



IQWiG-Berichte – Nr. 891

Neugeborenenenscreening auf 5q-assozierte spinale Muskelatrophie

Abschlussbericht

Auftrag: S18-02
Version: 1.0
Stand: 27.02.2020

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Neugeborenencreening auf 5q-assoziierte spinale Muskelatrophie

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

13.12.2018

Interne Auftragsnummer:

S18-02

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Im Mediapark 8
50670 Köln

Tel.: +49 221 35685-0

Fax: +49 221 35685-1

E-Mail: berichte@iqwig.de

Internet: www.iqwig.de

ISSN: 1864-2500

Dieser Bericht wurde unter Beteiligung externer Sachverständiger erstellt.

Für die Inhalte des Berichts ist allein das IQWiG verantwortlich.

Externe Sachverständige, die wissenschaftliche Forschungsaufträge für das Institut bearbeiten, haben gemäß § 139b Abs. 3 Satz 2 Sozialgesetzbuch (SGB) Fünftes Buch (V) – Gesetzliche Krankenversicherung „alle Beziehungen zu Interessenverbänden, Auftragsinstituten, insbesondere der pharmazeutischen Industrie und der Medizinprodukteindustrie, einschließlich Art und Höhe von Zuwendungen“ offenzulegen. Das Institut hat von jedem der Sachverständigen ein ausgefülltes Formular „Offenlegung potenzieller Interessenkonflikte“ erhalten. Die Angaben wurden durch das speziell für die Beurteilung der Interessenkonflikte eingerichtete Gremium des Instituts bewertet. Die Selbstangaben der externen Sachverständigen zu potenziellen Interessenkonflikten sind in Kapitel A8 zusammenfassend dargestellt. Es wurden keine Interessenkonflikte festgestellt, die die fachliche Unabhängigkeit im Hinblick auf eine Bearbeitung des vorliegenden Auftrags gefährden.

Externe Sachverständige

- Tim Mathes, Institut für Forschung in der Operativen Medizin, Witten/Herdecke
- Barbara Prediger, Institut für Forschung in der Operativen Medizin, Witten/Herdecke
- Benedikt Schoser, Friedrich-Baur-Institut, Neurologische Klinik, Ludwig-Maximilians-Universität München

Das IQWiG dankt den externen Beteiligten für ihre Mitarbeit am Projekt.

Mitarbeiterinnen und Mitarbeiter des IQWiG

- Lina Rodenhäuser
- Andrea Steinzen
- Inga Overesch
- Anne Rummer
- Lisa Schell
- Wiebke Sieben
- Sibylle Sturtz
- Vera Weingärtner

Schlagwörter: Neugeborenencreening, Spinale Muskelatrophie im Kindesalter, Nutzenbewertung, Systematische Übersicht

Keywords: Neonatal Screening, Spinal Muscular Atrophies of Childhood, Benefit Assessment, Systematic Review

Kernaussage

Fragestellung

Das Ziel der vorliegenden Untersuchung ist die Nutzenbewertung eines Neugeborenen-screenings auf 5q-assozierte spinale Muskelatrophie (SMA). Dabei wird das Neugeborenen-screening auf 5q-assozierte SMA in Kombination mit einer Vorverlegung der Diagnosestellung und Behandlung im Vergleich zu keinem 5q-assozierte-SMA-Screening hinsichtlich patientenrelevanter Endpunkte bewertet.

Fazit

Zum Vergleich eines Neugeborenencreenings auf 5q-assozierte spinale Muskelatrophie (SMA) versus kein Neugeborenencreening lagen keine vergleichenden Interventionsstudien der Screeningkette vor. Daher wurden Interventionsstudien, die einen Vergleich eines früheren mit einem späteren Therapiebeginn ermöglichen, sowie Studien zur diagnostischen Güte herangezogen, die mittels Linked-Evidence-Ansatz zusammengeführt wurden.

Für den **Vergleich eines früh- versus einen spätsymptomatischen Therapiebeginn** konnte 1 kleine randomisierte kontrollierte Studie mit kurzer Beobachtungsdauer herangezogen werden, die eine medikamentöse Therapie im Vergleich zu einer Scheinbehandlung bei Kindern mit infantiler 5q-assoziierter SMA untersuchte. Anhand von Subgruppenanalysen zeigten sich für den kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung sowie für den Endpunkt Erreichen motorischer Meilensteine Effektunterschiede zwischen Kindern mit frühsymptomatischem Therapiebeginn (Krankheitsdauer ≤ 12 Wochen) und Kindern mit späterem Therapiebeginn (Krankheitsdauer > 12 Wochen). Bezüglich beider Endpunkte profitieren Kinder mehr von einem frühsymptomatischen als von einem spätsymptomatischen Therapiebeginn. Zu den Endpunkten schwerwiegende unerwünschte Ereignisse, schwere unerwünschte Ereignisse und Therapieabbrüche wegen unerwünschter Ereignisse zeigten sich jeweils keine Effektunterschiede zwischen den Subgruppen. Für andere Endpunkte lagen keine verwertbaren Daten vor.

Zum **Vergleich eines prä- versus einen frühsymptomatischen Therapiebeginn** wurde 1 kleine, vom Hersteller auf Anfrage zur Verfügung gestellte, retrospektive vergleichende Studie eingeschlossen. Es konnten Daten zu Kindern mit diagnostizierter 5q-assoziierter SMA und 2 *SMN2*-Kopien (d. h. prognostizierter oder bestätigter infantiler SMA) herangezogen werden. Für den Endpunkt Erreichen motorischer Meilensteine zeigten sich große Effekte zugunsten des präsymptomatischen im Vergleich zum frühsymptomatischen Therapiebeginn (Krankheitsdauer ≤ 12 Wochen), die sich nicht allein durch Verzerrung erklären ließen (dramatischer Effekt). Bei den Endpunkten schwerwiegende unerwünschte Ereignisse und schwere unerwünschte Ereignisse zeigten sich statistisch signifikante Unterschiede zugunsten des präsymptomatischen Therapiebeginns. Diese beobachteten Unterschiede wurden als nicht groß genug eingeschätzt, als dass sie jeweils nicht allein durch die Einwirkung von Störgrößen zu erklären gewesen sein könnten. Bei den Endpunkten Gesamtüberleben, dauerhafte Beatmung sowie Therapieabbrüche wegen unerwünschter Ereignisse und Rückenschmerzen

zeigten sich jeweils keine statistisch signifikanten Unterschiede. Die Kriterien eines dramatischen Effekts waren nicht erfüllt. Für andere Endpunkte wurden keine Daten angefragt.

Zu **Patientinnen und Patienten mit einem späteren Krankheitsbeginn als bei der infantilen Form** konnten keine vergleichenden Interventionsstudien herangezogen werden.

Für die Bewertung der **diagnostischen Güte** konnten 4 Studien herangezogen werden, in denen jedoch allein die positiven Testergebnisse überprüft wurden (Verification-of-only-positive-Testers-Design). Die Ergebnisse dieser Studien deuten darauf hin, dass die untersuchten Testverfahren für ein Screening bei Neugeborenen auf 5q-assoziierte SMA geeignet sind. Unklar bleibt, wie viele erkrankte Kinder nicht gefunden werden.

Zusammenfassend ergibt sich ein Hinweis für einen Nutzen des Neugeborenencreenings auf 5q-assoziierte SMA im Vergleich zu keinem Screening. Dieses Ergebnis beruht zum einen auf Daten zur prä-, früh- und spätsymptomatischen Arzneimittelbehandlung von Kindern mit diagnostizierter 5q-assoziiierter SMA und prognostizierter oder bestätigter infantiler Form. Den herangezogenen Daten ist eine Zeitpunkt-Wirkungs-Beziehung zu entnehmen, der zufolge bessere Therapieergebnisse erzielt werden, je früher die Therapie einsetzt. Zum anderen basiert die Ableitung eines Hinweises für einen Nutzen des Neugeborenencreenings auf der Eignung diagnostischer Testverfahren und der Möglichkeit, mittels Neugeborenencreening eine Vorverlegung der Diagnosestellung (und damit der Therapie) zu erlangen. Die verfügbaren Daten erlauben keine Schlüsse, ob durch das Screening identifizierte Kinder mit spätem Krankheitsbeginn der SMA (d. h. Symptombeginn erst nach Jahren) bereits von einem präsymptomatischen Therapiebeginn profitieren würden. Derzeit unklar ist somit vor allem der Umgang mit im Screening identifizierten Neugeborenen mit erwartbar spätem Krankheitsbeginn (≥ 4 SMN2-Kopien). Ein adäquater Umgang mit diesen Kindern und ihren Familien – einschließlich der Erwägung einer selbstbestimmten Entscheidung über das (Nicht-)Wissen zum Vorliegen milde verlaufender SMA-Formen – ist bei Einführung eines Neugeborenencreenings zu berücksichtigen.

Inhaltsverzeichnis

	Seite
Kernaussage	iii
Tabellenverzeichnis	xi
Abbildungsverzeichnis	xiii
Abkürzungsverzeichnis	xiv
1 Hintergrund	1
2 Fragestellung	3
3 Methoden	4
4 Ergebnisse	7
4.1 Ergebnisse der umfassenden Informationsbeschaffung	7
4.2 Vergleichende Interventionsstudien der Screeningkette	7
4.3 Vergleichende Interventionsstudien zum Therapiebeginn	7
4.3.1 Charakteristika der in die Bewertung eingeschlossenen Studien	7
4.3.1.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn.....	7
4.3.1.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn.....	9
4.3.2 Übersicht der patientenrelevanten Endpunkte	11
4.3.3 Bewertung des Verzerrungspotenzials der Ergebnisse.....	13
4.3.3.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn.....	13
4.3.3.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn.....	13
4.3.4 Ergebnisse zu patientenrelevanten Endpunkten	13
4.3.4.1 Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung	13
4.3.4.1.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	13
4.3.4.1.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn	14
4.3.4.2 Ergebnisse zum Gesamtüberleben	14
4.3.4.2.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	14
4.3.4.2.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn	14
4.3.4.3 Ergebnisse zum krankheitsspezifischen Überleben	14
4.3.4.3.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	14
4.3.4.3.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn	15
4.3.4.4 Ergebnisse zum Erreichen motorischer Meilensteine.....	15
4.3.4.4.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	15
4.3.4.4.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn	16
4.3.4.5 Ergebnisse zu Entwicklungs- und Wachstumsstörungen.....	17

4.3.4.5.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	17
4.3.4.5.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	18
4.3.4.6	Ergebnisse zu Krankenhausaufenthalten	18
4.3.4.6.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	18
4.3.4.6.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	18
4.3.4.7	Ergebnisse zu Atemnot	18
4.3.4.7.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	18
4.3.4.7.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	18
4.3.4.8	Ergebnisse zur dauerhaften Beatmung.....	18
4.3.4.8.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	18
4.3.4.8.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	18
4.3.4.9	Ergebnisse zu schwerwiegenden respiratorischen Ereignissen	19
4.3.4.9.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	19
4.3.4.9.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	19
4.3.4.10	Ergebnisse zu (schwerwiegenden) unerwünschten Ereignissen	19
4.3.4.10.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	19
4.3.4.10.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	19
4.3.4.11	Ergebnisse zur gesundheitsbezogenen Lebensqualität des Kindes.....	20
4.3.4.11.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	20
4.3.4.11.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	20
4.4	Studien zur diagnostischen Güte	20
4.4.1	Charakteristika der in die Bewertung eingeschlossenen Studien	20
4.4.2	Vorhandene bewertungsrelevante Zielgrößen	22
4.4.3	Bewertung des Verzerrungspotenzials und der Übertragbarkeit.....	22
4.4.4	Ergebnisse zu den Zielgrößen.....	22
4.5	Zusammenfassung der Beleglage.....	23
5	Einordnung des Arbeitsergebnisses	25
6	Fazit.....	31
Details des Berichts		33
A1	Projektverlauf	33
A1.1	Zeitlicher Verlauf des Projekts	33
A1.2	Spezifizierungen und Änderungen im Projektverlauf.....	33
A2	Methodik gemäß Berichtsplan 1.0.....	35
A2.1	Kriterien für den Einschluss von vergleichenden Interventionsstudien der Screeningkette in die Untersuchung.....	36
A2.1.1	Population.....	36
A2.1.2	Prüf- und Vergleichsintervention	36
A2.1.3	Patientenrelevante Endpunkte	36

A2.1.4	Studientypen	37
A2.1.5	Studiendauer	38
A2.1.6	Publikationssprache	38
A2.1.7	Tabellarische Darstellung der Kriterien für den Studieneinschluss	38
A2.2	Kriterien für den Einschluss von vergleichenden Interventionsstudien zum Therapiebeginn in die Untersuchung	38
A2.2.1	Population.....	39
A2.2.2	Prüf- und Vergleichsintervention	39
A2.2.3	Patientenrelevante Endpunkte	39
A2.2.4	Studientypen	39
A2.2.5	Studiendauer	40
A2.2.6	Publikationssprache	40
A2.2.7	Tabellarische Darstellung der Kriterien für den Studieneinschluss	40
A2.3	Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Untersuchung.....	40
A2.3.1	Population.....	41
A2.3.2	Indextest.....	41
A2.3.3	Referenztest	41
A2.3.4	Zielgrößen.....	41
A2.3.5	Studientypen	41
A2.3.6	Studiendauer	42
A2.3.7	Publikationssprache	42
A2.3.8	Tabellarische Darstellung der Kriterien für den Studieneinschluss	42
A2.4	Vorgehen im Falle einer Zulassungsänderung im Projektverlauf	42
A2.5	Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen	43
A2.6	Umfassende Informationsbeschaffung	43
A2.6.1	Informationsquellen.....	43
A2.6.2	Selektion relevanter Studien.....	44
A2.7	Informationsbewertung und Synthese	45
A2.7.1	Darstellung der Einzelstudien.....	45
A2.7.1.1	Darstellung der Einzelstudien (vergleichende Interventionsstudien)	45
A2.7.1.2	Darstellung der Einzelstudien (Studien zur diagnostischen Güte).....	45
A2.7.2	Bewertung des Verzerrungspotenzials der Ergebnisse.....	46
A2.7.2.1	Bewertung des Verzerrungspotenzials der Ergebnisse (vergleichende Interventionsstudien).....	46
A2.7.2.2	Bewertung des Verzerrungspotenzials der Ergebnisse (Studien zur diagnostischen Güte).....	47
A2.7.3	Metaanalysen	47

A2.7.3.1	Metaanalysen (vergleichende Interventionsstudien).....	47
A2.7.3.2	Metaanalysen (Studien zur diagnostischen Güte).....	48
A2.7.4	Sensitivitätsanalysen.....	48
A2.7.5	Subgruppenmerkmale und andere Effektmodifikatoren.....	49
A2.7.6	Aussagen zur Beleglage.....	49
A3	Details der Ergebnisse	51
A3.1	Umfassende Informationsbeschaffung	51
A3.1.1	Primäre Informationsquellen	51
A3.1.1.1	Bibliografische Datenbanken.....	51
A3.1.1.2	Studienregister	53
A3.1.1.3	Herstellieranfragen.....	54
A3.1.2	Weitere Informationsquellen und Suchtechniken.....	56
A3.1.2.1	Zulassungsbehörden.....	57
A3.1.2.2	Durch den G-BA übermittelte Dokumente	57
A3.1.2.3	G-BA-Website und IQWiG-Website	57
A3.1.2.4	Anwendung weiterer Suchtechniken	57
A3.1.2.5	Anhörung	57
A3.1.2.6	Autorenanfragen.....	58
A3.1.3	Resultierender Studienpool.....	58
A3.1.4	Studien ohne berichtete Ergebnisse	58
A3.2	Vergleichende Interventionsstudien zur Screeningkette.....	59
A3.3	Vergleichende Interventionsstudien zum Therapiebeginn.....	59
A3.3.1	Charakteristika der in die Bewertung eingeschlossenen Studie	59
A3.3.1.1	Studiendesign und Studienpopulationen.....	59
A3.3.1.1.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	59
A3.3.1.1.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	64
A3.3.1.2	Bewertung endpunktübergreifender Kriterien des Verzerrungspotenzials.....	69
A3.3.1.2.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	69
A3.3.1.2.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	69
A3.3.2	Patientenrelevante Endpunkte	69
A3.3.2.1	Kombinierter Endpunkt Zeit bis Tod oder dauerhafte Beatmung	69
A3.3.2.1.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	69
A3.3.2.1.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	72
A3.3.2.2	Gesamtüberleben.....	72
A3.3.2.2.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	72
A3.3.2.2.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	72
A3.3.2.3	Krankheitsspezifisches Überleben	73

A3.3.2.3.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	73
A3.3.2.3.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	73
A3.3.2.4	Erreichen motorischer Meilensteine	73
A3.3.2.4.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	73
A3.3.2.4.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	76
A3.3.2.5	Entwicklungs- und Wachstumsstörungen.....	79
A3.3.2.5.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	79
A3.3.2.5.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	79
A3.3.2.6	Krankenhausaufenthalte.....	79
A3.3.2.6.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	79
A3.3.2.6.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	79
A3.3.2.7	Atemnot.....	79
A3.3.2.7.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	79
A3.3.2.7.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	79
A3.3.2.8	Dauerhafte Beatmung	79
A3.3.2.8.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	79
A3.3.2.8.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	80
A3.3.2.9	Schwerwiegende respiratorische Ereignisse	80
A3.3.2.9.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	80
A3.3.2.9.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	81
A3.3.2.10	(Schwerwiegende) unerwünschte Ereignisse.....	81
A3.3.2.10.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	81
A3.3.2.10.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	85
A3.3.2.11	Gesundheitsbezogene Lebensqualität des Kindes.....	86
A3.3.2.11.1	Frühsymptomatischer versus spätsymptomatischen Therapiebeginn ..	86
A3.3.2.11.2	Präsymptomatischer versus frühsymptomatischen Therapiebeginn	86
A3.3.2.12	Metaanalysen	86
A3.3.2.13	Sensitivitätsanalysen	86
A3.3.2.14	Subgruppenmerkmale und andere Effektmodifikatoren	86
A3.4	Studien zur diagnostischen Güte	86
A3.4.1	Charakteristika der in die Bewertung eingeschlossenen Studien	86
A3.4.1.1	Studiendesign und Studienpopulationen.....	86
A3.4.1.2	Einschätzung des Verzerrungspotenzials.....	89
A3.4.2	Ergebnisse zu den Zielgrößen.....	90
A3.4.2.1	Ergebnisse zum positiven prädiktiven Wert	90
A3.4.2.2	Sensitivitätsanalysen	92
A3.4.2.3	Subgruppenmerkmale und andere Effektmodifikatoren	92

