.

Zeit bis zum Tod oder zur dauerhaften Beatmung

Für den kombinierten und zweiten primären Endpunkt der ENDEAR-Studie zeigte sich ein signifikanter Unterschied zwischen den Studienarmen zugunsten einer Behandlung mit Nusinersen (HR: 0,47, 95%-KI: [0,28;0,78]; $p=0,0037$). Im Kontrollarm erlitten 68 % der Patienten ein Ereignis bei einer medianen Zeit bis zum Tod oder zur dauerhaften Beatmung von 22,6 Wochen (95%-KI: [13,6;31,3]). Unter Nusinersen trat bei 39 % der Patienten ein Ereignis auf, wobei die mediane Zeit bis zum Ereignis nicht ermittelt werden konnte.

Zeit bis zur dauerhaften Beatmung

Die Teilkomponente Zeit bis zur dauerhaften Beatmung wurde in der ENDEAR-Studie ebenfalls ergänzend zum primären kombinierten Endpunkt Zeit bis zum Tod oder zur dauerhaften Beatmung analysiert. In der Auswertung des Endpunktes zeigte sich, dass unter Nusinersen weniger Patienten einer dauerhaften Beatmung (Beatmung von ≥ 16 Stunden/Tag durchgehend für mehr als 21 Tage bei Abwesenheit eines akuten, reversiblen Ereignisses oder Tracheotomie) bedurften als unter der Scheinintervention (23 vs. 32 %), wobei kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen vorliegt. Zu Studienbeginn wies ein größerer Anteil Studienteilnehmer in der Interventionsgruppe eine Pneumonie bzw. respiratorische Symptome auf, jedoch musste niemand dauerhaft beatmet werden.

Hammersmith Infant Neurological Examination (HINE) (Subskala 2)

Der HINE ist ein Instrument zur Untersuchung motorischer Fähigkeiten im Säuglingsalter. Das Erreichen motorischer Meilensteine wurde in der ENDEAR-Studie über den primären Endpunkt HINE-Responder (Subskala 2) sowie den Endpunkt Veränderung des HINE-Scores im Vergleich zu Baseline bewertet. Zum Zeitpunkt der Analyse erreichten 51 % der Patienten unter einer Behandlung mit Nusinersen einen motorischen Meilenstein im Vergleich zu keinem Patienten unter einer Behandlung mit BSC. Die Responserate war in der Interventionsgruppe signifikant erhöht (RR: 43,04; 95%-KI: [2,72;682,32]; $p=0,0076$). Die Ergebnisse sind konsistent mit den Ergebnissen der Interimsanalyse. Die Analyse für die ES-Population sowie die durchgeführten Sensitivitätsanalysen stützen die Robustheit der Ergebnisse des primären Endpunktes. Es bleibt unklar, ob eine dauerhafte Response erforderlich war oder ein einmaliges Erreichen des Responsekriteriums ausreichte, um als Responder gezählt zu werden.

Auch für den Endpunkt Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline zeigte sich ein statistisch signifikanter Unterschied zwischen den Vergleichsarmen zugunsten von Nusinersen (Differenz MW: 3,60; 95%-KI: [2,29;4,90]; $p<0,0001$). Die Effektstärke mit einem

Hedges' g von 1,05 (95%-KI: [0,64;1,44]; $p < 0,0001$) wird als relevant interpretiert. Die Auswertung der ES-Population zeigte vergleichbare Ergebnisse.

Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease (CHOP INTEND)

In der ENDEAR-Studie werden die Ergebnisse zum CHOP INTEND in Form einer Responderanalyse und der Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline unterstützend zum primären Endpunkt HINE berichtet. Insgesamt erreichten statistisch signifikant deutlich mehr Probanden unter Nusinersen (73 %) die Responseschwelle als im Vergleichsarm (2 %) (RR: 29,73; 95%-KI: [4,27;206,98]; $p = 0,0006$). In der Auswertung des CHOP INTEND im Vergleich zu Baseline fiel auch die Veränderung des Gesamtwertes im Nusinersen-Arm deutlich höher aus als im Vergleichsarm. Der Unterschied zwischen den Vergleichsarmen ist statistisch signifikant (Differenz MW: 18,66; 95%-KI: [15,42;21,90]; $p < 0,0001$) bei nachweislich großer relevanter Effektstärke (Hedges' g: 2,19 (95%-KI: [1,72;2,66]; $p < 0,0001$) zugunsten von Nusinersen. Die Auswertungen der ES-Population und die Ergebnisse der durchgeführten Sensitivitätsanalysen bestätigen die Robustheit der Ergebnisse der Hauptanalyse.