A4	Kommentare.....	93
A4.1	Bericht im Vergleich zu anderen systematischen Übersichten	93
A4.2	Kritische Reflexion des Vorgehens	93
A4.3	Würdigung der Anhörung.....	96
A5	Literatur	133
A6	Studienlisten	142
A6.1	Liste der gesichteten systematischen Übersichten	142
A6.1.1	Vergleichende Interventionsstudien zum Therapiebeginn	142
A6.2	Liste der ausgeschlossenen Publikationen mit Ausschlussgründen	142
A6.2.1	Vergleichende Interventionsstudien der Screeningkette sowie Studien zur diagnostischen Güte.....	142
A6.2.2	Vergleichende Interventionsstudien zum Therapiebeginn	148
A6.3	Liste der ausgeschlossenen Dokumente aus den durch den G-BA übermittelten Dokumenten mit Ausschlussgründen.....	152
A7	Suchstrategien	153
A7.1	Suchstrategien in bibliografischen Datenbanken.....	153
A7.1.1	Suchstrategien für vergleichende Interventionsstudien der Screeningkette und Studien zur diagnostischen Güte	153
A7.1.2	Suchstrategien für vergleichende Interventionsstudien zum Therapiebeginn .	156
A7.2	Suche in Studienregistern.....	159
A8	Offenlegung potenzieller Interessenkonflikte der externen Sachverständigen.....	161

Tabellenverzeichnis

	Seite
Tabelle 1: Matrix der patientenrelevanten Endpunkte	12
Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien der Screeningkette)	38
Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien zum Therapiebeginn)	40
Tabelle 4: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)	42
Tabelle 5: Regelmäßig abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit	50
Tabelle 6: In Studienregistern identifizierte relevante Studien bzw. Dokumente	54
Tabelle 7: Durch Hersteller übermittelte relevante Studien bzw. Dokumente	54
Tabelle 8: Übersicht über Herstelleranfragen	55
Tabelle 9: In vom G-BA übermittelten Dokumenten identifizierte relevante Studien bzw. Dokumente	57
Tabelle 10: In IQWiG-Produkten, öffentlich zugänglichen Herstellerunterlagen oder Beschlüssen des G-BA identifizierte relevante Studien bzw. Dokumente	57
Tabelle 11: Studienpool der Nutzenbewertung	58
Tabelle 12: In der Informationsbeschaffung identifizierte Studien ohne berichtete Ergebnisse	59
Tabelle 13: Charakterisierung der eingeschlossenen Studie (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	61
Tabelle 14: Charakterisierung der Interventionen in der eingeschlossenen Studie (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	62
Tabelle 15: Ein- / Ausschlusskriterien für Patientinnen und Patienten in der Studie (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	62
Tabelle 16: Charakterisierung der Studienpopulationen (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	63
Tabelle 17: Charakterisierung der eingeschlossenen Studie (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)	65
Tabelle 18: Charakterisierung der Interventionen in der eingeschlossenen Studie (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)	66
Tabelle 19: Ein- / Ausschlusskriterien für Patientinnen und Patienten in der Studie (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)	67
Tabelle 20: Charakterisierung der Studienpopulation (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)	68
Tabelle 21: Endpunktübergreifendes Verzerrungspotenzial (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	69

Tabelle 22: Bewertung des endpunktspezifischen Verzerrungspotenzials: kombinierter Endpunkt Zeit bis Tod oder dauerhafte Beatmung (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	70
Tabelle 23: Ergebnisse – kombinierter Endpunkt Zeit bis Tod oder dauerhafte Beatmung (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	71
Tabelle 24: Ergebnisse – Mortalität (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)	72
Tabelle 25: Bewertung des endpunktspezifischen Verzerrungspotenzials: Erreichen motorischer Meilensteine (Vergleich früh- vs. spätsymptomatischen Therapiebeginn).....	74
Tabelle 26: Ergebnisse – Responder Erreichen motorischer Meilensteine HINE-2 (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	75
Tabelle 27: Ergebnisse – HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)	76
Tabelle 28: Ergebnisse – dauerhafte Beatmung (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)	80
Tabelle 29: Bewertung des endpunktspezifischen Verzerrungspotenzials: schwerwiegende unerwünschte Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	81
Tabelle 30: Ergebnisse – schwerwiegende unerwünschte Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	82
Tabelle 31: Bewertung des endpunktspezifischen Verzerrungspotenzials: schwere unerwünschte Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	82
Tabelle 32: Ergebnisse – schwere unerwünschte Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	83
Tabelle 33: Bewertung des endpunktspezifischen Verzerrungspotenzials: Therapieabbrüche wegen unerwünschter Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	83
Tabelle 34: Ergebnisse – Therapieabbrüche wegen unerwünschter Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)	84
Tabelle 35: Ergebnisse – (schwerwiegende) unerwünschte Ereignisse (Vergleich prä- vs. frühsymptomatischen Therapiebeginn).....	85
Tabelle 36: Charakterisierung der eingeschlossenen Studien zur diagnostischen Güte	87
Tabelle 37: Ein- / Ausschlusskriterien für Neugeborene in den Studien zur diagnostischen Güte	88
Tabelle 38: Indextest und Referenzstandard – Studien zur diagnostischen Güte	88
Tabelle 39: Verzerrungspotenzial nach QUADAS 2	89
Tabelle 40: Bedenken bezüglich der Übertragbarkeit QUADAS 2	89
Tabelle 41: Ergebnisse der Studien zur diagnostischen Güte	91
Tabelle 42: Würdigung der Stellungnahmen	97

Abbildungsverzeichnis

	Seite
Abbildung 1: Ergebnis der bibliografischen Recherche und der Studienselektion für die Suche nach vergleichenden Interventionsstudien der Screeningkette sowie nach Studien zur diagnostischen Güte	52
Abbildung 2: Ergebnis der bibliografischen Recherche und der Studienselektion für die Suche nach vergleichenden Interventionsstudien zum Therapiebeginn	53
Abbildung 3: Responder Erreichen des Maximalscores beim Einzelitem Drehen („dreht sich im Sitzen“) des HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn	77
Abbildung 4: Responder Erreichen des Maximalscores beim Einzelitem Krabbeln („krabbelt auf Händen und Knien“) des HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn.....	77
Abbildung 5: Responder Erreichen des Maximalscores beim Einzelitem Stehen („steht ohne Hilfe“) des HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn.....	78
Abbildung 6: Responder Erreichen des Maximalscores beim Einzelitem Gehen („geht selbstständig“) des HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn	78

Abkürzungsverzeichnis

Abkürzung	Bedeutung
aRQ	Average relative Quantity (relative Durchschnittsmenge)
BSC	Best supportive Care
CHOP	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
Cq	Cycle of quantification (Quantifizierungszyklus)
ddPCR	Droplet digital Polymerase Chain Reaction (digitale Tröpfchen-Polymerase-Kettenreaktion)
DNA	Desoxyribonukleinsäure
EMA	European Medicines Agency (Europäische Arzneimittel-Agentur)
FDA	U. S. Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
HINE-2	Hammersmith Infant Neurological Examination – Subscale 2 (Hammersmith neurologische Untersuchung bei Säuglingen – Subskala 2)
HR	Hazard Ratio
ICD-10	International Statistical Classification of Diseases and Related Health Problems, Revision 10 (Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Ausgabe)
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
MD	Mittelwertdifferenz
MID	Minimal important Difference
MLPA	Multiplex Ligation-dependent Probe Amplification (multiplexe ligationsabhängige Sondenamplifikation)
PCR	Polymerase Chain Reaction (Polymerase-Kettenreaktion)
PPV	Positive predictive Value (positiver prädiktiver Wert)
qPCR	Quantitative Real-Time Polymerase Chain Reaction (quantitative Echtzeit-Polymerase-Kettenreaktion)
RCT	Randomized controlled Trial (randomisierte kontrollierte Studie)
RT-PCR	Real-Time Polymerase Chain Reaction (Echtzeit-Polymerase-Kettenreaktion)
SD	Standardabweichung
SMA	spinale Muskelatrophie
SMN	Survival Motor Neuron
SMN1	Survival Motor Neuron 1
SMN2	Survival Motor Neuron 2
SOC	System Organ Class (Systemorganklasse)
SUE	schwerwiegendes unerwünschtes Ereignis

Abkürzung	Bedeutung
UE	unerwünschtes Ereignis
VOPT	Verification of only positive Testers

1 Hintergrund

Die 5q-assoziierte spinale Muskelatrophie (SMA) ist eine autosomal-rezessiv vererbte Erkrankung. Sie tritt selten auf; ihre Prävalenz wird mit 1–2:100 000 [1] bis zu 1:30 000 [2,3] angegeben, die Inzidenz für den SMA-Typ I mit 1:10 000 [1,2].

Für die Krankheit typisch ist der fortschreitende Untergang von motorischen Nervenzellen (Motoneuronen), der zu Muskelschwund und -schwäche führt.

Motoneuronen sind Nervenzellen, die spezifisch den Muskel stimulieren und so Muskelkontraktionen für willkürliche Bewegungen auslösen. Für das Fortbestehen dieser Nervenzellen ist das Survival-Motor-Neuron(SMN)-Protein unabdingbar. Für dieses SMN-Protein codiert das *SMN1*-Gen. Sind auf beiden Allelen des *SMN1*-Gens (homozygot) Deletionen zu finden, verursacht dies 5q-assoziierte SMA [4,5]. Das benachbarte und sehr ähnliche Gen *SMN2* produziert ebenfalls funktionierendes SMN-Protein, allerdings in geringerer Menge als das Gen *SMN1*. In den meisten Fällen tritt die Erkrankung in dieser homozygoten Form auf. Etwa 5 % der Patientinnen und Patienten weisen dagegen eine andere Auffälligkeit auf (heterozygote Deletion und Punktmutation auf dem vorhandenen *SMN1*-Gen) [6-8].

Typische Symptome einer 5q-assoziierten SMA sind Muskelschwäche, eine verzögerte motorische Entwicklung und je nach Typ teils schwerwiegende Einschränkungen der Lungenfunktion. In der Literatur werden 4 Typen unterschieden [9], die aber nicht trennscharf zu unterscheiden sind [10,11]. Die Typen unterscheiden sich in den Aspekten Symptombeginn und -schwere, erreichte motorische Meilensteine und Anzahl der *SMN2*-Kopien. So beginnt die Krankheit beim SMA-Typ I in den ersten Lebenswochen, die Kinder erlernen das freie Sitzen nicht. Ohne (supportive) Therapie versterben die Kinder meist bis zum Alter von 1 bis 2 Jahren [6,12,13]. Bei einem Krankheitsbeginn nach dem 6. Lebensmonat wird vom SMA-Typ II gesprochen. Hier wird eine höhere Lebenserwartung [6,12], teils von 10 bis 40 Jahren [13], beschrieben. Kinder mit dem SMA-Typ II können das freie Sitzen erlernen, nicht aber stehen und / oder gehen [12]. Bei Patientinnen und Patienten mit dem SMA-Typ III scheint die Lebenserwartung nicht eingeschränkt [13]. Die Patientinnen und Patienten erlernen das Laufen, verlieren diese Fähigkeit aber im Laufe ihres Lebens durch den progressiven Muskelschwund. Mildere Formen werden als SMA-Typ IV bezeichnet. Der SMA-Typ 0 bezeichnet den pränatalen Erkrankungsbeginn. Die Kinder bedürfen mit der Geburt einer Unterstützung der Atmung zum Überleben [6,11-14]. Der Zeitraum zwischen Symptombeginn und Diagnose beträgt durchschnittlich 3,6 Monate beim SMA-Typ I, 14,3 Monate beim SMA-Typ II und 43,6 Monate beim SMA-Typ III [15]. Die Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Ausgabe (ICD-10) unterteilt dieses Krankheitsbild in eine infantile SMA und eine sonstige vererbte SMA [16]. Während die infantile dem SMA-Typ I entspricht, werden unter der sonstigen vererbten SMA die Typen II, III und IV mit jeweils späterem Krankheitsbeginn zusammengefasst.

Die medikamentöse Therapie der 5q-assozierten SMA umfasst die genterapeutische Einbringung des *SMN1*-Gens sowie die *SMN2*-Modifikation. Durch die *SMN2*-Genmodifikation soll der vom *SMN2*-Gen codierte, funktionelle SMN-Proteinanteil erhöht werden. Diesen Ansatz verfolgt das einzige derzeit in Deutschland seit 2017 für die Therapie der 5q-assozierten SMA zugelassene Arzneimittel Nusinersen. Nusinersen wird intrathekal durch eine Punktion des Nervenwasserraumes (Lumbalkanalpunktion) verabreicht [17,18]. Einen weiteren Ansatz stellen SMN-unabhängige Therapien dar (beispielsweise Neuroprotektoren). Valproinsäure bei Kindern und andere SMN-unabhängige Behandlungsstrategien werden derzeit erforscht (beispielsweise [19]). Im Mai 2019 wurde die Genersatztherapie AVXS-101 [20] von der U. S. Food and Drug Administration (FDA) in den USA für Kinder unter 2 Jahren zugelassen [21]. Auf europäischer Ebene ist hierzu bislang keine Zulassung erfolgt. Die multidisziplinäre Therapie umfasst rehabilitative, orthopädische und psychologische Maßnahmen, Maßnahmen zum Muskelerhalt sowie die Unterstützung der Atmung und der Ernährung.

Die 5q-assozierte SMA lässt sich präsymptomatisch über eine Blutprobe diagnostizieren: Dazu wird die homozygote *SMN1*-Gen-Deletion nachgewiesen. Zusätzlich wird die Anzahl der vorhandenen Kopien des *SMN2*-Gens, das im Zusammenhang mit dem Schweregrad der Erkrankung steht, bestimmt [11,22,23].

Zur SMA-Diagnostik kann auf Filterpapierkarten aufgetropftes getrocknetes Blut verwendet werden. Beim in Deutschland gemäß der Kinder-Richtlinie des Gemeinsamen Bundesausschusses (G-BA) [24] durchgeführten erweiterten Neugeborenencreening wird in der 36. bis 72. Lebensstunde Venen- oder Fersenblut gewonnen, auf Filterpapierkarten aufgetropft und auf Zielerkrankungen untersucht. Die 5q-assozierte SMA gehört bisher nicht zu den Zielerkrankungen, die im Rahmen des erweiterten Neugeborenencreenings untersucht werden.

Ziel eines Neugeborenencreenings auf 5q-assozierte SMA ist die frühere Identifikation und Behandlung von betroffenen Kindern.

2 Fragestellung

Das Ziel der vorliegenden Untersuchung ist die Nutzenbewertung eines Neugeborenen-screensings auf 5q-assoziierte spinale Muskelatrophie (SMA). Dabei wird das Neugeborenen-screening auf 5q-assoziierte SMA in Kombination mit einer Vorverlegung der Diagnosestellung und Behandlung im Vergleich zu keinem 5q-assoziierte-SMA-Screening hinsichtlich patientenrelevanter Endpunkte bewertet.

3 Methoden

In die Nutzenbewertung wurden vergleichende Studien der Screeningkette eingeschlossen. Für den Fall, dass solche Studien nicht oder in nicht ausreichender Quantität und Qualität vorliegen sollten, war eine Bewertung von Interventionsstudien, die einen Vergleich eines früheren gegenüber einem späteren Therapiebeginn ermöglichen, sowie von Studien zur diagnostischen Güte als die einzelnen Bausteine der Screeningkette vorgesehen (Linked Evidence).

Vergleichende Interventionsstudien der Screeningkette

Die Zielpopulation der Nutzenbewertung bildeten Neugeborene. Die Prüflintervention bildete das Neugeborenencreening auf 5q-assozierte SMA in Kombination mit einer Vorverlegung der Diagnosestellung und Behandlung. Als Vergleichsintervention galt keine Screeningstrategie.

Für die Untersuchung wurden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (Gesamtüberleben, krankheitsspezifisches Überleben),
- Morbidität (zum Beispiel Entwicklungs- und Wachstumsstörungen wie das Erreichen motorischer Meilensteine, Krankenhausaufenthalte, durch 5q-assozierte SMA hervorgerufene Atemnot),
- (schwerwiegende) unerwünschte Ereignisse ([S]UEs),
- gesundheitsbezogene Lebensqualität des Kindes.

Es sollten randomisierte kontrollierte Studien (RCTs) in die Nutzenbewertung eingeschlossen werden. Sofern die auf RCTs basierende Datenlage zur Nutzenbewertung nicht ausreichte, sollten auch nicht randomisierte vergleichende Interventionsstudien und vergleichende Kohortenstudien (auch retrospektive oder mit historischem Vergleich) eingeschlossen werden. Hinsichtlich der Studiendauer bestand keine Einschränkung.

Vergleichende Interventionsstudien zum Therapiebeginn

Sofern keine vergleichende Interventionsstudie der Screeningkette identifiziert wurde, wurden in die Bewertung auch Interventionsstudien mit Patientinnen und Patienten mit 5q-assoziierter SMA aufgenommen, die einen Vergleich eines früheren gegenüber einem späteren Therapiebeginn mit Nusinersen ermöglichten. Die Diagnosestellung bei Patientinnen und Patienten mit früherem Therapiebeginn musste auf die Screeningsituation bei Neugeborenen übertragbar sein. Für die zu prüfende Intervention wurden 2 verschiedene Zeitpunkte des Therapiebeginns berücksichtigt. Zum einen wurde (a) ein präsymptomatischer Behandlungsbeginn als zu prüfende Intervention berücksichtigt. Zum anderen wurde (b) ein früherer Therapiebeginn bei bereits vorliegender Symptomatik (im Folgenden: frühsymptomatisch) als zu prüfende Intervention herangezogen. Als Vergleichsintervention galt jeweils ein späterer Therapiebeginn bei vorliegender Symptomatik (im Folgenden: [spät]symptomatisch). Für die Untersuchung wurden die oben genannten patientenrelevanten Endpunkte betrachtet. Es wurden RCTs in die

Nutzenbewertung eingeschlossen. Sofern zu einer Fragestellung keine RCTs vorlagen, wurden Studien mit niedrigerem Evidenzniveau (retrospektive vergleichende Studien) in die Nutzenbewertung eingeschlossen. Hinsichtlich der Studiendauer bestand keine Einschränkung.

Studien zur diagnostischen Güte

Sofern sich aus der Vorverlegung des Therapiebeginns eine positive Aussage zum Nutzen ergab (siehe Abschnitt „Informationsbeschaffung, Informationsbewertung und Synthese“), wurden im Rahmen des vorliegenden Berichts auch Studien zur diagnostischen Güte zur Nutzenbewertung herangezogen. In der Bewertung wurden Studien mit Neugeborenen eingeschlossen. Der Indextest war die Testung auf 5q-assozierte SMA unter Verwendung von Filterpapierkarten. Den Referenztest bildeten genetische Analysen. Bei unauffälligem Befund im Indextest konnte alternativ auch die Nachbeobachtung akzeptiert werden. Eingeschlossen wurden Studien, aus denen Daten zur Berechnung der diagnostischen Güte im Hinblick auf die Entdeckung von 5q-assoziierter SMA ableitbar waren.

Informationsbeschaffung, Informationsbewertung und Synthese

Eine systematische Literaturrecherche nach Studien wurde in den Datenbanken MEDLINE, Embase und Cochrane Central Register of Controlled Trials durchgeführt. Parallel erfolgte eine Suche nach relevanten systematischen Übersichten in den Datenbanken MEDLINE, Embase, Cochrane Database of Systematic Reviews und HTA Database.

Darüber hinaus wurden folgende Informationsquellen und Suchtechniken berücksichtigt: Studienregister, Herstelleranfragen, öffentlich zugängliche Dokumente von Zulassungsbehörden, vom Gemeinsamen Bundesausschuss (G-BA) übermittelte Dokumente, die G-BA-Website und IQWiG-Website, die Sichtung von Referenzlisten, aus Anhörungsverfahren zur Verfügung gestellte Dokumente und Autorenanfragen.

Die Selektion relevanter Studien erfolgte von 3 Personen unabhängig voneinander. Die Ergebnisse der Selektion wurden nach der Volltextbewertung zusammengefasst. Die Datenextraktion erfolgte in standardisierte Tabellen. Zur Einschätzung der qualitativen Ergebnissicherheit wurden endpunktübergreifende und endpunktspezifische Kriterien des Verzerrungspotenzials bewertet und dieses jeweils in niedrig oder hoch eingestuft. Bei Studien zur diagnostischen Güte wurde neben dem Verzerrungspotenzial der Ergebnisse zudem die Übertragbarkeit auf die deutsche Versorgungssituation überprüft. Die Ergebnisse der einzelnen Studien wurden nach Endpunkten geordnet beschrieben.

Sofern die Studien hinsichtlich der Fragestellung und relevanter Charakteristika vergleichbar waren und keine bedeutsame Heterogenität beobachtet wurde, sollten die Einzelergebnisse mithilfe von Metaanalysen quantitativ zusammengefasst werden.

Für jeden Endpunkt wurde eine Aussage zur Beleglage des (höheren) Nutzens und (höheren) Schadens in 4 Abstufungen bezüglich der jeweiligen Aussagesicherheit getroffen: Es lag entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein

Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall trat ein, wenn keine Daten vorlagen oder die vorliegenden Daten keine der 3 übrigen Aussagen zuließen. In diesem Fall wurde die Aussage „Es liegt kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden vor“ getroffen. Eine positive Aussage zum Nutzen der Therapievorverlegung wurde in der vorliegenden Bewertung insofern spezifiziert und weit gefasst, als hierfür auch RCTs berücksichtigt wurden, die mittels Subgruppenanalysen den Vergleich eines früheren versus einen späteren Therapiebeginn ermöglichten. Eine positive Aussage zum Nutzen der Therapievorverlegung konnte sich hierbei durch eine statistisch signifikante Interaktion (Anhaltspunkt für Unterschiede der Effekte je nach Therapiebeginn) bei gleichzeitigem Vorliegen statistisch signifikanter Behandlungseffekte zugunsten der Intervention mindestens in der Gruppe mit früherem Therapiebeginn ergeben. Aussagen zum Nutzen auf Basis von Studien mit niedrigerer Evidenzstufe (retrospektive vergleichende Studien) waren nur möglich, wenn die vorliegenden Effekte so groß waren, dass sie sich nicht allein durch die Einwirkung von Störgrößen erklären ließen (dramatischer Effekt). Der Nutzen des Screenings wurde durch die Gegenüberstellung der gesundheitsbezogenen Konsequenzen der möglichen Testergebnisse und ihrer Wahrscheinlichkeiten zusammen mit einer Aussage zum Nutzen eines früheren Therapiebeginns abgeleitet (Linked-Evidence-Ansatz). Die Aussagesicherheit bezüglich des Nutzens des Screenings berücksichtigte auf diese Weise sowohl die Aussagesicherheit bezüglich des Nutzens eines früheren Therapiebeginns als auch bezüglich der diagnostischen Güte.

4 Ergebnisse

4.1 Ergebnisse der umfassenden Informationsbeschaffung

Die Informationsbeschaffung identifizierte keine Studie zur Screeningkette als relevant für die Fragestellung der vorliegenden Nutzenbewertung. Es wurden keine geplanten oder laufenden Studien zur Screeningkette identifiziert. Die letzte Suche nach Studien zur Screeningkette fand am 23.10.2019 statt.

Die Informationsbeschaffung identifizierte 1 Studie (14 Dokumente) mit Daten zum Vergleich eines früh- versus einen spätsymptomatischen Therapiebeginn als relevant für die Fragestellung der vorliegenden Nutzenbewertung. Des Weiteren wurden für den Vergleich eines präsymptomatischen mit einem frühsymptomatischen Therapiebeginn Daten vom Hersteller zur Verfügung gestellt, die im Sinne von 1 retrospektiven vergleichenden Studie (1 Dokument) als relevant für die vorliegende Nutzenbewertung eingeschlossen wurden. Weitere geplante oder laufende vergleichende Interventionsstudien zum Therapiebeginn wurden nicht identifiziert. Die letzte Suche nach vergleichenden Interventionsstudien zum Therapiebeginn fand am 23.10.2019 statt.

Die Informationsbeschaffung identifizierte 4 Studien (7 Dokumente) zur diagnostischen Güte als relevant für die Fragestellung der vorliegenden Nutzenbewertung. Es wurden 3 laufende Studien zur diagnostischen Güte identifiziert. Die letzte Suche nach Studien zur diagnostischen Güte fand am 23.10.2019 statt.

4.2 Vergleichende Interventionsstudien der Screeningkette

Vergleichende Interventionsstudien der Screeningkette konnten nicht identifiziert werden. Daher erfolgte eine Bewertung der einzelnen Bausteine der Screeningkette – zum einen anhand von vergleichenden Interventionsstudien zum Therapiebeginn (Abschnitt 4.3), zum anderen anhand von Studien zur diagnostischen Güte (Abschnitt 4.4).

4.3 Vergleichende Interventionsstudien zum Therapiebeginn

Zum Vergleich eines früh- versus einen spätsymptomatischen Therapiebeginn konnte 1 RCT in die Nutzenbewertung eingeschlossen werden. Zum Vergleich eines präsymptomatischen versus einen frühsymptomatischen Therapiebeginn wurde 1 retrospektive vergleichende Studie in die Nutzenbewertung eingeschlossen. Beide werden im Folgenden dargestellt.

4.3.1 Charakteristika der in die Bewertung eingeschlossenen Studien

4.3.1.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Es wurde 1 Studie mit Daten zum Vergleich eines früh- versus einen spätsymptomatischen Therapiebeginn herangezogen (ENDEAR) [25]. Es handelt sich hierbei um 1 RCT zum Vergleich einer Nusinersenbehandlung versus eine Scheinbehandlung bei Kindern mit infantiler 5q-assoziierter SMA. Aus dieser konnten Daten zu dem Vergleich eines früh- versus einen spätsymptomatischen Therapiebeginn aus geplanten Subgruppenanalysen herangezogen

werden. In der vorliegenden Nutzenbewertung wurde dieser Vergleich im Sinne einer Vorverlegung des Therapiebeginns gedeutet. Zur Einbettung dieser Daten wird zunächst der Kontext der Studie dargestellt und werden anschließend die herangezogenen Subgruppen charakterisiert.

Die Studie ENDEAR

Die Studie wurde multizentrisch mit insgesamt 31 Zentren in Nordamerika, Europa, Asien und Australien durchgeführt. Eingeschlossen wurden 122 Kinder mit einem frühen Krankheitsbeginn der 5q-assoziierten SMA (2 *SMN2*-Genkopien), d. h. mit einem Alter bei Symptombeginn ≤ 6 Monate, und einem Alter bei Studieneinschluss ≤ 7 Monate. Die 122 Kinder mit infantiler SMA wurden im Verhältnis 2:1 auf eine Behandlung mit Nusinersen oder eine Scheinbehandlung stratifiziert nach der Krankheitsdauer (≤ 12 oder > 12 Wochen zwischen Alter des Kindes bei Symptombeginn und Alter bei Studieneinschluss) randomisiert. Für 1 Kind wurde die Einwilligung zur Teilnahme zurückgezogen, sodass 121 Kinder in die Intention-to-treat(ITT)-Analyse eingingen.

In der Nusinersengruppe erhielten die Kinder Nusinersen als intrathekale Bolusinjektion mit altersadjustierter Dosierung an den Studientagen 1, 15, 29, 64 (Aufsättigung) sowie 183 und 302 (Erhaltung). Die Kontrollgruppe erhielt stattdessen eine Scheinbehandlung in Form eines Nadelstichs am unteren Rücken (keine Lumbalpunktion) zu den entsprechenden Zeitpunkten. In beiden Behandlungsgruppen erhielten die Kinder nach Bedarf eine Behandlung nach Best supportive Care (BSC).

Insgesamt war eine 10-monatige Behandlungsdauer der Kinder geplant. Daran sollte sich eine 3-monatige Nachbeobachtung anschließen. Die Studie wurde jedoch nach einer geplanten Interimsanalyse vorzeitig beendet aufgrund eines Wirksamkeitsnachweises von Nusinersen. Die mittlere Beobachtungsdauer zum finalen Datenschnitt am 16.12.2016 betrug 40 (Minimum 1; Maximum 63) Wochen in der Nusinersengruppe und 27 (Minimum 3; Maximum 60) Wochen in der Scheinbehandlungsgruppe. Der Wirksamkeitsnachweis wurde durch positive Effekte der Nusinersentherapie für den Endpunkt Erreichen motorischer Meilensteine begründet, der unter anderem mittels der Hammersmith Infant Neurological Examination – Subscale 2 (HINE-2) erhoben und als Responderanalyse ausgewertet wurde. Dieser Endpunkt wurde nachträglich als primärer Endpunkt der Studie definiert. Die Begründung für diese Protokolländerung war, dass Phase-2-Daten darauf hindeuteten, dass eine funktionelle Antwort einen frühen Nachweis der Wirksamkeit liefern könne und somit eine frühere Interimsanalyse ermögliche [26]. Der ursprünglich primäre und final koprimäre Endpunkt war der kombinierte Endpunkt Zeit bis Tod oder dauerhafte Beatmung. Dieser umfasste die Ereignisse Tod und dauerhafte Beatmung, die definiert war als Beatmung über ≥ 16 Stunden pro Tag durchgehend für > 21 Tage bei Abwesenheit von akuten reversiblen Ereignissen oder Tracheotomie. Die akuten reversiblen Ereignisse waren prädefiniert und jeder Fall wurde durch ein verblindetes, zentrales und unabhängiges Gremium (Endpoint Adjudication Committee) geprüft. Weitere patientenrelevante Endpunkte umfassten u. a. weitere Operationalisierungen zu den End-

punkten Erreichen motorischer Meilensteine, Entwicklungs- und Wachstumsstörungen, Krankenhausaufenthalte, SUEs, schwere UEs und Therapieabbrüche wegen UEs. Die Definition schwerer UEs umfasste Handlungsunfähigkeit einerseits sowie beträchtliche Beeinträchtigungen in den Alltagsaktivitäten andererseits [25].

Im Rahmen der Studie wurden präspezifizierte Subgruppenanalysen u. a. zur Krankheitsdauer (≤ 12 oder > 12 Wochen zwischen Alter bei Symptomeintritt und Alter bei Studieneinschluss) durchgeführt [25]. Diese konnten für den Vergleich eines früh- versus einen spätsymptomatischen Therapiebeginn herangezogen werden.

Für die vorliegende Bewertung relevante Subgruppenanalyse aus der Studie ENDEAR

Die Subgruppe mit einer Krankheitsdauer ≤ 12 Wochen zwischen Symptombeginn und Studieneinschluss wurde für die Prüflintervention eines frühsymptomatischen Therapiebeginns herangezogen, die Subgruppe mit einer Krankheitsdauer > 12 Wochen für die Vergleichsintervention eines spätsymptomatischen Therapiebeginns. Beim Hersteller wurden entsprechende Daten, z. B. zu den Patientencharakteristika, sowie Ergebnisse zu patientenrelevanten Endpunkten für diese Subgruppen angefragt (siehe Abschnitt A3.1.1.3).

Die Subgruppe mit frühsymptomatischem Therapiebeginn umfasste insgesamt 52 Kinder (Nusinersen $n = 34$ versus Scheinbehandlung $n = 18$), die Subgruppe mit spätsymptomatischem Therapiebeginn 69 Kinder (Nusinersen $n = 46$ versus Scheinbehandlung $n = 23$). Während sich die Kinder hinsichtlich des Altersmedians bei Symptombeginn kaum unterschieden, lag das mittlere Alter bei der 1. Dosis in der Subgruppe frühsymptomatischer Therapiebeginn bei 16 (Spanne: 7 bis 34) Wochen in der Nusinersengruppe versus 19 (4 bis 33) Wochen in der Scheinbehandlungsgruppe. In der Subgruppe mit spätsymptomatischem Therapiebeginn lag das Alter bei 28 (18 bis 35) versus 30 (20 bis 37) Wochen.

4.3.1.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Es wurde 1 retrospektive vergleichende Studie mit Daten zum Vergleich eines prä- versus einen frühsymptomatischen Therapiebeginn mit Nusinersen herangezogen (Biogen 2019 [27]). Die Studie wurde vom Hersteller auf Anfrage erstellt (siehe Abschnitt A3.1.1.3) und zur Verfügung gestellt. In der vorliegenden Nutzenbewertung wurde dieser Vergleich im Sinne einer Vorverlegung des Therapiebeginns gedeutet.

Die Interventionsgruppe mit präsymptomatischem Therapiebeginn der Studie Biogen 2019 basiert auf der Population der Studie NURTURE. Die Vergleichsgruppe mit frühsymptomatischem Therapiebeginn setzt sich zusammen aus Kindern der oben beschriebenen Studie ENDEAR. Letzteres Kollektiv ist somit Bestandteil beider betrachteten Vergleiche zur Therapievorverlegung in diesem Bericht.

Zur Einbettung der Ergebnisse der Studie Biogen 2019 wird im Folgenden zunächst die Studie NURTURE skizziert. Daran anschließend erfolgt eine detaillierte Beschreibung der Studie Biogen 2019.

Die Studie NURTURE

Die Studie NURTURE ist eine laufende, offene, 1-armige Verlaufsbeobachtung zur präsymptomatischen Behandlung von SMA mit Nusinersen, die multizentrisch mit insgesamt 15 Zentren in 7 Ländern durchgeführt wird. Die Rekrutierung ist bereits abgeschlossen. Gemäß Studienprotokoll geplant und eingeschlossen wurden 25 Kinder mit genetischem Nachweis von 5q-assoziierter SMA, die zum Zeitpunkt des Studieneinschlusses noch keine klinischen Symptome der Erkrankung aufwiesen und ≤ 6 Wochen alt waren. 15 Kinder hatten 2 *SMN2*-Kopien und 10 Kinder 3 *SMN2*-Kopien. Es handelt sich somit überwiegend um Kinder, deren natürlicher Verlauf (d. h. Krankheitsverlauf ohne Therapie) sehr wahrscheinlich in einem frühen Krankheitsbeginn münden würde (SMA-Typ I oder II).

Die Kinder erhielten an den Studientagen 1, 15, 29 und 64 jeweils 12 mg Nusinersen als intrathekale Bolusinjektion (Aufsättigung) und im Anschluss ab Studientag 183 erfolgt alle 4 Monate jeweils 1 Erhaltungsdosis über einen Zeitraum von insgesamt 5 Jahren. Daran anschließend ist eine Follow-up-Evaluation 3 Monate nach der letzten Dosis geplant. Eine adäquate supportive Therapie ist im Studienverlauf erlaubt. Der derzeit jüngste Datenschnitt vom 29.03.2019 umfasst alle 25 Kinder und das mediane Alter betrug etwa 34 Monate.

Primärer Endpunkt der Studie ist der kombinierte Endpunkt Zeit bis Tod oder respiratorische Intervention, welche definiert ist als invasive oder nicht invasive Beatmung für ≥ 6 Stunden pro Tag durchgehend für ≥ 7 Tage oder Tracheotomie. Weitere patientenrelevante Endpunkte umfassen u. a. den kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung (und dessen Einzelkomponenten), verschiedene Operationalisierungen zu dem Endpunkt Erreichen motorischer Meilensteine, die Entwicklung einer klinisch manifesten SMA, Entwicklungs- und Wachstumsstörungen, Krankenhausaufenthalte aufgrund von respiratorischen Ereignissen, SUEs, schwere UEs und Therapieabbrüche wegen UEs. Die Definition schwerer UEs umfasste Handlungsunfähigkeit einerseits sowie beträchtliche Beeinträchtigungen in den Alltagsaktivitäten andererseits [28,29].