Hammersmith Functional Motor Scale Expanded (HFMSSE)

Der HFMSSE ist ebenfalls ein Instrument zur Untersuchung der motorischen Funktion bei Kindern mit späterem Krankheitsbeginn der SMA. Für den primären Endpunkt der CHERISH-Studie Änderung der motorischen Funktion nach 15 Monaten im Vergleich zu Baseline zeigte sich eine signifikante Veränderung des HFMSSE-Scores zugunsten einer Behandlung mit Nusinersen (Differenz LS MW: 4,42; 95%-KI: [2,46;6,38]; $p < 0,0001$). Das 95%-KI des Hedges' g liegt dabei vollständig oberhalb der Irrelevanzschwelle von 0,2, womit sich ein relevanter Effekt ableiten lässt. Eine ergänzend dargestellte Auswertung in der ES-Population sowie die durchgeführten Sensitivitätsanalysen zum primären Endpunkt unterstützen die Ergebnisse der Hauptanalyse in der ITT-Population.

In der Responderanalyse zeigt sich, dass das Erreichen der Responseschwelle in der Nusinersen-Gruppe mehr als zweimal höher war als in der Scheinkontrollgruppe. Das Ergebnis ist statistisch signifikant (RR: 2,17; 95%-KI: [1,23;3,81]); $p = 0,0073$). Die Sensitivitätsanalysen des Endpunktes (inklusive einer Auswertung der ES-Population) zeigen gleichgerichtete Effekte und bestätigen die Ergebnisse der Hauptanalyse.

Häufigkeit schwerwiegender respiratorischer Ereignisse

In der ENDEAR-Studie zeigte sich für den Endpunkt kein Unterschied zwischen den Behandlungsarmen. In der CHERISH-Studie ist der Endpunkt post hoc ausgewertet worden. Wie in der ENDEAR-Studie ergab sich kein Unterschied zwischen den Behandlungsarmen. Bezugnehmend auf die Operationalisierung des Endpunktes bleibt unklar, ob es zu Doppelzählungen mit der Häufigkeit von unerwünschten Ereignissen (SOC) gekommen ist und wie mit der Problematik umgegangen wurde.

Häufigkeit von Hospitalisierungen

Für den Endpunkt Häufigkeit von Hospitalisierungen zeigte sich in der ENDEAR-Studie kein Unterschied zwischen den Behandlungsarmen. In der CHERISH-Studie ergab sich ein Unterschied zwischen den Behandlungsarmen zugunsten von Nusinersen (Rate Ratio: 0,35; 95%-KI:

[0,14;0,87]; $p=0,02$). Unter der Scheinintervention traten 16 Ereignisse, unter Nusinersen 11 Ereignisse auf. Wie für den vorherigen Endpunkt bleibt es hinsichtlich der Operationalisierung des Endpunktes unklar, ob es zu Doppelzählungen mit der Häufigkeit von unerwünschten Ereignissen gekommen ist, und wie mit der Problematik umgegangen wurde.