Die Studie Biogen 2019

Wie angefragt, lieferte der Hersteller separate Auswertungen für 2 Zielpopulationen: 1.) Kinder mit 2 *SMN2*-Kopien und 2.) Kinder mit 3 *SMN2*-Kopien. Auf diese Weise wurde eine mögliche Strukturungleichheit bezüglich dieses Merkmals vermieden. Da die Fallzahlen der Kinder mit 3 *SMN2*-Kopien jedoch zu gering waren (6 Kinder mit prä- und 5 Kinder mit frühsymptomatischem Therapiebeginn), um auf dieser Basis verlässliche Aussagen zum Nutzen (oder Schaden) treffen zu können, wurde auf eine Darstellung dieser Daten im Bericht verzichtet (siehe dafür Anhang der Stellungnahme Biogen zum Vorbericht [27]). Im Folgenden wird unter der Studie Biogen 2019 somit ausschließlich der Vergleich für die Kinder mit 2 *SMN2*-Kopien beschrieben.

In die herangezogene Studie Biogen 2019 gingen die Daten von 49 Kindern mit jeweils 2 *SMN2*-Kopien ein. 15 Kinder wurden präsymptomatisch im Rahmen der 1-armigen Studie

NURTURE mit Nusinersen behandelt. 34 Kinder wurden frühsymptomatisch im Rahmen der Studie ENDEAR mit Nusinersen behandelt (siehe oben).

Der zugrunde liegenden Fragestellung immanent, unterschieden sich die Vergleichskollektive im Hinblick auf das Vorliegen bzw. Nichtvorliegen von Symptomen bei Studienbeginn. Des Weiteren wiesen die Kollektive entsprechend den Einschlusskriterien der zugrunde liegenden Studien NURTURE und ENDEAR bei Studieneinschluss ein unterschiedliches Alter auf. Kinder mit präsymptomatischem Therapiebeginn waren bei Studieneinschluss im Mittel 14 Wochen jünger und wurden im Mittel 15 Wochen früher behandelt als Kinder mit frühsymptomatischem Therapiebeginn.

Um die Ergebnisse beider Teilkollektive angemessen miteinander vergleichen zu können, wurden die Daten nicht zu Nachbeobachtungszeitpunkten ab Studienbeginn angefragt, sondern zeitlich am Lebensalter der Kinder verankert (z. B. „Lebensalter 1 Jahr“).

Die Therapieregimes der beiden Studienarme wurden als ausreichend vergleichbar eingeschätzt.

Es wurden Daten zu den Endpunkten Gesamtüberleben, dauerhafte Beatmung, SUEs, schwere UEs, Therapieabbrüche wegen UEs, Rückenschmerzen (als SUE und schweres UE) sowie das Erreichen motorischer Meilensteine übermittelt und dargestellt.

4.3.2 Übersicht der patientenrelevanten Endpunkte

Aus 2 Studien konnten Daten zu patientenrelevanten Endpunkten extrahiert werden. Tabelle 1 zeigt die Übersicht der verfügbaren Daten zu patientenrelevanten Endpunkten aus den eingeschlossenen Studien.

Tabelle 1: Matrix der patientenrelevanten Endpunkte

Studie	Endpunkte										
	Mortalität / Morbidität		Mortalität		Morbidität					LQ	
	Zeit bis Tod oder dauerhafte Beatmung	Gesamtüberleben	Krankheitsspezifisches Überleben	Erreichen motorischer Meilensteine	Entwicklungs- und Wachstumsstörungen	Krankenhaus-aufenthalte	Atemnot	Dauerhafte Beatmung	Schwerwiegende respiratorische Ereignisse	(Schwerwiegende) UEs	Gesundheitsbezogene Lebensqualität
Vergleich früh- vs. spätsymptomatischen Therapiebeginn											
ENDEAR	● ^a	● ^b	-	● ^c	○ ^d	○ ^e	- ^f	● ^g	○ ^h	● ⁱ	-
Vergleich prä- vs. frühsymptomatischen Therapiebeginn											
Biogen 2019	- ^j	●	-	●	- ^j	- ^j	- ^j	●	- ^j	●	-
<ul style="list-style-type: none"> ● Daten wurden berichtet und waren verwertbar. ○ Daten wurden berichtet, aber waren nicht für die Nutzenbewertung verwertbar. - Es wurden keine Daten berichtet (keine weiteren Angaben). / Der Endpunkt wurde nicht erhoben. <p>a: kombinierter Endpunkt, bestehend aus den Einzelkomponenten Zeit bis Tod und dauerhafte Beatmung, die definiert war als Beatmung über ≥ 16 Stunden pro Tag durchgehend für > 21 Tage bei Abwesenheit von akuten reversiblen Ereignissen oder Tracheotomie</p> <p>b: Dieser Endpunkt wird als Teilkomponente des kombinierten Endpunkts Zeit bis Tod oder dauerhafte Beatmung berücksichtigt und dargestellt. Die Daten für die berichtsrelevanten Subgruppen zur Krankheitsdauer wurden (teilweise) nicht öffentlichen Unterlagen entnommen.</p> <p>c: Die Daten für die berichtsrelevanten Subgruppen zur Krankheitsdauer wurden (teilweise) nicht öffentlichen Unterlagen entnommen. Für die berichtsrelevanten Subgruppen lagen Ergebnisse zu den Instrumenten CHOP und HINE-2 in Form von Responderanalysen sowie der Veränderung der Scores im Vergleich zu Baseline und Ergebnisse des Interaktionstests vor. Ergebnisse zum HINE-2 in Form von Responderanalysen wurden berücksichtigt und dargestellt.</p> <p>d: Berichtet wurde der Endpunkt in Form der Operationalisierungen HINE-1 und HINE-3. Es waren keine verwertbaren Daten für die berichtsrelevanten Subgruppen zur Krankheitsdauer verfügbar.</p> <p>e: Für die berichtsrelevanten Subgruppen zur Krankheitsdauer lagen lediglich p-Werte zum Interaktionstest vor. Die Ergebnisse waren auch aufgrund der fehlenden patientenbasierten Auswertung nicht interpretierbar und wurden daher nicht herangezogen.</p> <p>f: Atemnot wurde im Rahmen der Studie als UE (Preferred Term) erfasst. Für die berichtsrelevanten Subgruppen zur Krankheitsdauer lagen keine Daten vor.</p> <p>g: Dieser Endpunkt wird als Teilkomponente des kombinierten Endpunkts Zeit bis Tod oder dauerhafte Beatmung berücksichtigt und dargestellt. Die Daten für die berichtsrelevanten Subgruppen zur Krankheitsdauer wurden (teilweise) nicht öffentlichen Unterlagen entnommen.</p> <p>h: Der Endpunkt umfasst alle UEs, die während des Studienzeitraums in die Systemorganklasse (SOC) der Erkrankungen der Atemwege, des Brustraums und Mediastinums als primäre SOC oder sekundäre SOC und als schwerwiegend eingestuft wurden. Seltene Ereignisse in den SOC's Infektionen und parasitäre Erkrankungen und Untersuchungen wurden (falls vorhanden) nicht berücksichtigt. Für die berichtsrelevanten Subgruppen zur Krankheitsdauer lagen lediglich p-Werte zum Interaktionstest vor. Die Ergebnisse waren auch aufgrund der fehlenden patientenbasierten Auswertung nicht interpretierbar und wurden daher nicht herangezogen.</p> <p>i: Hierunter fallen SUEs, schwere UEs sowie Abbrüche wegen UEs. Die Daten für die berichtsrelevanten Subgruppen zur Krankheitsdauer wurden (teilweise) nicht öffentlichen Unterlagen entnommen.</p> <p>j: Ergebnisse zum Endpunkt nicht angefragt</p> <p>CHOP: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; HINE: Hammersmith Infant Neurological Examination; HINE-1: HINE – Subscale 1; HINE-2: HINE – Subscale 2; HINE-3: HINE – Subscale 3; LQ: gesundheitsbezogene Lebensqualität; SOC: Systemorganklasse; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis</p>											

4.3.3 Bewertung des Verzerrungspotenzials der Ergebnisse

4.3.3.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Das Verzerrungspotenzial wurde endpunktübergreifend als niedrig eingestuft.

Das endpunktspezifische Verzerrungspotenzial für die Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung, zum Endpunkt Erreichen motorischer Meilensteine sowie zu den Endpunkten SUEs, schwere UEs und Therapieabbrüche wegen UEs wurde als hoch bewertet. Dies lag daran, dass keine Angaben zur Verteilung der Begleitbehandlung in den Subgruppen vorlagen. Da in der Studie jedoch eine Begleitbehandlung nach BSC erfolgte und verfügbare Angaben zur Gesamtpopulation auch vielfältige Begleitbehandlungen dokumentieren, wären zur Einschätzung eines möglichen Kointerventionsbias entsprechende Angaben für die Subgruppen erforderlich gewesen. Damit wurde die qualitative Ergebnissicherheit für alle Endpunkte als mäßig bewertet.

4.3.3.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Aufgrund des retrospektiven Studiendesigns ist das Verzerrungspotenzial endpunktübergreifend als hoch einzustufen. Auch das endpunktspezifische Verzerrungspotenzial wurde daher für alle berichteten Endpunkte als hoch und die qualitative Ergebnissicherheit für alle Endpunkte damit als sehr gering bewertet.

4.3.4 Ergebnisse zu patientenrelevanten Endpunkten

4.3.4.1 Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung

4.3.4.1.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Für den kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung lagen aus der Studie ENDEAR Ergebnisse zu den berichtsrelevanten Subgruppen vor. Der Interaktionstest zeigte eine statistisch signifikante Interaktion zwischen den Subgruppen früh- versus spätsymptomatischen Therapiebeginn (Krankheitsdauer ≤ 12 / > 12 Wochen) bei Kindern mit infantiler SMA ($p = 0,008$). Der Effekt zugunsten von Nusinersen beim frühsymptomatischen Therapiebeginn (HR: 0,16; 95 %-KI: [0,06; 0,44]; $p < 0,001$) war dabei statistisch signifikant im Gegensatz zum nicht statistisch signifikanten Unterschied beim spätsymptomatischen Therapiebeginn (HR: 0,82; 95 %-KI: [0,43; 1,54]; $p = 0,533$).

Für die Einzelkomponenten zeigte sich die gleiche Effektrichtung, wobei für die Einzelkomponente Zeit bis dauerhafte Beatmung ebenfalls eine statistisch signifikante Interaktion vorlag ($p = 0,040$). Das hier von der Studiengruppe verwendete Modell zur Berechnung des HR unterscheidet sich allerdings zwischen dieser Einzelkomponente und dem kombinierten Endpunkt (HR basierend auf Cox-Regression bzw. a priori geplante Cox-Regression adjustiert für die Krankheitsdauer zum Rekrutierungszeitpunkt). Für die Einzelkomponente Zeit bis Tod war die Interaktion nicht statistisch signifikant, die Effektschätzer in den Subgruppen zeigen ein mit dem der Gesamtgruppe vergleichbares Bild.

Insgesamt ergibt sich somit für den kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung ein Anhaltspunkt für unterschiedliche Effekte in Bezug auf den Therapiebeginn bei symptomatischen Kindern mit infantiler SMA, wobei der Effekt bei einem frühsymptomatischen Therapiebeginn statistisch signifikant war zugunsten von Nusinersen.

4.3.4.1.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung wurden beim Hersteller nicht angefragt. Ergebnisse zu diesem Endpunkt lagen für die Studie Biogen 2019 folglich nicht vor.

Aufgrund der geringen Ereigniszahlen bei den Einzelkomponenten (siehe Abschnitte 4.3.4.2.2 und 4.3.4.8.2) ist für den kombinierten Endpunkt nicht mit einem dramatischen Effekt zu rechnen, welcher wegen der sehr geringen qualitativen Ergebnissicherheit der Studie für eine Nutzensaussage erforderlich wäre.

4.3.4.2 Ergebnisse zum Gesamtüberleben

4.3.4.2.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Für den Endpunkt Gesamtüberleben lagen verwertbare Ergebnisse aus der Studie ENDEAR für die berichtsrelevanten Subgruppen vor. Dieser Endpunkt wird als Teilkomponente des kombinierten Endpunkts Zeit bis Tod oder dauerhafte Beatmung berücksichtigt und dargestellt (siehe Abschnitte 4.3.4.1).

4.3.4.2.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Für den Endpunkt Gesamtüberleben lagen aus der Studie Biogen 2019 Daten zur Anzahl von Todesfällen vor.

Im Lebensalter von 1 Jahr leben alle 15 Kinder der Interventionsgruppe mit präsymptomatischem Therapiebeginn. In der Vergleichsgruppe mit frühsymptomatischem Therapiebeginn sind bis zu diesem Zeitpunkt 3 Kinder (9 %) verstorben.

Dieser numerische Unterschied im Gesamtüberleben war jedoch nicht statistisch signifikant. Die Kriterien eines dramatischen Effekts waren nicht erfüllt (siehe Allgemeine Methoden 5.0 [30]). Auf Basis dessen ergibt sich für den Endpunkt Gesamtüberleben kein Anhaltspunkt für einen höheren Nutzen oder Schaden eines präsymptomatischen gegenüber einem frühsymptomatischen Therapiebeginn.

4.3.4.3 Ergebnisse zum krankheitsspezifischen Überleben

4.3.4.3.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Zum Endpunkt krankheitsspezifisches Überleben lagen keine Ergebnisse für die berichtsrelevanten Subgruppen zur Krankheitsdauer vor.

4.3.4.3.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Zum Endpunkt krankheitsspezifisches Überleben erfolgte beim Hersteller keine Anfrage, da aus den zugrunde liegenden Studien ENDEAR und NURTURE keine entsprechenden Ergebnisse zu erwarten waren.

4.3.4.4 Ergebnisse zum Erreichen motorischer Meilensteine

4.3.4.4.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Zu dem Endpunkt Erreichen motorischer Meilensteine lagen aus der Studie ENDEAR für die berichtsrelevanten Subgruppen Ergebnisse zu den Instrumenten HINE-2 (als primärem Endpunkt der Studie) und CHOP (Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders) INTEND in unterschiedlichen Operationalisierungen vor. Die Instrumente wurden bislang beide nicht abschließend validiert. Darüber hinaus gibt es keine validierte Minimal important Difference (MID) für die Instrumente. Dennoch wird der HINE-2 in der vorliegenden Nutzenbewertung berücksichtigt. Angesichts der Seltenheit der Erkrankung und der Verfügbarkeit von Ergebnissen der einzelnen Items, die jeweils alleinstehend als patientenrelevant erachtet werden und zur Interpretation der Responder des Gesamtscores beitragen können, wird der HINE-2 in Form von Responderanalysen für den Endpunkt Erreichen motorischer Meilensteine herangezogen. Von einer Darstellung des CHOP INTEND wird darüber hinaus abgesehen, da es sich um eine andere Operationalisierung desselben Endpunkts handelt. Die Ergebnisse weisen in dieselbe Richtung wie die des HINE-2.

Ergebnisse zum HINE-2 in Form von Responderanalysen zum Gesamtscore wurden für die berichtsrelevanten Subgruppen auf Basis des ITT-Sets herangezogen. Des Weiteren wurden Ergebnisse zu Responderanalysen für die erfassten einzelnen Items des HINE-2 dargestellt. Diese lagen allerdings nur auf Basis des Efficacy Sets vor (Patientinnen und Patienten mit einer Nachbeobachtung bis mindestens Studientag 183). Die Definition der Gesamtscore-Responder basierte auf 7 von 8 einzelnen Meilensteinkategorien des HINE-2, die jeweils anhand von Skalen von 3 bis 5 möglichen Entwicklungsstufen gemessen werden. Das bewusste Greifen wurde von der Studiengruppe aus den Responderanalysen ausgeschlossen, da dieser motorische Meilenstein auch von Kindern mit SMA ohne Behandlung zu großen Teilen erreicht werden kann. Als Gesamtscore-Responder wurden Kinder gewertet, die die folgenden Kriterien erfüllten: (1) ≥ 1 -Punkt-Verbesserung in den Kategorien Kopfkontrolle, Drehen, Sitzen, Robben / Krabbeln, Stehen oder Laufen oder ≥ 2 -Punkt-Verbesserung in der Kategorie Strampeln und / oder Erreichen des Maximalscores in der Kategorie Strampeln, (2) mehr Kategorien mit einer Verbesserung als Kategorien mit einer Verschlechterung. Verstorbene Kinder sowie Studienabbrecher wurden dabei als Non-Responder gewertet. Responder bei den einzelnen Items waren analog zu den Einzelteilen beim Gesamtscore definiert.

Für die HINE-2-Responder zum Gesamtscore zeigte sich eine statistisch signifikante Interaktion zwischen den Subgruppen früh- versus spätsymptomatischen Therapiebeginn ($p = 0,003$), wobei Kinder mehr von einem frühsymptomatischen als von einem spätsymptomatischen Therapiebeginn profitieren. Innerhalb der Subgruppen zeigte sich jeweils ein

statistisch signifikanter Effekt zugunsten von Nusinersen (frühsymptomatischer Therapiebeginn: RD: 0,71; 95 %-KI: [0,55; 0,86]; $p < 0,001$; spätsymptomatischer Therapiebeginn: RD: 0,37; 95 %-KI: [0,23; 0,51]; $p < 0,00$). In der Gruppe der Kinder mit einer Scheinbehandlung gab es jeweils keine Responder.

Anhand der Ergebnisse zu den einzelnen Items des HINE-2 ist erkennbar, dass sich mit Ausnahme 1 Kindes mit Scheinbehandlung ausschließlich bei Kindern mit einer Nusinersenbehandlung Verbesserungen in Form von Respondern bei einzelnen motorischen Meilensteinen zeigten. Des Weiteren zeigten sich numerisch über alle Items hinweg mehr Responder in der Subgruppe mit frühsymptomatischem Therapiebeginn als in der Subgruppe mit spätsymptomatischem Therapiebeginn. Anhand zusätzlicher Daten im Studienbericht ist erkennbar, dass sich diese Unterschiede nicht durch ungleiche Voraussetzungen in Bezug auf den Entwicklungsstand zu Studienbeginn erklären lassen; denn zu Studienbeginn hatte der Entwicklungsprozess bei den verschiedenen motorischen Meilensteinen im Gesamtkollektiv mit wenigen Ausnahmen noch nicht eingesetzt. Insgesamt stützen diese Daten das statistisch signifikante Ergebnis des Interaktionstests der Responderanalyse zum Gesamtscore.

Auf Basis dieser Ergebnisse wird für den Endpunkt Erreichen motorischer Meilensteine ein Anhaltspunkt für unterschiedliche Effekte in Bezug auf den Therapiebeginn bei symptomatischen Kindern mit infantiler SMA abgeleitet, wobei der Effekt bei einem frühsymptomatischen Therapiebeginn größer ist als bei einem spätsymptomatischen Therapiebeginn.