Zusammenfassende Einschätzung zur Wirksamkeit

In der Zusammenschau zeigen sich für die ENDEAR-Studie bei Patienten mit infantiler SMA über alle ausgewerteten Endpunkte, abgesehen von der Zeit bis zur dauerhaften Beatmung, Häufigkeit von Hospitalisierungen und schwerwiegender respiratorischer Ereignisse signifikante Ergebnisse zugunsten der Wirksamkeit von Nusinersen. Das Verzerrungspotential wurde für alle Endpunkte, abgesehen von den Endpunkten HINE und CHOP INTEND, mit niedrig bewertet. Weder der HINE noch der CHOP INTEND werden als ausreichend validiert eingestuft und für beide Instrumente fehlen Untersuchungen zur Responseschwelle (MCID) (vgl. 2.5.3). Problematisch ist auch, dass unklar bleibt, ob für den HINE eine dauerhafte Response erforderlich oder ein einmaliges Erreichen des Responsekriteriums ausreichend war. Auch wenn bei dem vorliegenden Krankheitsbild von einer stetigen Verschlechterung des Gesundheitszustandes auszugehen ist und jegliche dauerhafte Verbesserung der motorischen Funktion bei SMA-I-Patienten als erstrebenswert angesehen werden kann, ist der Einsatz von validierten Instrumenten (HINE, CHOP INTEND) in Studien nötig. Aufgrund dieser Limitationen sind die Ergebnisse, hinsichtlich der Validität der Erhebungsinstrumente bzw. der Präzision und Größe des Effekts, mit Unsicherheiten behaftet. Für eine Robustheit der Ergebnisse und den Nachweis einer dauerhaften Wirksamkeit von Nusinersen können zusätzliche Daten noch laufender Studien unterstützend hinzugezogen werden. Es zeigte sich für keine der untersuchten Effektmodifikatoren eine Subgruppeninteraktion konstant über alle dargestellten Endpunkte. Darüber hinaus sind die Stichprobengrößen der Subgruppen klein und zeigen ein Ungleichgewicht, weshalb die Ergebnisse mit Vorsicht zu interpretieren sind.

In der CHERISH-Studie zeigte sich bei Patienten mit spätem Krankheitsbeginn der SMA ebenfalls über alle Wirksamkeitsendpunkte, abgesehen vom Endpunkt Häufigkeit schwerwiegender respiratorischer Ereignisse, ein statistisch signifikanter Behandlungsvorteil zugunsten von Nusinersen. Es wurden keine Todesfälle berichtet. Der HFMSE gilt als validiertes Instrument, jedoch fehlen auch hier bestätigende Untersuchungen zur MCID. Die Ergebnisse der Responderanalyse sind daher mit Vorsicht zu interpretieren. Das Verzerrungspotential ist für alle Endpunkte der CHERISH-Studie als niedrig eingestuft. Um die langfristige Wirksamkeit von Nusinersen zu bewerten sind, für die Robustheit der Ergebnisse und die Übertragbarkeit der Effekte auf die Population, unterstützende Langzeitdaten wünschenswert.

3.4 Lebensqualität

Die gesundheitsbezogene Lebensqualität ist in der CHERISH-Studie anhand der Generic-Core-Skala und dem Neuromuskulären Modul 3.0 des Pediatric Quality of Life Inventory (PedsQL), Version 4.0 untersucht worden.

Aufgrund der geringen Rücklaufquoten (< 70 %) in beiden Studienarmen, maximal 25 % im Nusinersen-Arm und maximal 19 % im Kontrollarm, ist seitens des pU auf eine ausführliche Ergebnisdarstellung in Modul 4 des Herstellerdossiers verzichtet worden. Der Einschätzung des pU ist nach abschließender Sichtung der Studienergebnisse gefolgt worden.

Zusammenfassende Einschätzung zur Lebensqualität

Aufgrund geringer Rücklaufquoten ist eine Einschätzung zur Lebensqualität für den Vergleich von Nusinersen vs. BSC nicht möglich.

3.5 Sicherheit

Die Sicherheitspopulationen der pivotalen Studien ENDEAR und CHERISH waren identisch mit den entsprechenden ITT-Populationen der Studien. In der ENDEAR-Studie gingen 121 Patienten, in der CHERISH-Studie 126 Patienten in die Sicherheitsbewertung ein. Das Verzerrungspotential der Sicherheitsendpunkte wird für beide Studien als niedrig eingestuft.

Angesichts der vermehrten Anzahl Todesfälle in der Kontrollgruppe, waren die Studiendauer und somit auch die Expositionsdauer in der ENDEAR-Studie unterschiedlich lang. Die Probanden in der Nusinersen-Gruppe waren im Median länger in der Studie als Probanden in der Kontrollgruppe.

Die vorliegenden Daten zur Sicherheit sind nicht ausschließlich als Nebenwirkungen anzusehen, da diese auch Ereignisse der Grunderkrankung umfassen.