4.3.4.4.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Für den Endpunkt Erreichen motorischer Meilensteine lagen aus der Studie Biogen 2019 verwertbare Daten zu dem Instrument HINE-2 wie folgt vor: 1.) in Form einer Veränderung des Gesamtscores vom Zeitpunkt des Studienbeginns zum Lebensalter von 1 Jahr und 2.) in Form von Abbildungen zum Erreichen des höchsten Levels der 4 Einzelitems „Sitzen: dreht sich im Sitzen“, „Krabbeln: krabbelt auf Händen und Knien“, „Stehen: steht ohne Hilfe“ und „Gehen: geht selbstständig“. Die Änderung des HINE-2-Gesamtscores basierte auf allen 8 Meilensteinkategorien des HINE-2 (und unterscheidet sich dahin gehend von den im vorausgehenden Abschnitt dargestellten Responderanalysen zum HINE-2, in denen nur 7 der 8 Einzelitems berücksichtigt wurden).

In der Interventionsgruppe mit präsymptomatischem Therapiebeginn lag der Wert des HINE-2-Gesamtscores zu Studienbeginn durchschnittlich bei 2,67 Punkten (Standardabweichung [SD] 1,59) und in der Vergleichsgruppe mit frühsymptomatischem Therapiebeginn bei 1,15 Punkten (SD 1,10). Bereits zu diesem Zeitpunkt unterschieden sich die Werte beider Gruppen statistisch signifikant (Mittelwertdifferenz [MD]: 1,52; 95 %-KI: [0,58; 2,46]; $p = 0,003$). Einerseits wäre zu erwarten gewesen, dass die Kinder mit frühsymptomatischem Therapiebeginn zu Studienbeginn höhere Werte hätten aufweisen müssen, da sie zu diesem Zeitpunkt ca. 3 Monate älter waren als die Kinder mit präsymptomatischem Therapiebeginn. Andererseits lässt sich der niedrigere Wert bei diesen Kindern mit der bereits vorliegenden Symptomatik erklären, die möglicherweise schon zu einem Rückschritt der motorischen Entwicklung

aufgrund des bereits eingesetzten Untergangs beziehungsweise einer Nichtentwicklung der motorischen Nervenzellen geführt haben könnte. Im Lebensalter von 1 Jahr ist der Mittelwert der Interventionsgruppe mit präsymptomatischem Therapiebeginn auf 19,67 Punkte gestiegen (SD 4,15). In der Vergleichsgruppe mit frühsymptomatischem Therapiebeginn hat ein Anstieg auf einen Mittelwert von 5,86 Punkten (SD 3,16) stattgefunden. Der Unterschied zwischen beiden Gruppen im Lebensalter von 1 Jahr fällt damit größer aus (MD: 13,81; 95 %-KI: [11,01; 16,61]; $p < 0,001$) als der Unterschied zwischen beiden Gruppen zu Studienbeginn.

Die Veränderungen im HINE-2-Gesamtscore vom Zeitpunkt des Studienbeginns zum Lebensalter von 1 Jahr unterscheiden sich sehr deutlich zwischen den Gruppen zugunsten des präsymptomatischen im Vergleich zum frühsymptomatischen Therapiebeginn (Hedges' g : 3,62; 95 %-KI: [2,38; 4,86]). Die untere Konfidenzgrenze ist mehr als 10-mal so groß wie die übliche Schwelle von 0,2, die im Rahmen der klinischen Relevanzbewertung verwendet wird (siehe Allgemeine Methoden 5.0 [30]). Dieser Unterschied wird als groß genug eingeschätzt, als dass er nicht allein durch die Einwirkung von Störgrößen zu erklären ist.

Anhand der grafisch dargestellten Ergebnisse zum Erreichen des höchsten Levels der Einzelitems ist erkennbar, dass in der Vergleichsgruppe mit frühsymptomatischem Therapiebeginn mit Ausnahme 1 Kindes, das mit 15 Monaten das höchste Level beim Einzelitem „Sitzen“ erreicht, keines der Kinder 1 oder mehrere dieser 4 Meilensteine erreicht. In der Gruppe mit präsymptomatischem Therapiebeginn hingegen erreichen je nach Einzelitem nahezu alle (Sitzen), die Mehrheit (Krabbeln und Stehen) beziehungsweise einige (Gehen) das höchste Level der Meilensteine. Die unterschiedlich lange Behandlungs- und Beobachtungsdauer und dadurch ein unterschiedliches Alter der Gruppen beim letzten Follow-up erklären einen Teil des Unterschieds. Es bleibt aber ein deutlicher Unterschied bestehen, wenn nur diejenigen Kinder betrachtet werden, für die der Responsewert bis zum Lebensalter von etwa 13 Monaten vorliegt. Bei diesem Zeitpunkt handelt es sich um das im Mittel letzte Follow-up der Vergleichsgruppe mit frühsymptomatischem Therapiebeginn. Während zu diesem Zeitpunkt kein Kind der Gruppe mit frühsymptomatischem Therapiebeginn einen Meilenstein erreicht, kann sich ungefähr die Hälfte der Kinder mit präsymptomatischem Therapiebeginn zu diesem Zeitpunkt im Sitzen drehen, ein Drittel der Kinder krabbelt auf Händen und Füßen und 1 Kind kann bereits ohne Hilfe stehen.

Insgesamt stützen diese Daten das Ergebnis zu den Veränderungen im HINE-2-Gesamtscore. In der Gesamtschau zeigt sich ein dramatischer Effekt, auf Basis dessen für den Endpunkt Erreichen motorischer Meilensteine ein Anhaltspunkt für einen höheren Nutzen eines präsymptomatischen gegenüber einem frühsymptomatischen Therapiebeginn abgeleitet wird.

4.3.4.5 Ergebnisse zu Entwicklungs- und Wachstumsstörungen

4.3.4.5.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Neben dem Endpunkt Erreichen motorischer Meilensteine wurden in der Studie ENDEAR weitere Entwicklungs- und Wachstumsstörungen mittels HINE (Subskalen 1 und 3) erhoben.

Für die berichtsrelevanten Subgruppen zur Krankheitsdauer lagen keine verwertbaren Daten vor.

4.3.4.5.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Ergebnisse zum Endpunkt Entwicklungs- und Wachstumsstörungen wurden beim Hersteller nicht angefragt. Ergebnisse zu diesem Endpunkt lagen folglich nicht vor.

4.3.4.6 Ergebnisse zu Krankenhausaufenthalten

4.3.4.6.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Für den Endpunkt Krankenhausaufenthalte lagen für die berichtsrelevanten Subgruppen zur Krankheitsdauer keine verwertbaren Daten aus der Studie ENDEAR vor (siehe Details in Abschnitt A3.3.2.6).

4.3.4.6.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Ergebnisse zum Endpunkt Krankenhausaufenthalte wurden beim Hersteller nicht angefragt. Ergebnisse zu diesem Endpunkt lagen folglich nicht vor.

4.3.4.7 Ergebnisse zu Atemnot

4.3.4.7.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Der Endpunkt Atemnot wurde im Rahmen der Studie ENDEAR in Form von UEs erfasst. Für die berichtsrelevanten Subgruppen zur Krankheitsdauer lagen keine Daten vor.

4.3.4.7.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Der Endpunkt Atemnot wurde im Rahmen der Studien ENDEAR und NURTURE, die Biogen 2019 zugrunde liegen, in Form von UEs erfasst. Er wurde beim Hersteller nicht gesondert angefragt. Ergebnisse zu diesem Endpunkt lagen folglich nicht vor.

4.3.4.8 Ergebnisse zur dauerhaften Beatmung

4.3.4.8.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Der Endpunkt Beatmung wurde in der Studie ENDEAR in unterschiedlichen Operationalisierungen erfasst. Im Rahmen der vorliegenden Bewertung wurden Ergebnisse zur dauerhaften Beatmung berücksichtigt, die als Teilkomponente des kombinierten Endpunkts Zeit bis Tod oder dauerhafte Beatmung herangezogen wurden (vgl. Abschnitte 4.3.4.1).

4.3.4.8.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Für den Endpunkt dauerhafte Beatmung lagen aus der Studie Biogen 2019 verwertbare Daten zur Anzahl von Kindern mit Ereignissen vor.

Im Lebensalter von 1 Jahr benötigte kein Kind (0 %) der Interventionsgruppe mit präsymptomatischem Therapiebeginn eine dauerhafte Beatmung. In der Vergleichsgruppe mit

frühsymptomatischem Therapiebeginn wurden zu diesem Zeitpunkt 2 Kinder (6 %) dauerhaft beatmet.

Dieser Unterschied in der dauerhaften Beatmung war nicht statistisch signifikant. Die Kriterien eines dramatischen Effekts waren nicht erfüllt (siehe Allgemeine Methoden 5.0 [30]). Auf Basis dessen ergibt sich für den Endpunkt dauerhafte Beatmung kein Anhaltspunkt für einen höheren Nutzen oder Schaden eines präsymptomatischen gegenüber einem frühsymptomatischen Therapiebeginn.

4.3.4.9 Ergebnisse zu schwerwiegenden respiratorischen Ereignissen

4.3.4.9.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Der Endpunkt schwerwiegende respiratorische Ereignisse wurde in der Studie ENDEAR erfasst, jedoch lagen für die berichtsrelevanten Subgruppen zur Krankheitsdauer keine verwertbaren Daten vor (siehe Abschnitt A3.3.2.9).

4.3.4.9.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Ergebnisse zum Endpunkt schwerwiegende respiratorische Ereignisse wurden beim Hersteller nicht angefragt. Ergebnisse zu SUEs werden davon ungeachtet im folgenden Abschnitt dargestellt.

4.3.4.10 Ergebnisse zu (schwerwiegenden) unerwünschten Ereignissen

4.3.4.10.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Für die Endpunkte SUEs, schwere UEs und Therapieabbrüche wegen UEs wurden zusätzliche Daten aus der Studie ENDEAR vom Hersteller für die berichtsrelevanten Subgruppen zur Krankheitsdauer übermittelt. Der Interaktionstest zeigte bei keinem der 3 Endpunkte eine statistisch signifikante Interaktion zwischen den Subgruppen früh- versus spätsymptomatischen Therapiebeginn (Krankheitsdauer ≤ 12 / > 12 Wochen) bei Kindern mit infantiler SMA.

In beiden Subgruppen war die Anzahl der Kinder mit Ereignissen in der Nusinersengruppe jeweils geringer als in der Gruppe der Scheinbehandlung, wobei der Effekt in der Subgruppe mit frühsymptomatischem Therapiebeginn für die Endpunkte schwere UEs und Therapieabbrüche wegen UEs statistisch signifikant war (siehe auch Kapitel 5).

Daten zu einzelnen (S)UEs lagen für die berichtsrelevanten Subgruppen nicht vor.

Insgesamt ergibt sich für die Endpunkte SUEs, schwere UEs und Therapieabbrüche wegen UEs kein Anhaltspunkt für unterschiedliche Effekte in Bezug auf den Therapiebeginn bei symptomatischen Kindern mit infantiler SMA.

4.3.4.10.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Für die Endpunkte SUEs, schwere UEs, Therapieabbrüche wegen UEs und Rückenschmerzen (als SUE und schweres UE) lagen verwertbare Ergebnisse für die Studie Biogen 2019 vor.

Nummerisch war die Anzahl der Kinder mit Ereignissen in der Vergleichsgruppe mit fröhsymptomatischem Therapiebeginn bei allen 4 Endpunkten höher als in der Interventionsgruppe mit präsymptomatischem Therapiebeginn. Bei SUEs und schweren UEs waren die Unterschiede statistisch signifikant zugunsten der Gruppe mit präsymptomatischem Therapiebeginn. Diese beobachteten Unterschiede werden als nicht groß genug eingeschätzt, als dass sie jeweils nicht allein durch die Einwirkung von Störgrößen zu erklären sein könnten. Auf Basis dessen ergibt sich für die Endpunkte SUEs, schwere UEs, Therapieabbrüche wegen UEs und Rückenschmerzen (als SUE und schweres UE) kein Anhaltspunkt für einen höheren oder geringeren Schaden eines präsymptomatischen gegenüber einem fröhsymptomatischen Therapiebeginn.

4.3.4.11 Ergebnisse zur gesundheitsbezogenen Lebensqualität des Kindes

4.3.4.11.1 Fröhsymptomatischer versus spätsymptomatischen Therapiebeginn

Der Endpunkt gesundheitsbezogene Lebensqualität des Kindes wurde in der eingeschlossenen Studie nicht erhoben.

4.3.4.11.2 Präsymptomatischer versus fröhsymptomatischen Therapiebeginn

Zum Endpunkt gesundheitsbezogene Lebensqualität des Kindes erfolgte beim Hersteller keine Anfrage, da dieser Endpunkt in der Studie ENDEAR, die Biogen 2019 unter anderem zugrunde liegt, nicht erhoben wurde und somit keine Daten für das Teilkollektiv mit fröhsymptomatischem Therapiebeginn verfügbar waren. Ergebnisse zu diesem Endpunkt lagen folglich nicht vor.

4.4 Studien zur diagnostischen Güte

4.4.1 Charakteristika der in die Bewertung eingeschlossenen Studien

Es wurden 4 Studien zur diagnostischen Güte eingeschlossen. Bei allen handelt es sich um prospektive diagnostische Kohortenstudien im Verification-of-only-positive-Testers(VOPT)-Design, in denen für testnegative Befunde keine systematische Nachbeobachtung bzw. kein Referenztest beschrieben werden.

In der prospektiven diagnostischen Kohortenstudie Chien 2017 [14] werden die Ergebnisse eines Neugeborenencreening-Programms in Taiwan beschrieben. Im Zeitraum November 2014 bis September 2016 wurde in 1 Screeningzentrum Filterkartenblut von 120 267 Neugeborenen in einer konsekutiven Serie untersucht. Mittels Echtzeit-Polymerase-Kettenreaktion (RT-PCR) sollte das Fehlen von *SMN1* nachgewiesen werden. Dafür wurde ein Grenzwert von *SMN1* $\Delta Rn < 1$ für testpositive Befunde festgesetzt. Bei unzureichender Desoxyribonukleinsäure(DNA)-Qualität wurde der Test wiederholt. Bei einem positiven Testergebnis wurde dasselbe Probenmaterial mit einer nachgeschalteten digitalen Tröpfchen-Polymerase-Kettenreaktion (ddPCR) untersucht zwecks Absicherung und Bestimmung der Anzahl der *SMN2*-Kopien. Auf Basis dieses 2-stufigen Screeningtests wurden abschließend positive Befunde

ermittelt. Abschließend testpositive Befunde wurden dann mittels multiplexer ligations-abhängiger Sondenamplifikation (MLPA) aus frischem Vollblut im Hinblick auf eine homozygote *SMN1*-Exon-7-Deletion untersucht und die jeweilige Anzahl von *SMN2*-Kopien bestimmt.

Bei der Studie Czibere/Vill 2019 [31,32] handelt es sich um ein laufendes Neugeborenencreening-Programm in Deutschland. Für den vorliegenden Bericht wird diese Studie als Studie zur diagnostischen Güte herangezogen. Im Rahmen dieses Programms wird seit Januar 2018 in den Bundesländern Bayern und Nordrhein-Westfalen Filterkartenblut von Neugeborenen im Hinblick auf einen Nachweis einer homozygoten Deletion des Exons 7 von *SMN1* untersucht. Gemäß den Autorinnen und Autoren werden circa 80 % aller Neugeborenen in Bayern und circa 40 % aller Neugeborenen in Nordrhein-Westfalen mit diesem Programm erfasst [32]. Bis zum Zeitpunkt der Publikation Czibere 2019 waren dies 213 279 Neugeborene und bis zum Zeitpunkt der Publikation Vill 2019 241 270. Da es sich den Autorinnen und Autoren zufolge bei Letztgenannten um vorläufige Ergebnisse handelt und der Publikation zudem keine Angaben zur genetischen Überprüfung dieser vorläufigen Ergebnisse zu entnehmen sind, werden im vorliegenden Bericht die Daten zu den 213 279 Neugeborenen mit verifizierten Screeningtestergebnissen der Testpositiven aus Czibere 2019 herangezogen. Die Bestimmung der homozygoten *SMN1* Exon 7 Deletion erfolgte mittels quantitativer Polymerase-Kettenreaktion (qPCR). Die *SMN1*-Genkopie wurde als vorhanden angesehen, wenn ein Produkt vor dem PCR-Zyklus 36 (Quantifizierungszyklus-Wert, Cq-Wert < 36) nachgewiesen wurde, während für die Amplifikation der internen Kontrolle ein Cq-Wert < 34 festgelegt wurde. Wenn nur bei der Kontrolle ein Produkt nachgewiesen wurde, wurde das Testergebnis als positiv angesehen, d. h., beim Neugeborenen lag wahrscheinlich eine homozygote Deletion des *SMN1* Exons 7 vor. In diesem Fall sowie bei invalidem Testergebnis (d. h. bei fehlendem Produkt sowohl für die Probe als auch die Kontrolle) wurde ein neues Stück Filterkartenblut ausgestanzt und erneut getestet. Abschließend testpositive Befunde wurden dann mittels MLPA aus frischem Vollblut in einem unabhängigen Labor im Hinblick auf homozygote *SMN1* Exon 7 Deletion untersucht und die jeweilige Anzahl an *SMN2*-Kopien mit doppelter Testung bestimmt.