Unerwünschte Ereignisse

Insgesamt trat bei 96 % der Probanden im Nusinersen-Arm und 98 % im BSC-Arm in der ENDEAR-Studie ein UE auf. Ein signifikanter Unterschied zugunsten von Nusinersen zeigte sich für die Endpunkte schwere UE, SUE und Therapieabbruch aufgrund von UE. Das Relative Risiko war in der Interventionsgruppe entsprechend um 30 %, 20 % und 58 % geringer als in der Kontrollgruppe, wobei eine adäquatere Ergebnisdarstellung Ereigniszeitanalysen wären, da der Effekt aufgrund der beschriebenen unterschiedlichen Expositionsdauer und der Möglichkeit der Doppelzählungen schwierig zu interpretieren ist.

In der CHERISH-Studie zeigt sich für die Endpunkte UE, milde, moderate und schwere UE, SUE und Therapieabbrüche aufgrund von UE jeweils kein statistisch signifikanter Unterschied zwischen den Vergleichsarmen.

Häufige unerwünschte Ereignisse nach Systemorganklassen und Preferred Term

Die am häufigsten aufgetretenen UE nach Systemorganklassen und Preferred Term waren in der ENDEAR-Studie Atemwegserkrankungen und/oder Infektionen. Dazu gehören Infektion der oberen Atemwege, Nasopharyngitis, Pneumonie, Bronchitis, Stauung der oberen Atemwege, Bronchitis viral, Grippe, Bronchiolitis und virale Infektionen, virale Infektion der oberen Atemwege, Infektion der Atemwege, Rhinovirusinfektionen, Atemnot, Dyspnoe, Atemstillstand, akutes Atemversagen, Atelektase und Husten. Patienten in der Nusinersen-Gruppen wiesen zu Studienbeginn jedoch einen schlechteren klinischen Zustand in Form von paradoxer Atmung, Pneumonie / respiratorischen Symptomen und Schluckbeschwerden / Probleme bei der Nahrungsaufnahme als Patienten der Kontrollgruppe auf.

Andere häufig gemeldete Ereignisse waren Verstopfung, gastroösophageale Refluxkrankheit, Erbrechen, Zahnen, Diarrhö, Verstopfung, Dysphagie, Pyrexie, Abnahme der Sauerstoffsättigung und Apirationspneumonie.

Auch in der CHERISH-Studie waren die häufigsten UE nach Systemorganklassen und Preferred Term respiratorische Erkrankungen oder Infektionen. Andere häufig aufgetretene Ereignisse waren

Gastroenteritis, virale Gastroenteritis, Influenza, Ohreninfektion, Pyrexie, Erbrechen, Diarrhö, Verstopfung, Kopfschmerzen und Rückenschmerzen.

Kopfschmerzen, Rückenschmerzen und Erbrechen waren im Rahmen der Applikation der Studienmedikation in Form einer Lumbalpunktion zu erwarten.

Zusammenfassende Einschätzung zur Sicherheit

In der ENDEAR-Studie zeigte sich für die Endpunkte schwere UE, SUE und Therapieabbruch aufgrund von UE ein statistisch signifikanter Unterschied zugunsten der Nusinersen-Behandlung. Dieser Vorteil für Nusinersen ergibt sich trotz einer längeren Beobachtungszeit als im Kontrollarm (Intervention Median 280,0 Tage vs. Kontrolle Median 187,0 Tage) und somit längerer Exposition für die Studienteilnehmer im Interventionsarm. Es bleibt jedoch unklar, ob die UE mit der Symptomatik der Grunderkrankung einhergehen oder als Nebenwirkungen der Behandlung auftraten. Die Häufigkeit schwerwiegender Atemwegereignisse und die Häufigkeit und Dauer der Hospitalisierungen sind in der ENDEAR-Studie als Wirksamkeitsendpunkt verortet. Es wird jedoch nicht adressiert, inwieweit Doppelzählungen vermieden wurden. Die Bewertung eines Gruppenunterschieds für die Sicherheit hinsichtlich der Zuordnung von Symptomen und Nebenwirkungen ist insofern nicht möglich. In der CHERISH-Studie ergab sich für die Sicherheitsendpunkte kein Unterschied zwischen den Behandlungsarmen. Passend zum klinischen Krankheitsbild der SMA zählen zu den häufigsten UE in beiden Studien respiratorische Erkrankungen und Infektionen.