In der Publikation zu der prospektiven diagnostischen Kohortenstudie Kariyawasam 2019 [33] werden Ergebnisse aus einem Neugeborenencreening-Programm in Australien berichtet. Im Zeitraum August 2018 bis Juli 2019 wurde dieses flächendeckend im Bundesstaat New South Wales sowie im Australien Capital Territory durchgeführt. Insgesamt wurde das Filterkartenblut von 103 903 Neugeborenen untersucht. Mittels RT-PCR sollte das Fehlen von *SMN1* nachgewiesen werden, wobei der Publikation keine Angaben zum verwendeten, vorab validierten Grenzwert zu entnehmen sind. Bei einem positiven Testergebnis wurde ein neues Stück Filterkartenblut ausgestanzt und mit einer nachgeschalteten ddPCR zwecks Absicherung untersucht. Zudem wurde in diesem Schritt anhand der 1. und 2. Probe die Anzahl der *SMN2*-Kopien bestimmt. Als testpositiv wurden ausschließlich Neugeborene mit < 4 *SMN2*-Kopien definiert. Individuen mit ≥ 4 *SMN2*-Kopien wurden als testnegativ beschrieben und wurden

nicht gemeldet. In dieser Studie liegt somit eine enger gefasste Zielerkrankungsdefinition vor als in den 3 anderen eingeschlossenen Studien zur diagnostischen Güte. Auf Basis des beschriebenen 2-stufigen Screeningtests wurden abschließend positive Befunde ermittelt, welche daran anschließend mittels genetischen Nachweises aus frischem Vollblut untersucht wurden.

In der prospektiven diagnostischen Kohortenstudie Kraszewski 2018 [34] werden Ergebnisse aus einem Neugeborenencreening-Programm im New York State berichtet. Dazu wurde Filterkartenblut von 3826 Neugeborenen im Zeitraum Januar 2016 bis Januar 2017 im Hinblick auf einen Nachweis einer homozygoten Deletion des Exons 7 von *SMN1* untersucht. Mittels qPCR wurde die relative Durchschnittsmenge der *SMN1*-Exon-7-Kopien (Average relative Quantity, aRQ) bestimmt. Neugeborene, deren Proben eine $aRQ \geq 0,8$ aufwiesen, wurden als testnegativ eingestuft. Bei Proben mit einer $aRQ < 0,8$ und Proben, die den Qualitätsanforderungen nicht genügten, wurde ein neues Stück Filterkartenblut ausgestanzt und erneut getestet. Abschließend testpositive Befunde ($RQ = 0$) wurden klinisch untersucht und das Screeningergebnis sowie das Vorhandensein von 2 *SMN2*-Kopien in einem externen Labor bestätigt.

4.4.2 Vorhandene bewertungsrelevante Zielgrößen

Für die Beurteilung geeigneter diagnostischer Testverfahren wurden 4 Studien zur diagnostischen Güte im VOPT-Design betrachtet (Chien 2017, Czibere/Vill 2019, Kariyawasam 2019, Kraszewski 2018). Dieses Studiendesign erlaubt allein eine Berechnung des PPV. Nur in dem Fall, dass keine falsch-positiven Ereignisse auftreten, sind darüber hinaus Aussagen zur Spezifität möglich.

4.4.3 Bewertung des Verzerrungspotenzials und der Übertragbarkeit

Das Verzerrungspotenzial der Studien Chien 2017, Czibere/Vill 2019 und Kariyawasam 2019 zur diagnostischen Güte wurde jeweils als niedrig eingestuft. Ebenso wurden die Bedenken bezüglich der Übertragbarkeit der Ergebnisse dieser Studien als gering bewertet.

Die Studie Kraszewski 2018 wurde als mit einem hohen Verzerrungspotenzial behaftet bewertet, da sowohl bei der Patientenselektion als auch beim Indextest Unklarheiten bestehen. Zum einen erfolgte der Studieneinschluss von Kindern Englisch oder Spanisch sprechender Mütter ausschließlich wochentags. Daher bleibt unklar, ob die Neugeborenen konsekutiv eingeschlossen wurden und ob inadäquate Studienausschlüsse vermieden wurden. Zum anderen wurden Teile des Indextests manuell durchgeführt. Diese manuelle Durchführung führte ebenfalls zu einer Einschätzung als unklar im Hinblick auf die Übertragbarkeit des Indextests. In einer zusammenfassenden Einschätzung wurden die Bedenken bezüglich der Übertragbarkeit der Ergebnisse der Studie Kraszewski 2018 dennoch als gering bewertet.

4.4.4 Ergebnisse zu den Zielgrößen

In der Studie Chien 2017 wurde zur Auswertung der Filterkartenblutproben ein 2-stufiges Verfahren angewendet. In einem 1. Schritt wurden mittels RT-PCR 15 Neugeborene positiv

getestet. In einem 2. Schritt wurde dasselbe Probenmaterial dieser 15 Neugeborenen mit einer nachgeschalteten ddPCR untersucht. Während diese bei 8 Neugeborenen zu einem abschließend negativen Befund führte, blieb es bei 7 Neugeborenen bei einem positiven Testergebnis. Mittels dieses 2-stufigen Verfahrens traten keine falsch-positiven Ergebnisse auf, das heißt, sowohl PPV als auch Spezifität liegen bei 100 % (PPV 100; 95 %-KI: [64,6; 100]). In der laufenden Studie Czibere/Vill 2019, in der die Auswertung der Filterkartenblutproben mittels qPCR erfolgte, wurden zu einem Datenschnitt nach eineinhalb Jahren ebenfalls keine falsch-positiven Befunde berichtet, das heißt, sowohl PPV als auch Spezifität liegen bei 100 % (PPV 100; 95 %-KI: [88,6; 100]). Insgesamt wurden in dieser Studie bislang 30 Neugeborene mit 5q-assoziierter SMA identifiziert. In der Studie Kraszewski 2018, in der die Auswertung ebenfalls mittels qPCR erfolgte, wurden keine falsch-positiven Befunde berichtet. Das heißt, auch in dieser Studie liegen sowohl PPV als auch Spezifität bei 100 % (PPV 100; 95 %-KI: [20,7; 100] bei 1 positiv getesteten Neugeborenen).

In der Studie Kariyawasam 2019 wurde eine enger gefasste Zielerkrankungsdefinition zugrunde gelegt als in den anderen Studien, da mit dem Indextest nur Neugeborene mit < 4 *SMN2*-Kopien als testpositiv klassifiziert wurden. Zum Einsatz kam ein sehr ähnliches 2-stufiges Verfahren wie in der Studie Chien 2017. 10 Neugeborene wurden sowohl im 1. Schritt mittels RT-PCR als auch im 2. Schritt mittels nachgeschalteter ddPCR positiv getestet. Eines dieser Kinder erwies sich im Referenztest als falsch-positiv (PPV 90; 95 %-KI: [59,6; 98,2]).

Die Berechnung eines gepoolten Effekts ist aufgrund der unterschiedlichen Indextests sowie der unterschiedlichen Definition der Zielerkrankung nicht sinnvoll.

Unklar bleibt, wie viele erkrankte Kinder nicht gefunden werden. Aufgrund der geringen Prävalenz ist davon auszugehen, dass der überwiegende Anteil negativer Testergebnisse im Indextest richtig ist. Insgesamt deuten die Ergebnisse der Studien zur diagnostischen Güte darauf hin, dass die untersuchten Testverfahren für ein Screening bei Neugeborenen auf 5q-assozierte SMA geeignet sind.

4.5 Zusammenfassung der Beleglage

Vergleichende Interventionsstudien der Screeningkette konnten nicht identifiziert werden.

Zu dem Vergleich eines **früh- versus einen spätsymptomatischen Therapiebeginn** lagen Daten vor. Für den kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung sowie für den Endpunkt Erreichen motorischer Meilensteine zeigt sich jeweils ein Anhaltspunkt für unterschiedliche Effekte in Bezug auf den Therapiebeginn bei symptomatischen Kindern. Bezüglich beider Endpunkte profitieren Kinder mehr von einem frühsymptomatischen als von einem spätsymptomatischen Therapiebeginn. Für die Endpunkte SUEs, schwere UEs und Therapieabbrüche wegen UEs ergibt sich kein Anhaltspunkt für unterschiedliche Effekte in Bezug auf den Therapiebeginn bei symptomatischen Kindern mit infantiler SMA. Für andere Endpunkte lagen keine oder keine verwertbaren Daten vor.

Zu dem Vergleich eines **prä- versus einen frühsymptomatischen Therapiebeginn** lagen Daten vor. Es zeigt sich ein dramatischer Unterschied für den Endpunkt Erreichen motorischer Meilensteine und somit ein Anhaltspunkt für einen höheren Nutzen eines präsymptomatischen Therapiebeginns im Vergleich zu einem frühsymptomatischen Therapiebeginn. Für die Endpunkte schwerwiegende unerwünschte Ereignisse und schwere unerwünschte Ereignisse zeigen sich statistisch signifikante Unterschiede zugunsten des präsymptomatischen Therapiebeginns. Diese beobachteten Unterschiede werden als nicht groß genug eingeschätzt, als dass sie jeweils nicht allein durch die Einwirkung von Störgrößen zu erklären sein könnten. Es ergibt sich somit jeweils kein Anhaltspunkt für einen geringeren Schaden des präsymptomatischen im Vergleich zu einem frühsymptomatischen Therapiebeginn. Für die Endpunkte Gesamtüberleben, dauerhafte Beatmung sowie Therapieabbrüche wegen UEs und Rückenschmerzen (als SUE und schweres UE) zeigen sich jeweils keine statistisch signifikanten Unterschiede. Die Kriterien eines dramatischen Effekts sind nicht erfüllt. Es ergibt sich somit jeweils kein Anhaltspunkt für einen höheren oder geringeren Nutzen oder Schaden des präsymptomatischen im Vergleich zu einem frühsymptomatischen Therapiebeginn. Für andere Endpunkte lagen keine verwertbaren Daten vor.

Die Ergebnisse der **Studien zur diagnostischen Güte** deuten darauf hin, dass die untersuchten Testverfahren für ein Screening bei Neugeborenen auf 5q-assozierte SMA geeignet sind. Daten zur Anzahl falsch-negativer Ergebnisse liegen nicht vor.

Insgesamt wird anhand der Zusammenführung der verfügbaren Ergebnisse zum Therapiebeginn sowie zur diagnostischen Güte mittels Linked-Evidence-Ansatz ein Hinweis für einen Nutzen des Neugeborenencreenings auf 5q-assozierte SMA im Vergleich zu keinem Neugeborenencreening abgeleitet.

5 Einordnung des Arbeitsergebnisses

Datenlage zur Vorverlegung des Therapiebeginns

Früh- versus spätsymptomatischen Therapiebeginn mit Nusinersen

Es lag ausschließlich 1 kleine RCT mit potenziell hochverzerrten Ergebnissen zur Bewertung eines früh- versus einen spätsymptomatischen Therapiebeginn mit dem einzigen bislang in Deutschland zugelassenen Arzneimittel Nusinersen für Kinder mit einem Krankheitsbeginn im Neugeborenenalter (< 7 Monaten) vor. Die Studie beleuchtet demnach ausschließlich einen Teil der an 5q-assoziierter SMA erkrankten Kinder, nämlich die mit der infantilen Form der SMA (auch: SMA-Typ I). Durch den vorzeitigen Abbruch der Studie liegen zudem lediglich Daten zu einer begrenzten Beobachtungsdauer (Median: < 10 Monate) vor, Langzeitergebnisse beispielsweise zu unerwünschten Wirkungen bleiben demnach offen.

Die Ergebnisse der Studie ENDEAR lassen darauf schließen, dass bei Kindern mit der infantilen Form der 5q-assozierten SMA ein Therapiebeginn mit Nusinersen innerhalb von 12 Wochen nach Symptomeintritt im Vergleich zu einem späteren Therapiebeginn vorteilhaft ist im Hinblick auf das Risiko, zu versterben oder dauerhaft beatmet werden zu müssen, sowie im Hinblick auf das Erreichen motorischer Meilensteine.

Daten zu (S)UEs konnten aus dem veröffentlichten Modul 4 des Dossiers [26] zur Nutzenbewertung des G-BA zum Wirkstoff Nusinersen [35] sowie aus ergänzenden Unterlagen des Herstellers für die berichtsrelevanten Subgruppen (früh- versus spätsymptomatischen Therapiebeginn) zu den Endpunkten SUEs, Therapieabbrüche aufgrund von UEs und schweren UEs herangezogen werden. Für alle 3 Endpunkte zeigte sich keine statistisch signifikante Interaktion, womit anzunehmen ist, dass Unterschiede zwischen einem früh- und einem spätsymptomatischen Therapiebeginn hinsichtlich dieser Endpunkte – wenn überhaupt welche vorhanden sind – nicht wesentlich sind. Diese Ergebnisse sind mangels Beschreibung der einzelnen aufgetretenen (S)UEs eingeschränkt interpretierbar. Es bleibt unklar, welche (S)UEs bei Kindern mit einem früh- versus einen spätsymptomatischen Therapiebeginn häufiger oder seltener vorkamen. Anhand der Ergebnisse zur Gesamtpopulation der Studie können jedoch folgende Informationen dazu abgeleitet werden: Im Studienbericht wird kein SUE als durch die Studienmedikation bedingt klassifiziert [25]. Einige wenige (nicht als schwerwiegend klassifizierte) UEs wurden mit der Art der Applikation (Lumbalpunktion) assoziiert, z. B. Übelkeit und postprozedurale Schwellung [25,36]. Schwerwiegende respiratorische Ereignisse kamen in beiden Behandlungsgruppen bei allen ausgewerteten Kindern vor [26], jedoch bestehen Unklarheiten bezüglich der Auswertung und Operationalisierung dieses Endpunkts (siehe auch [35]). Das Auftreten schwerer UEs unterschied sich statistisch signifikant zugunsten von Nusinersen (n = 45 [56 %] versus Scheinbehandlung n = 33 [80 %]; OR: 0,31; 95 %-KI: [0,13; 0,76]; p < 0,010 [26]). Durch die höhere Anzahl von Todesfällen in der Scheinbehandlungsgruppe war die Beobachtungs- und somit auch die Expositionsdauer der Kinder in der Nusinersengruppe tendenziell länger, was die Ergebnisse zugunsten der Nusinersentherapie unterstreicht. Diese Ergebnisse verdeutlichen, dass die beobachteten Ereignisse nicht ausschließlich als Nebenwirkungen der Therapie einzustufen sind, sondern

sich weitgehend durch Symptome oder das Fortschreiten der Grunderkrankung erklären lassen. Selbst wenn sich bei einzelnen UEs nachteilige Effekte zeigen würden, könnte im vorliegenden Fall eines schwerkranken Patientenkollektivs davon ausgegangen werden, dass die gezeigten Vorteile des früheren Therapiebeginns hinsichtlich des kombinierten Endpunkts Zeit bis Tod oder dauerhafte Beatmung sowie hinsichtlich des Endpunkts Erreichen motorischer Meilensteine dadurch nicht gänzlich aufgewogen werden würden. Einschränkend ist zu allen vorgenannten Punkten anzumerken, dass die verfügbaren Ergebnisse auf einem stark limitierten Beobachtungszeitraum beruhen, der mögliche Langzeitschäden nicht abdeckt, sodass hierüber Unklarheit besteht. Hieran schließt sich an, dass es keine Langzeitdaten zur Verträglichkeit des Arzneimittels gibt und dazu, ob langfristig möglicherweise immunologische Abwehrreaktionen auftreten können [37]. Langfristig könnte insbesondere auch die Art der Anwendung – die wiederholte intrathekale Injektion (Lumbalpunktion) – zu (S)UEs führen. In der Studie CHERISH an älteren Kleinkindern mit 5q-assoziierter SMA wurde im Zusammenhang mit der Lumbalpunktion z. B. das postpunktionelle Syndrom als Nebenwirkung beobachtet [26]. Mit der laufenden Beobachtungsstudie SHINE [38] könnte ein Teil dieser Fragen in einigen Jahren beantwortet werden. In dieser Studie werden Langzeitdaten zu Patientinnen und Patienten erfasst, die in der Interventionsgruppe der eingeschlossenen Studie ENDEAR sowie weiterer Studien wie CHERISH [39-42] oder EMBRACE [43,44] mit Nusinersen behandelt wurden.

Es bleiben auch viele weitere Fragen unbeantwortet. Bislang ist beispielsweise ungeklärt, wie lange die Behandlung fortgesetzt werden sollte. Die Verabreichung bei Patientinnen und Patienten mit Skoliose [23] oder Spondylodese [45] gestaltet sich zudem häufig schwierig und bedarf häufig einer Steuerung mithilfe bildgebender Verfahren mit Kontrastmittel, Anästhesie und meist auch einer stationären Überwachung [46,47]. Auch werden mögliche Auswirkungen der mit der intrathekalen Verabreichung einhergehenden wiederholten Anästhesien auf die Hirnentwicklung von Kleinkindern diskutiert [48-50]. Ferner steigt bei einer Steuerung der Lumbalpunktionen mittels radiologischer Verfahren das Krebsrisiko [45,51,52].

Präsymptomatischer versus frühsymptomatischen Therapiebeginn mit Nusinersen

Um die Frage zu beantworten, ob eine Therapievorverlegung von der symptomatischen auf die präsymptomatische Situation nützt, lag ausschließlich 1 kleine retrospektive vergleichende Studie vor, die der Hersteller auf Anfrage zur Verfügung gestellt hat. Diese Studie umfasst ausschließlich Kinder mit 2 *SMN2*-Kopien (Daten zu Kindern mit 3 *SMN2*-Kopien konnten aufgrund zu geringer Fallzahlen nicht zur vorliegenden Nutzenbewertung herangezogen werden). Sie beleuchtet demnach ausschließlich den Teil der Kinder mit der schwersten Verlaufsform einer 5q-assozierten SMA, nämlich diejenigen, bei denen bereits eine infantile SMA diagnostiziert wurde (Vergleichsgruppe mit frühsymptomatischem Therapiebeginn), beziehungsweise diejenigen, die ohne Therapie höchstwahrscheinlich eine infantile SMA entwickeln würden (Interventionsgruppe mit präsymptomatischem Therapiebeginn). Ergebnisse lagen lediglich zum Lebensalter von 1 Jahr vor, Langzeitergebnisse beispielsweise zu unerwünschten Wirkungen bleiben demnach offen.

Die Ergebnisse der Studie Biogen 2019 lassen darauf schließen, dass Kinder mit diagnostizierter 5q-assoziierter SMA, 2 *SMN2*-Kopien und damit einer wahrscheinlichen Prognose für eine infantile SMA am meisten von einer Nusinersentherapie profitieren, wenn diese bereits präsymptomatisch begonnen wird. Die Vorteile zeigten sich insbesondere im Hinblick auf das Erreichen motorischer Meilensteine und ließen sich nicht allein durch Verzerrung erklären (dramatischer Effekt).