4 Anforderung an eine qualitätsgesicherte Anwendung

Die Vorgaben der Fachinformation sind zu berücksichtigen.

Die Behandlung mit Spinraza® sollte nur von einem Arzt eingeleitet werden, der Erfahrung in der Behandlung der spinalen Muskelatrophie (SMA) hat.

Da Spinraza® zur intrathekalen Anwendung bestimmt ist, sollte die Behandlung nur durch Ärzte erfolgen, die Erfahrung in der Durchführung von Lumbalpunktionen haben.

5 Zusammenfassung der Bewertung

Nusinersen (Spinraza®) ist zugelassen für die Behandlung der 5q-assoziierten spinalen Muskelatrophie [12]. Die Nutzenbewertung von Nusinersen basiert auf den pivotalen Studien ENDEAR (CS3B) und CHERISH (CS4). Bei beiden Studien handelt es sich um randomisierte, doppelblinde, kontrollierte, multizentrische Phase-III-Studien zur Beurteilung der Wirksamkeit und Sicherheit von Nusinersen im Vergleich zu BSC [10,11].

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

Tabelle 36: Zusammenfassende Darstellung der Ergebnisse der Studien ENDEAR und CHERISH

Endpunkt	Nusinersen vs. BSC ¹⁾ Effektschätzer [95%-KI]; p-Wert	Effekt
Mortalität		
<i>ENDEAR</i>		
Zeit bis zum Tod	HR: 0,367 [0,18;0,76]; 0,0074	↑↑
Morbidität		
<i>ENDEAR</i>		
Zeit bis zum Tod oder zur dauerhaften Beatmung	HR: 0,47 [0,28;0,78]; 0,0037	↑↑
Zeit bis zur dauerhaften Beatmung	HR: 0,56 [0,27;1,15]; 0,1142	↔
HINE-Responder (Subskala 2)	RR: 43,04 [2,72;682,32]; 0,0076	↑
Veränderung des HINE-Scores (Subskala 2) im Vergleich zu Baseline	Differenz MW: 3,60 [2,29;4,90]; < 0,0001 Hedges' g: 1,05 [0,64;1,44]; < 0,0001	↑
CHOP-INTEND-Responder	RR: 29,73 [4,27;206,98]; 0,0006	↑
Veränderung des CHOP-INTEND-Gesamtwertes im Vergleich zu Baseline	Differenz MW: 18,66 [15,42;21,90]; < 0,0001 Hedges' g: 2,19 [1,72;2,66]; < 0,0001	↑
Häufigkeit schwerwiegender respiratorischer Ereignisse	Rate Ratio: 0,81 [0,53;1,25]; 0,35	↔
Häufigkeit von Hospitalisierungen	Rate Ratio: 0,91 [0,66;1,26]; 0,57	↔
<i>CHERISH</i>		
Veränderung des HFMSE-Scores	Differenz MW: 4,42 [2,46;6,38]; < 0,0001 Hedges' g: 0,99 [0,55;1,42]; < 0,0001	↑↑
HFMSE-Responder	RR: 2,17 [1,23;3,81]; 0,0073	↑↑
Häufigkeit schwerwiegender respiratorischer Ereignisse	Rate Ratio: 0,40 [0,14;1,14]; 0,09	↔
Häufigkeit von Hospitalisierungen	Rate Ratio: 0,35 [0,14;0,87]; 0,02	↑↑

Endpunkt	Nusinersen vs. BSC ¹⁾ Effektschätzer [95%-KI]; p-Wert	Effekt
Lebensqualität		
<i>CHERISH</i>		
PedsQL ²⁾		-
Sicherheit³⁾		
<i>ENDEAR</i>		
Unerwünschte Ereignisse (UE)		-
- mild	RR ⁴⁾ : 3,59 [0,46;28,18]; 0,22	↔
- moderat oder schwer	RR ⁴⁾ : 0,92 [0,83;1,03]; 0,13	↔
- schwer	RR ⁴⁾ : 0,70 [0,55;0,89]; 0,004	↑↑
Schwerwiegende UE (SUE)	RR ⁴⁾ : 0,80 [0,70;0,92]; 0,002	↑↑
Therapieabbruch aufgrund UE	RR ⁴⁾ : 0,42 [0,22;0,78]; 0,006	↑↑
<i>CHERISH</i>		
Unerwünschte Ereignisse (UE)		-
- mild	RR: 1,03 [0,69;1,54]; 0,90	↔
- moderat oder schwer	RR: 0,85 [0,59;1,21]; 0,37	↔
- schwer	RR: 0,67 [0,16;2,84]; 0,58	↔
Schwerwiegende UE (SUE)	RR: 0,58 [0,30;1,15]; 0,12	↔
Therapieabbruch aufgrund UE		-