Die vom Hersteller zur Verfügung gestellten Ergebnisse zu Kindern mit 3 *SMN2*-Kopien weisen in dieselbe Richtung wie die Ergebnisse der Kinder mit 2 *SMN2*-Kopien [27] und unterstützen somit die gezeigten Vorteile des präsymptomatischen Therapiebeginns.

Wären für die Vergleichsgruppe der Studie Biogen 2019 nicht nur die frühsymptomatisch, sondern alle symptomatisch behandelten Kinder aus der Studie ENDEAR berücksichtigt worden, ist davon auszugehen, dass die beobachteten Vorteile des präsymptomatischen Therapiebeginns noch größer ausgefallen wären.

Zusammenfassende Einordnung der Daten zum Therapiebeginn

In der Gesamtschau ist den in der vorliegenden Nutzenbewertung eingeschlossenen Daten zur Therapie im Sinne einer „Dosis-Wirkungs-Beziehung“ eine „Zeitpunkt-Wirkungs-Beziehung“ zu entnehmen, der zufolge bessere Therapieergebnisse erzielt werden, je früher die Therapie eingesetzt wird.

Datenlage zur diagnostischen Güte

Die 4 eingeschlossenen Studien zur diagnostischen Güte weisen ein VOPT-Design auf, sodass unklar bleibt, wie viele Kinder durch die diagnostische Testung auf 5q-assozierte SMA nicht gefunden werden. Grundsätzlich deuten die Ergebnisse der Studien zur diagnostischen Güte aber darauf hin, dass die untersuchten Testverfahren für ein Screening bei Neugeborenen auf 5q-assozierte SMA geeignet sind.

Mit den untersuchten Testverfahren können heterozygote Träger identifiziert werden. Heterozygote Träger erkranken nicht an einer SMA. Die Erhebung von Informationen zum Trägerstatus im Rahmen einer Reihenuntersuchung ist nach § 16 Gendiagnostikgesetz problematisch, denn nach der Gesetzesbegründung soll ein Screening im Hinblick auf Anlageträger für rezessive Erkrankungen in Deutschland nicht zulässig sein [53]. Sofern die Einführung eines Neugeborenencreenings auf SMA in Deutschland erwogen wird, muss dies Beachtung finden.

Konsequenzen der Ergebnisse für ein Neugeborenencreening

Patientinnen und Patienten mit späterem SMA-Krankheitsbeginn

Die für die Fragestellung des vorliegenden Berichts identifizierte Evidenz zeigt auf, dass Kinder mit diagnostizierter 5q-assoziierter SMA und prognostizierter oder bestätigter infantiler Form von einem Neugeborenencreening profitieren würden. Zu Patientinnen und Patienten mit späterem SMA-Krankheitsbeginn konnten keine relevanten Therapiestudien identifiziert

werden, sodass unklar bleibt, ob auch für sie ein Neugeborenencreening von Nutzen wäre. Im Rahmen eines solchen Screenings werden grundsätzlich jedoch alle SMA-Typen identifiziert (mit Ausnahme derjenigen mit heterozygoter Deletion und Punktmutation auf dem vorhandenen *SMN1*-Gen, was ca. 5 % der Patientinnen und Patienten betrifft). Die frühe Diagnose von später (erst nach Jahren) auftretender SMA hat ethische Implikationen, denn die Vorhersage, in welchem Alter die SMA bei Patientinnen und Patienten mit zum Beispiel ≥ 4 *SMN2*-Kopien ausbricht, ist weniger sicher als bei der infantilen SMA [14]. In den eingeschlossenen Diagnosestudien wurden unterschiedliche Herangehensweisen für den Umgang mit diesen Patientinnen und Patienten gewählt. In Chien 2017 hat die Studiengruppe versucht, Patientinnen und Patienten mit später auftretender SMA (beziehungsweise deren Eltern) nicht zu beunruhigen. Sie wurden gebeten, sich nur beim Auftreten von Symptomen zu melden. In ihrer Publikation fordert die Studiengruppe einen kritischen Umgang bezüglich des Abwägens des möglichen Nutzens und der möglichen psychologischen Schäden eines Neugeborenencreenings für Patientinnen und Patienten mit ≥ 4 *SMN2*-Kopien ([14], siehe auch [54]). Auch die Autorinnen und Autoren der Studie Kraszewski 2018 adressieren dieses Teilkollektiv [34]. Sie fordern mehr Daten, um herauszufinden, ob Patientinnen und Patienten mit ≥ 4 *SMN2*-Kopien bereits im Neugeborenenalter behandelt werden sollten und wie häufig dies geschehen sollte. Im deutschen Neugeborenencreeningprojekt (Czibere/Vill 2019 [31,32]) wurde eine Behandlungsstrategie gewählt, die auf aktuellen Empfehlungen einer amerikanischen Arbeitsgruppe um Glascock basiert [55]. Diese sieht eine Überwachung der Kinder mit ≥ 4 *SMN2*-Kopien mittels konservativer Behandlungsstrategie (Beobachten) vor. Auf diese Weise wurden bei einem betroffenen Säugling im Alter von 8 Monaten erste klinische Symptome erkannt und es wurde unmittelbar eine Behandlung initiiert. Dieser Fall bestätigt, dass die Vorhersage, in welchem Alter die SMA bei Patientinnen und Patienten mit ≥ 4 *SMN2*-Kopien ausbricht, weniger sicher ist als bei Patientinnen und Patienten mit < 4 *SMN2*-Kopien. Mit dem im australischen Neugeborenencreeningprogramm nach Kariyawasam 2019 [33] angewendeten Vorgehen wäre dieser Säugling nicht identifiziert worden. Denn dort wurde die Zielerkrankungsdefinition für den Screeningtest enger gefasst, um Patientinnen und Patienten mit späterem SMA-Krankheitsbeginn explizit gar nicht zu identifizieren. Nur diejenigen Säuglinge, die < 4 *SMN2*-Kopien aufwiesen, wurden als testpositiv definiert. Individuen mit ≥ 4 *SMN2*-Kopien wurden als testnegativ beschrieben und wurden nicht gemeldet.

Säuglinge, die über ein Neugeborenencreening mit ≥ 4 *SMN2*-Kopien identifiziert werden, haben nicht zwangsläufig einen späten SMA-Krankheitsbeginn. Für diejenigen Patientinnen und Patienten mit tatsächlich spätem SMA-Krankheitsbeginn bleibt derzeit unklar, ob sie von einem Neugeborenencreening profitieren würden. Eine Übertragbarkeit der identifizierten Daten zum Therapiebeginn bei Kindern mit diagnostizierter 5q-assoziierter SMA und prognostizierter oder bestätigter infantiler Form ist nicht ohne Weiteres auf Patientinnen und Patienten mit spätem SMA-Krankheitsbeginn möglich (siehe auch [55]). Ein adäquater Umgang mit diesen Kindern und ihren Familien – einschließlich der Erwägung einer selbstbestimmten Entscheidung über den Erhalt dieser Information – ist bei Einführung eines Neugeborenencreenings dringend zu berücksichtigen.

Zentrale Therapiestudien ohne Berichtsrelevanz

Neben der eingeschlossenen Studie ENDEAR ist als weitere zentrale Studie CHERISH zu nennen [39-42], in der Nusinersen zur Behandlung von Patientinnen und Patienten mit später auftretender 5q-assoziierter SMA zum Einsatz kam. In dieser multizentrischen RCT wurden 126 symptomatische Kinder auf eine Behandlung mit Nusinersen oder einer Sham-Injektion randomisiert. Einschlusskriterien waren unter anderem ein Symptombeginn im Alter von ≥ 6 Monaten sowie ein Alter bei Studieneinschluss zwischen 2 und 12 Jahren. Das mediane Alter der Patientinnen und Patienten in der Nusinersengruppe bei Studieneinschluss liegt bei 4 Jahren und in der Scheinbehandlungsgruppe bei 3 Jahren. Erste Symptome zeigten sich in der Nusinersengruppe in einem mittleren Alter von 10 Monaten und in der Scheinbehandlungsgruppe von 11 Monaten. Die Krankheitsdauer (Alter bei Studieneinschluss abzüglich Alter bei Symptombeginn) betrug im Median 39 Monate (Nusinersengruppe) beziehungsweise 30 Monate (Scheinbehandlungsgruppe). Das Zeitfenster zwischen Diagnosestellung und Therapiebeginn betrug etwa 30 Monate (Nusinersengruppe) beziehungsweise 18 Monate (Scheinbehandlungsgruppe). Dies entspricht nicht der aktuellen Versorgungssituation in Deutschland, da seit der Zulassung von Nusinersen mit der Therapie heutzutage in der Regel direkt nach der Diagnosestellung begonnen wird. Aus diesem Grund wurde die Studie als nicht relevant ausgeschlossen. Zudem bietet diese Studie keinen Bezug zum Neugeborenencreening. Sie hätte nicht zur Klärung der Frage beitragen können, wann bei Kindern mit später auftretender SMA, die im Rahmen eines Neugeborenencreenings identifiziert wurden, idealerweise mit der Therapie begonnen werden sollte – beispielsweise direkt nach der Geburt, beim Auftreten erster Symptome oder kurz vorher, sofern der Symptombeginn ausreichend sicher geschätzt werden kann.

Des Weiteren ist die Studie EMBRACE zum Einsatz von Nusinersen bei Patientinnen und Patienten mit 5q-assoziierter SMA zu nennen [43,44,56]. In dieser wurden 21 Kinder auf eine Behandlung mit Nusinersen oder einer Sham-Injektion randomisiert, die weder für einen Einschluss in die Studie ENDEAR noch für einen Einschluss in die Studie CHERISH geeignet waren. Die Studie wurde vor Kurzem abgeschlossen. Eine Vollpublikation liegt nicht vor, im Studienregistereintrag werden jedoch erste Ergebnisse berichtet [56]. Auch diese Studie wurde als nicht relevant für die vorliegende Nutzenbewertung eingestuft. Die Frage eines früheren versus einen späteren Therapiebeginn war nicht Gegenstand dieser Studie und hätte gegebenenfalls anhand von Subgruppenanalysen beantwortet werden können. Allerdings war die Anzahl der Kinder pro Behandlungsgruppe (Nusinersen $n = 14$, Sham-Injektion $n = 7$) zu gering, um Subgruppenanalysen gemäß der Berichtsfragestellung (früherer versus späterer Therapiebeginn) durchzuführen.

Neue Therapieansätze in der Entwicklung

Neben dem seit 2017 in Deutschland zugelassenen Arzneimittel Nusinersen gibt es weitere Therapieansätze zur Behandlung der SMA, die sich derzeit noch in der Entwicklung befinden (siehe Kapitel 1). Nennenswert sind 2 Therapieansätze, die derzeit im Zulassungsverfahren (Prime-Status) der Europäischen Arzneimittel-Agentur (EMA) bewertet werden. Hier handelt es sich zum einen um ein orales Arzneimittel (Risdiplam), zum anderen um eine

Genersatztherapie (AVXS-101), die in den USA im Mai 2019 von der FDA für Kinder unter 2 Jahren zugelassen wurde [20,21]. Nachdem erste Studienergebnisse zu AVXS-101 darauf hindeuten, dass bei Kindern mit infantiler SMA ein früherer Behandlungsbeginn bezogen auf das Kindesalter zu einer besseren Entwicklung motorischer Meilensteine beitragen könnte, betonen Autorinnen und Autoren die Relevanz eines Neugeborenencreenings, um Kinder mit SMA frühzeitiger behandeln zu können [57]. Inwiefern Ergebnisse zu diesen Therapien zukünftig zur vorliegenden Nutzenbewertung beitragen können, kann derzeit jedoch nicht bewertet werden, da es zunächst einer Zulassung der Therapie in Deutschland bedarf. Möglicherweise könnten für die Beantwortung der Frage des Nutzens eines Neugeborenencreenings auf 5q-assozierte SMA jedoch zukünftig auch vergleichende Interventionsstudien (der Screeningkette) zu anderen Therapien als Nusinersen infrage kommen.

Laufendes Pilotprojekt zum Neugeborenencreening auf SMA in Deutschland

Seit Januar 2018 läuft in Bayern ein Pilotprojekt zum Neugeborenencreening auf Cystinose und SMA [31,32,58]. Bei diesem Pilotprojekt handelt es sich um eine Studie, die als 1-armige Interventionsstudie der Screeningkette betrachtet werden kann. Neugeborene, bei denen im Rahmen des Screenings eine SMA diagnostiziert wird, werden standardisiert in Abhängigkeit der *SMN2*-Kopienzahl entweder unmittelbar mit Nusinersen behandelt (≤ 3 *SMN2*-Kopien) oder standardisiert beobachtet, um zeitnah bei Auftreten von Symptomen die Therapie beginnen zu können (≥ 4 *SMN2*-Kopien). Aufgrund des 1-armigen Designs erfüllt die Studie nicht die Anforderungen, um als vergleichende Interventionsstudie der Screeningkette eingeschlossen zu werden. Im Sinne einer prospektiven diagnostischen Kohortenstudie entspricht sie jedoch den Einschlusskriterien von Studien zur diagnostischen Güte, sodass sie zu deren Bewertung herangezogen wurde (siehe Abschnitt 4.4).

Hinsichtlich der Ergebnisse dieser frühestmöglichen Behandlung der bislang im Screeningprojekt identifizierten Kinder handelt es sich in der Publikation Vill 2019 [32] um erste vorläufige Ergebnisse mit sehr kurzer Nachbeobachtungsdauer. Zur Beantwortung der Frage nach einem präsymptomatischen gegenüber einem späteren Therapiebeginn können diese mangels Vergleichsintervention nicht herangezogen werden. Dasselbe trifft zu auf Ergebnisse, die im Rahmen des Stellungnahmeverfahrens zur Verfügung gestellt wurden (siehe [59]). Diese Daten bieten Einblick in die patientenindividuellen Krankheitsverläufe und das Erreichen motorischer Meilensteine der bislang identifizierten Kinder mit 2, 3 oder 4 *SMN2*-Kopien. Es wurde vonseiten der Studiengruppe mitgeteilt, dass geplant ist, die im Rahmen dieses Neugeborenencreening-Programms detektierten Kinder mit einer Kohorte von Kindern zu vergleichen, die nach Symptombeginn erfasst wurden und im Register SMARtCARE dokumentiert werden (siehe A4.3). Ein solcher Vergleich könnte im Sinne einer retrospektiven vergleichenden Interventionsstudie der Screeningkette grundsätzlich weitere Erkenntnisse zur vorliegenden Nutzenbewertung liefern, sofern die Kriterien zum Einschluss von ebendiesen Studien erfüllt wären (siehe Kapitel 3 sowie ausführlich Abschnitt A2.1).

Publication Bias

Die vorliegenden Informationen deuten nicht auf einen Publication Bias hin.

6 Fazit

Zum Vergleich eines Neugeborenencreenings auf 5q-assoziierte spinale Muskelatrophie (SMA) versus kein Neugeborenencreening lagen keine vergleichenden Interventionsstudien der Screeningkette vor. Daher wurden Interventionsstudien, die einen Vergleich eines früheren mit einem späteren Therapiebeginn ermöglichen, sowie Studien zur diagnostischen Güte herangezogen, die mittels Linked-Evidence-Ansatz zusammengeführt wurden.

Für den **Vergleich eines früh- versus einen spätsymptomatischen Therapiebeginn** konnte 1 kleine randomisierte kontrollierte Studie mit kurzer Beobachtungsdauer herangezogen werden, die eine medikamentöse Therapie im Vergleich zu einer Scheinbehandlung bei Kindern mit infantiler 5q-assoziiertes SMA untersuchte. Anhand von Subgruppenanalysen zeigten sich für den kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung sowie für den Endpunkt Erreichen motorischer Meilensteine Effektunterschiede zwischen Kindern mit frühsymptomatischem Therapiebeginn (Krankheitsdauer ≤ 12 Wochen) und Kindern mit späterem Therapiebeginn (Krankheitsdauer > 12 Wochen). Bezüglich beider Endpunkte profitieren Kinder mehr von einem frühsymptomatischen als von einem spätsymptomatischen Therapiebeginn. Zu den Endpunkten schwerwiegende unerwünschte Ereignisse, schwere unerwünschte Ereignisse und Therapieabbrüche wegen unerwünschter Ereignisse zeigten sich jeweils keine Effektunterschiede zwischen den Subgruppen. Für andere Endpunkte lagen keine verwertbaren Daten vor.

Zum **Vergleich eines prä- versus einen frühsymptomatischen Therapiebeginn** wurde 1 kleine, vom Hersteller auf Anfrage zur Verfügung gestellte, retrospektive vergleichende Studie eingeschlossen. Es konnten Daten zu Kindern mit diagnostizierter 5q-assoziiertes SMA und 2 *SMN2*-Kopien (d. h. prognostizierter oder bestätigter infantiler SMA) herangezogen werden. Für den Endpunkt Erreichen motorischer Meilensteine zeigten sich große Effekte zugunsten des präsymptomatischen im Vergleich zum frühsymptomatischen Therapiebeginn (Krankheitsdauer ≤ 12 Wochen), die sich nicht allein durch Verzerrung erklären ließen (dramatischer Effekt). Bei den Endpunkten schwerwiegende unerwünschte Ereignisse und schwere unerwünschte Ereignisse zeigten sich statistisch signifikante Unterschiede zugunsten des präsymptomatischen Therapiebeginns. Diese beobachteten Unterschiede wurden als nicht groß genug eingeschätzt, als dass sie jeweils nicht allein durch die Einwirkung von Störgrößen zu erklären gewesen sein könnten. Bei den Endpunkten Gesamtüberleben, dauerhafte Beatmung sowie Therapieabbrüche wegen unerwünschter Ereignisse und Rückenschmerzen zeigten sich jeweils keine statistisch signifikanten Unterschiede. Die Kriterien eines dramatischen Effekts waren nicht erfüllt. Für andere Endpunkte wurden keine Daten angefragt.

Zu **Patientinnen und Patienten mit einem späteren Krankheitsbeginn als bei der infantilen Form** konnten keine vergleichenden Interventionsstudien herangezogen werden.

Für die Bewertung der **diagnostischen Güte** konnten 4 Studien herangezogen werden, in denen jedoch allein die positiven Testergebnisse überprüft wurden (Verification-of-only-positive-Testers-Design). Die Ergebnisse dieser Studien deuten darauf hin, dass die untersuchten

Testverfahren für ein Screening bei Neugeborenen auf 5q-assozierte SMA geeignet sind. Unklar bleibt, wie viele erkrankte Kinder nicht gefunden werden.