¹⁾ Die dargestellten Ergebnisse beziehen sich jeweils auf die ITT-Population.

²⁾ Aufgrund der geringen Rücklaufquoten entfällt eine Ergebnisdarstellung für den Endpunkt Lebensqualität mittels PedsQL.

³⁾ Die Ergebnisse zur Sicherheit von Nusinersen sind nicht ausschließlich als Nebenwirkungen anzusehen, da diese auch Ereignisse der Grunderkrankung umfassen.

⁴⁾ Eine adäquatere Ergebnisdarstellung wären Ereigniszeitanalysen, da der Effekt aufgrund der beschriebenen unterschiedlichen Expositionsdauer der Behandlungsarme und der Möglichkeit der Doppelzählungen schwierig zu interpretieren ist.

Abkürzungen: BSC: Best Supportive Care; CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disease; HFMSE: Hammersmith Functional Motor Scale Expanded; HINE: Hammersmith Infant Neurological Examination; HR: Hazard Ratio; MW: Mittelwert; PedsQL: Pediatric Quality of Life Inventory; RR: Relatives Risiko, (S)UE: (schwerwiegende/s) unerwünschte/s Ereignis/se.

Referenzen

1. **Biogen.** Dossier zur Nutzenbewertung gemäß § 35a SGB Nusinersen (Spinraza®). 2017.
2. **Biogen.** Expanded Access Program (EAP) to Provide Nusinersen to Patients With Infantile-onset Spinal Muscular Atrophy (SMA) [online]. NCT02865109. In: ClinicalTrials.gov. 24.08.2017. 2016. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT02865109>.
3. **Biogen.** An Open-Label Extension Study for Patients With Spinal Muscular Atrophy Who Previously Participated in Investigational Studies of ISIS 396443(CS11) [online]. NCT02594124. In: ClinicalTrials.gov. 16.08.2017. 2015. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT02594124>.
4. **Biogen.** An Open-Label Study to Assess the Efficacy, Safety, Tolerability, and Pharmacokinetics of Multiple Doses of ISIS 396443 Delivered Intrathecally to Subjects With Genetically Diagnosed and Presymptomatic Spinal Muscular Atrophy (NUTURE) [online]. NCT02386553. In: ClinicalTrials.gov. 15.08.2017. 2015. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT02386553>.
5. **Biogen.** An Open-label Study to Assess the Safety and Tolerability of a Single Intrathecal Dose of ISIS 396443 in Patients With Spinal Muscular Atrophy Who Previously Participated in ISIS 396443-CS1(CS10) [online]. In: ClinicalTrials.gov. 30.03.2017. 2013. [Zugriff: 11.07.2017]. (Band NCT01780246). URL: <https://clinicaltrials.gov/show/NCT01780246>.
6. **Biogen.** An Open-label Study to Assess the Safety and Tolerability of ISIS 396443 in Patients With Spinal Muscular Atrophy Who Previously Participated in 396443-CS2 or 396443-CS10 (CS12) [online]. NCT02052791. In: ClinicalTrials.gov. 22.03.2017. 2014. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT02052791>.
7. **Biogen.** An Open-Label, Dose Escalation Study to Assess the Safety, Tolerability and Dose-Range Finding of Multiple Doses of ISIS 396443 Delivered Intrathecally to Patients With Spinal Muscular Atrophy (CS2) [online]. NCT01703988. In: ClinicalTrials.gov. 20.01.2017. 2012. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT01703988>.
8. **Biogen.** An Open-label, Escalating Dose Study to Assess the Safety, Tolerability and Dose-range Finding of a Single Intrathecal Dose of ISIS 396443 in Patients With Spinal Muscular Atrophy (CS1) [online]. NCT01494701. In: ClinicalTrials.gov. 05.04.2017. 2011. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT01494701>.
9. **Biogen.** A Phase 2, Randomized, Double-blind, Sham-procedure Controlled Study to Assess the Safety and Tolerability and Explore the Efficacy of ISIS 396443 (BIIB058) Administered Intrathecally in Subjects With Spinal Muscular Atrophy Who Are Not Eligible to Participate in the Clinical Studies ISIS 396443-CS3B or ISIS 396443-CS4 (EMBRACE) [online]. NCT02292537. In: ClinicalTrials.gov. 04.05.2017. 2015. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT02462759>.
10. **Biogen.** A Phase 3, Randomized, Double-Blind, Sham-Procedure Controlled Study to Assess the Clinical Efficacy and Safety of ISIS 396443 Administered Intrathecally in Patients With Infantile-onset Spinal Muscular Atrophy (ENDEAR); Clinical Study Report [online]. NCT02193074. In: ClinicalTrials.gov. 29.06.2017. 2014. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT02193074>.
11. **Biogen.** A Phase 3, Randomized, Double-blind, Sham-Procedure Controlled Study to Assess the Clinical Efficacy and Safety of ISIS 396443 Administered Intrathecally in Patients With Later-onset Spinal Muscular Atrophy (CHERISH); Clinical Study Report [online]. NCT02292537. In: ClinicalTrials.gov. 22.03.2017. 2014. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT02292537>.
12. **Biogen.** Spinraza® 12 mg Injektionslösung [online]. 05.2017. Berlin (GER): 2017. [Zugriff: 11.07.2017]. URL: <http://www.fachinfo.de>.

13. **Biogen.** A Study to Assess the Efficacy, Safety, Tolerability, and Pharmacokinetics of Multiple Doses of ISIS 396443 Delivered Intrathecally to Patients With Infantile-Onset Spinal Muscular Atrophy (CS3A) [online]. NCT01839656. In: ClinicalTrials.gov. 28.07.2017. 2013. [Zugriff: 11.07.2017]. URL: <https://clinicaltrials.gov/show/NCT01839656>.
14. **De Sanctis R, Coratti G, Pasternak A, Montes J, Pane M, Mazzone ES, et al.** Developmental milestones in type I spinal muscular atrophy. *Neuromuscul Disord* 2016;26(11):754-759.
15. **DynaMed Plus.** Spinal muscular atrophy (SMA) [online]. Record No. 114849. 05.01.2017. Ipswich (USA): EBSCO Information Services. [Zugriff: 11.07.2017]. URL: <https://www.dynamed.com/topics/dmp~AN~T114849/Spinal-muscular-atrophy-SMA>.
16. **European Medicines Agency (EMA).** CHMP summary of positive opinion for Spinraza; EMA/181654/2017 [online]. 21.04.2017. London (GBR): EMA. [Zugriff: 11.07.2017]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Summary_of_opinion_-_Initial_authorisation/human/004312/WC500226191.pdf.
17. **European Medicines Agency (EMA).** EMA decision P/0251/2016 of 23 September 2016 on the acceptance of a modification of an agreed paediatric investigation plan for nusinersen (EMEA-001448-PIP01-13-M02) [online]. 23.09.2016. London (GBR): EMA; 2016. [Zugriff: 11.07.2017]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/PIP_decision/WC500215458.pdf.
18. **European Medicines Agency (EMA).** Spinraza: European public assessment report EMA/289068/2017 [online]. 21.04.2017. London (GBR): EMA. [Zugriff: 11.07.2017]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/004312/WC500229707.pdf.
19. **Finkel R, Bertini E, Muntoni F, Mercuri E.** 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.
20. **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.
21. **Finkel RS, Sejersen T, Mercuri E, Enmc Sma Workshop Study Group.** 218th ENMC International Workshop: Revisiting the Consensus on Standards of Care in SMA February 19–21, 2016, Naarden, The Netherlands. *Neuromuscular Disorders* 2017.
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. **Haataja L, Mercuri E, Regev R, Cowan F, Rutherford M, Dubowitz V, et al.** Optimality score for the neurologic examination of the infant at 12 and 18 months of age. *J Pediatr* 1999;135(2 Pt 1):153-161.
25. **Iannaccone ST, Hynan LS, Morton A, Buchanan R, Limbers CA, Varni JW, et al.** The PedsQL™ in Pediatric Patients with Spinal Muscular Atrophy: Feasibility, Reliability, and Validity of the Pediatric Quality of Life Inventory™ Generic Core Scales and Neuromuscular Module. *Neuromuscular disorders : NMD* 2009;19(12):805-812.
26. **Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG).** Allgemeine Methoden; Vers. 5.0 [online]. 10.07.2017. Köln (GER): IQWiG; 2017. [Zugriff: 11.07.2017]. URL: https://www.iqwig.de/download/Allgemeine-Methoden_Version-5-0.pdf.

27. **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.
28. **Krosschell KJ, Maczulski JA, Crawford TO, Scott C, Swoboda KJ.** A modified Hammersmith functional motor scale for use in multi-center research on spinal muscular atrophy. *Neuromuscul Disord* 2006;16(7):417-426.
29. **Little R, Rubin D.** *Statistical Analysis with Missing Data.* 2nd edition ed; 2002.
30. **Mercuri E, Finkel R, Montes J, Mazzone ES, Sormani MP, Main M, et al.** Patterns of disease progression in type 2 and 3 SMA: Implications for clinical trials. *Neuromuscul Disord* 2016;26(2):126-131.
31. **O'Hagen JM, Glanzman AM, McDermott MP, Ryan PA, Flickinger J, Quigley J, et al.** An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. *Neuromuscul Disord* 2007;17(9-10):693-697.
32. **Ramsey D.** Outcome Measures: Information for Professionals. SMA Type 2 & 3 - The Revised Hammersmith Scale [online]. 07.05.2017. 2017. [Zugriff: 11.07.2017]. URL: <http://www.smareachuk.com/information-for-professionals/category/hammersmith-functional-motor-scale-expanded>.
33. **Romeo DM, Cioni M, Scoto M, Mazzone L, Palermo F, Romeo MG.** Neuromotor development in infants with cerebral palsy investigated by the Hammersmith Infant Neurological Examination during the first year of age. *Eur J Paediatr Neurol* 2007;12(1):24-31.
34. **Schafer J.** *Analysis of incomplete multivariate data.* 1 ed. London (GBR); Chapman & Hall; CRC Press; 1997.
35. **Schafer JL.** Multiple imputation: a primer. *Stat Methods Med Res* 1999;8(1):3-15.
36. **Schröder E.** Die sensomotorische Entwicklung des Kindes. In: Zalpour C (Ed). *Anatomie Physiologie Für die Physiotherapie.* Elsevier; 2016. S. 559-595.
37. **Straßburg H.** Beurteilung der normalen Entwicklung. In: Straßburg HM, Dacheneder W, Kreß W (Eds). *Praxisleitfaden für die interdisziplinäre Betreuung. Entwicklungsstörungen bei Kindern.* Elsevier; 2008. S. 55-73.
38. **Swoboda KJ, Scott CB, Crawford TO, Simard LR, Reyna SP, Krosschell KJ, et al.** SMA CARNI-VAL trial part I: double-blind, randomized, placebo-controlled trial of L-carnitine and valproic acid in spinal muscular atrophy. *PLoS One* 2010;5(8):e12140.
39. **U.S. Food and Drug Administration (FDA), Center for Drug Evaluation and Research.** Spinraza (Nusinersen) Injection: 207500Orig1s000 / 207501Orig1s000; summary review(s) [online]. 23.12.2016. Silver Spring (USA): FDA. [Zugriff: 11.07.2017]. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/209531Orig1s000SumR.pdf.
40. **Varni JW.** Scaling and scoring of the Pediatric Quality of Life Inventory (TM) PedsQL (TM) [online]. Lyon (FRA): Mapi Research Trust; 2017. [Zugriff: 11.07.2017]. URL: <http://www.pedsq.org/PedsQL-Scoring.pdf>.
41. **Varni JW, Seid M, Rode CA.** The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care* 1999;37(2):126-139.