Zusammenfassend ergibt sich ein Hinweis für einen Nutzen des Neugeborenencreenings auf 5q-assozierte SMA im Vergleich zu keinem Screening. Dieses Ergebnis beruht zum einen auf Daten zur prä-, früh- und spätsymptomatischen Arzneimittelbehandlung von Kindern mit diagnostizierter 5q-assoziierter SMA und prognostizierter oder bestätigter infantiler Form. Den herangezogenen Daten ist eine Zeitpunkt-Wirkungs-Beziehung zu entnehmen, der zufolge bessere Therapieergebnisse erzielt werden, je früher die Therapie einsetzt. Zum anderen basiert die Ableitung eines Hinweises für einen Nutzen des Neugeborenencreenings auf der Eignung diagnostischer Testverfahren und der Möglichkeit, mittels Neugeborenencreening eine Vorverlegung der Diagnosestellung (und damit der Therapie) zu erlangen. Die verfügbaren Daten erlauben keine Schlüsse, ob durch das Screening identifizierte Kinder mit spätem Krankheitsbeginn der SMA (d. h. Symptombeginn erst nach Jahren) bereits von einem präsymptomatischen Therapiebeginn profitieren würden. Derzeit unklar ist somit vor allem der Umgang mit im Screening identifizierten Neugeborenen mit erwartbar spätem Krankheitsbeginn (≥ 4 *SMN2*-Kopien). Ein adäquater Umgang mit diesen Kindern und ihren Familien – einschließlich der Erwägung einer selbstbestimmten Entscheidung über das (Nicht-)Wissen zum Vorliegen milde verlaufender SMA-Formen – ist bei Einführung eines Neugeborenencreenings zu berücksichtigen.

Details des Berichts

A1 Projektverlauf

A1.1 Zeitlicher Verlauf des Projekts

Der Gemeinsame Bundesausschuss (G-BA) hat am 13.12.2018 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Bewertung eines Neugeborenen-screensings zur Früherkennung der 5q-assozierten SMA beauftragt.

In die Bearbeitung des Projekts wurden externe Sachverständige eingebunden.

Während der Erstellung des Berichtsplans wurden am 08.01.2019 Betroffene zur Diskussion von patientenrelevanten Endpunkten und relevanten Subgruppen konsultiert.

Der Berichtsplan in der Version 1.0 vom 20.02.2019 wurde am 27.02.2019 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 27.03.2019 konnten schriftliche Stellungnahmen eingereicht werden. Die Dokumentation der Anhörung zum Berichtsplan ist auf der Website des IQWiG veröffentlicht.

Eine Überarbeitung der Methoden des Berichtsplans war nicht notwendig.

Die vorläufige Bewertung, der Vorbericht in der Version 1.0 vom 23.10.2019, wurde am 30.10.2019 auf der Website des IQWiG veröffentlicht und zur Anhörung gestellt. Bis zum 28.11.2019 konnten schriftliche Stellungnahmen eingereicht werden. Unklare Aspekte aus den schriftlichen Stellungnahmen zum Vorbericht wurden am 09.01.2020 in einer wissenschaftlichen Erörterung mit den Stellungnehmenden diskutiert. Die wesentlichen Argumente aus den Stellungnahmen werden im Kapitel „Kommentare“ des vorliegenden Abschlussberichts gewürdigt.

Der vorliegende Abschlussbericht beinhaltet die Änderungen, die sich aus der Anhörung ergeben haben.

Im Anschluss an die Anhörung erstellte das IQWiG den vorliegenden Abschlussbericht, der 4 Wochen nach Übermittlung an den G-BA auf der Website des IQWiG veröffentlicht wird. Die zum Vorbericht eingegangenen Stellungnahmen und das Protokoll der wissenschaftlichen Erörterung werden in einem gesonderten Dokument „Dokumentation der Anhörung zum Vorbericht“ zeitgleich mit dem Abschlussbericht im Internet bereitgestellt.

A1.2 Spezifizierungen und Änderungen im Projektverlauf

Vorbericht im Vergleich zum Berichtsplan 1.0

Neben redaktionellen Änderungen ergaben sich folgende Spezifizierungen oder Änderungen im Vorbericht:

- Der Hintergrund (Kapitel 1) wurde um die Diagnoseklassifikation der ICD-10 erweitert; des Weiteren wurde ein Hinweis auf eine neue Gentherapie ergänzt, die im Mai 2019 in den USA zugelassen wurde; ferner wurde die Beschreibung der Diagnose mittels Blutprobe spezifiziert.
- Als Informationsquelle wurde neben der EMA auch die FDA durchsucht.
- Eine positive Aussage zum Nutzen der Therapievorverlegung wurde in der vorliegenden Bewertung insofern spezifiziert und weit gefasst, als hierfür auch Studien berücksichtigt wurden, die mittels Subgruppenanalysen den Vergleich eines früheren versus einen späteren Therapiebeginn ermöglichten. Eine positive Aussage zum Nutzen der Therapievorverlegung konnte sich hierbei durch eine statistisch signifikante Interaktion (Anhaltspunkt für Unterschiede der Effekte je nach Therapiebeginn) bei gleichzeitigem Vorliegen statistisch signifikanter Behandlungseffekte zugunsten der Intervention mindestens in der Gruppe mit früherem Therapiebeginn ergeben.

Abschlussbericht im Vergleich zum Vorbericht

Neben redaktionellen Änderungen ergaben sich folgende Spezifizierungen oder Änderungen im Abschlussbericht:

- Vom Hersteller wurde der angefragte Vergleich zu einem prä- versus einen frühsymptomatischen Therapiebeginn zur Verfügung gestellt (Biogen 2019); diese Studie wurde in die vorliegende Nutzenbewertung eingeschlossen.
- Infolge der Nachrecherche wurden 2 weitere Studien zur diagnostischen Güte (Czibere/Vill 2019, Kariyawasam 2019) identifiziert und in die vorliegende Nutzenbewertung eingeschlossen.
- In der Einordnung des Arbeitsergebnisses (Kapitel 5) wurden Aktualisierungen aufgrund der neuen Daten vorgenommen. Darüber hinaus wurde dieses Kapitel um weitere Aspekte erweitert.
- Es erfolgte eine Aktualisierung des Fazits (Kapitel 6).
- Es erfolgte eine Aktualisierung der Ergebnisse zur umfassenden Informationsbeschaffung (Abschnitt A3.1).
- Abschnitt A4.1 wurde um 1 systematische Übersicht ergänzt und Abschnitt A4.2 zur kritischen Reflexion des Vorgehens wurde um weitere Aspekte erweitert.
- Es wurde ein Abschnitt A4.3 zur Würdigung der Anhörung ergänzt.

A2 Methodik gemäß Berichtsplan 1.0

Die folgenden Abschnitte geben den Wortlaut der Berichtsmethodik aus dem Berichtsplan wieder. Über diese Methodik hinausgehende Spezifizierungen oder Änderungen der Methoden im Projektverlauf werden in Abschnitt A1.2 erläutert. Im folgenden Text wird an den entsprechenden Stellen auf diesen Abschnitt verwiesen.

Diese Bewertung wird auf Grundlage der Allgemeinen Methoden 5.0 [30] erstellt.

Der Nutzen eines Neugeborenencreenings auf 5q-assozierte SMA kann a) in der Identifikation betroffener Neugeborener im präsymptomatischen Zustand liegen, um sie entsprechend früher einer Therapie zuführen zu können (Therapiebeginn bereits im präsymptomatischen Zustand), oder b) auch darin, dass die Therapie unmittelbar mit dem Symptomaufreten eingeleitet werden kann (kürzere Dauer zwischen Symptomen und Therapiebeginn). Es sollen beide möglichen Nutzenszenarien [a) und b)] des Neugeborenencreenings auf 5q-assozierte SMA betrachtet werden, die auf 2 Wegen bewertet werden können. Diese Herangehensweisen werden im Folgenden beschrieben.

Nutzenbewertung anhand von vergleichenden Interventionsstudien der Screeningkette

Der Nutzen von Screeningmaßnahmen lässt sich anhand von prospektiv geplanten vergleichenden Interventionsstudien mit einer (idealerweise randomisierten) Zuteilung von Personen zu einer Strategie mit beziehungsweise ohne Anwendung der Screeningmaßnahme wie auch anhand von vergleichenden Kohortenstudien der gesamten Screeningkette und der Betrachtung patientenrelevanter Endpunkte bewerten [30]. In solchen Studien erhalten die Neugeborenen in der Kontrollgruppe kein SMA-Screening. Die der Interventionsgruppe zugeteilten Neugeborenen erhalten ein 5q-assozierte-SMA-Screening und ihnen werden entsprechend dem Untersuchungsergebnis des Screenings gegebenenfalls eine diagnostische Abklärung und Therapie zugewiesen.

Nutzenbewertung anhand von vergleichenden Studien zum Therapiebeginn und Studien zur Bewertung der diagnostischen Güte

Liegen vergleichende Interventionsstudien der Screeningkette für die Nutzenbewertung nicht oder in nicht ausreichender Quantität und Qualität vor, kann eine Bewertung der einzelnen Bausteine der Screeningkette erfolgen. Für die Nutzenbewertung wird der Nutzen eines früheren gegenüber einem späteren Therapiebeginn erfasst, die diagnostische Güte untersucht und gesundheitsbezogene Konsequenzen für falsch-positive, richtig-positive, falsch-negative sowie richtig-negative Befunde gegenübergestellt. Dazu werden vergleichende Interventionsstudien zum Therapiebeginn und Studien zur diagnostischen Güte herangezogen.

Für die Untersuchung des Therapiebeginns können a) Studien zum Vergleich der Behandlung präsymptomatischer mit der Behandlung symptomatischer Patientinnen und Patienten herangezogen werden. Dann sind in einzuschließenden Studien konkrete Angaben zum Diagnosezeitpunkt und -anlass erforderlich, die für die Gruppe des früheren Therapiebeginns

den Rückschluss zulassen, dass die Betroffenen bereits im präsymptomatischen Zustand identifiziert und entsprechend präsymptomatisch einer Therapie zugeführt wurden. Der Nutzen des Screenings kann dann dadurch abgeleitet werden, dass ein früherer (präsymptomatischer) gegenüber einem späteren (symptomatischen) Therapiebeginn einen höheren Nutzen zeigt und gleichzeitig der Screeningtest eine hinreichende diagnostische Güte aufweist.

Weiter können b) Studien zum Vergleich früherer versus späteren Therapiebeginn in Bezug auf die Zeitspanne zwischen Symptom- und Therapiebeginn herangezogen werden. Dann sind in den einzuschließenden Studien konkrete Angaben zur Dauer zwischen dem Auftreten erster Symptome und dem Therapiebeginn erforderlich, aus denen Ergebnisse für einen Vergleich früherer versus späteren Therapiebeginn hergeleitet werden können. Der Zeitpunkt der Diagnosestellung ist dabei unerheblich. Der Nutzen des Screenings kann in diesem Fall ebenfalls dadurch abgeleitet werden, dass ein früherer (mit Einsetzen erster Symptome) gegenüber einem späteren (Therapiebeginn bei jeweils symptomatischen Patientinnen und Patienten) einen höheren Nutzen zeigt und gleichzeitig der Screeningtest eine hinreichende diagnostische Güte aufweist. Hinzukommen muss, dass das Ergebnis auf die Screeningsituation übertragbar ist.

A2.1 Kriterien für den Einschluss von vergleichenden Interventionsstudien der Screeningkette in die Untersuchung

A2.1.1 Population

In die Bewertung werden Studien mit Neugeborenen aufgenommen.

A2.1.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention ist das 5q-assozierte-SMA-Screening bei Neugeborenen. Der Zeitpunkt der Probenentnahme soll auf den in der Kinder-Richtlinie des G-BA genannten Zeitrahmen für das erweiterte Neugeborenencreening übertragbar sein (siehe Kapitel 1). Als Vergleichsintervention gilt kein Screening.

Die Therapieoptionen und die Maßnahmen, die sich an ein positives Testergebnis anschließen, müssen auf die in Deutschland etablierten Maßnahmen und Therapiemethoden übertragbar sein (siehe Kapitel 1). Die Anwendung der in den Studien eingesetzten Arzneimittel muss im Rahmen des für Deutschland gültigen Zulassungsstatus erfolgen.

Die in den Studien zur Screeningkette angewendeten Diagnoseverfahren müssen auf die Situation des Neugeborenencreenings im Rahmen der Kinder-Richtlinie in Deutschland übertragbar sein.

A2.1.3 Patientenrelevante Endpunkte

Für die Untersuchung werden folgende patientenrelevante Endpunkte betrachtet:

- Mortalität (Gesamtüberleben, krankheitsspezifisches Überleben),

- Morbidität (zum Beispiel Entwicklungs- und Wachstumsstörungen wie das Erreichen motorischer Meilensteine, Krankenhausaufenthalte, durch 5q-assozierte SMA hervorgerufene Atemnot),
- (schwerwiegende) unerwünschte Ereignisse,
- gesundheitsbezogene Lebensqualität des Kindes.

A2.1.4 Studientypen

Randomisierte kontrollierte Studien (RCTs) sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention.

Für alle unter Abschnitt A2.1.2 genannten Interventionen und alle unter Abschnitt A2.1.3 genannten Endpunkte ist eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar.

Für den zu erstellenden Bericht werden daher in erster Linie RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.

Es ist möglich, dass aufgrund der Seltenheit der 5q-assozierten SMA keine RCTs vorliegen. In diesem Fall werden nicht randomisierte vergleichende Interventionsstudien für die unter Abschnitt A2.1.2 genannte Intervention ausgewertet. Gleichzeitig ist es denkbar, dass zu prüfende Interventionen einen so großen Effekt aufweisen, dass sich dieser in Studien mit niedrigerem Evidenzniveau nicht allein durch Verzerrung erklären lässt (dramatischer Effekt). Wenn die auf RCTs und nicht randomisierten vergleichenden Interventionsstudien basierende Datenlage nicht reicht, um den patientenrelevanten Nutzen und Schaden des 5q-assozierte-SMA-Screenings bei Neugeborenen mit ausreichender Ergebnissicherheit schätzen zu können, werden zu dieser Fragestellung auch vergleichende Kohortenstudien (auch retrospektive oder mit historischem Vergleich) als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen. Dies gilt, sofern das Problem einer möglichen Strukturungleichheit bei der Planung und Auswertung der entsprechenden Studien berücksichtigt wurde und zwischen den Kollektiven vergleichbare Bedingungen vorlagen. Weiterhin müssen Daten zu wesentlichen Basischarakteristika beider Gruppen verfügbar sein, um die Übertragbarkeit der Ergebnisse auf andere Populationen und den Einfluss wichtiger Störgrößen und Verzerrungen abschätzen zu können. Auch Publikationen von Registerauswertungen (hier verstanden als retrospektive oder historische vergleichende Kohortenstudien), in denen 2 Kollektive miteinander verglichen werden und die Auswertung prospektiv geplant worden ist, werden berücksichtigt, wenn sie die oben genannten Kriterien erfüllen. Liegen keine RCTs oder prospektiv geplanten Kohortenstudien mit zeitlich paralleler Kontrollgruppe vor, so sind auf Basis vergleichender Kohortenstudien (retrospektive oder mit historischem Vergleich) Nutzensaussagen nur möglich, wenn die vorliegenden Effekte so groß sind, dass sie sich nicht allein durch Verzerrung erklären lassen (dramatischer Effekt).

A2.1.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.1.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfügbar sein.

A2.1.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 2: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien der Screeningkette)

Einschlusskriterien	
INS1	Neugeborene (siehe auch Abschnitt A2.1.1)
INS2	Prüfintervention: Neugeborenencreening auf 5q-assoziierte SMA (siehe auch Abschnitt A2.1.2)
INS3	Vergleichsintervention: kein Neugeborenencreening auf 5q-assoziierte SMA (siehe auch Abschnitt A2.1.2)
INS4	patientenrelevante Endpunkte wie in Abschnitt A2.1.3 formuliert
INS5	Studientypen: RCTs, nicht randomisierte vergleichende Interventionsstudien, vergleichende Kohortenstudien (auch retrospektiv oder mit historischem Vergleich) (siehe auch Abschnitt A2.1.4)
INS6	Die Publikationssprache ist Deutsch oder Englisch.
INS7	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [60] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [61], TREND- [62] oder STROBE-Statements [63] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; RCT: randomisierte kontrollierte Studie; SMA: spinale Muskelatrophie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs</p>	

A2.2 Kriterien für den Einschluss von vergleichenden Interventionsstudien zum Therapiebeginn in die Untersuchung

Studien, die einen Vergleich eines früheren versus einen späteren Therapiebeginn ermöglichen, werden im Rahmen des vorliegenden Berichts systematisch recherchiert und ausgewertet, wenn vergleichende Interventionsstudien der Screeningkette nicht oder in nicht ausreichender Quantität und Qualität vorliegen.

A2.2.1 Population

In die Bewertung werden Studien mit Patientinnen und Patienten mit 5q-assoziierter SMA aufgenommen. Die Diagnosestellung bei Patientinnen und Patienten mit früherem Therapiebeginn muss auf die Screeningsituation bei Neugeborenen übertragbar sein. Studien mit ausschließlich symptomatischen Patientinnen und Patienten müssen konkrete Angaben zur Dauer zwischen dem Auftreten erster Symptome und dem Therapiebeginn enthalten.

A2.2.2 Prüf- und Vergleichsintervention

Die zu prüfende Intervention bildet ein früherer, gegebenenfalls präsymptomatischer Therapiebeginn. Als Vergleichsintervention gilt ein späterer Therapiebeginn. Bezüglich des Behandlungsbeginns bei späterem Therapiebeginn besteht keine Einschränkung. Dieser muss aber der Versorgungssituation in Deutschland entsprechen. Die Therapie besteht in einer medikamentösen Behandlung und / oder Physiotherapie.

Die Therapieoptionen und Maßnahmen, die sich an die Diagnosestellung anschließen, müssen auf die in Deutschland etablierten Maßnahmen und Therapiemethoden übertragbar sein. Die Anwendung der in den Studien eingesetzten Arzneimittel muss im Rahmen des für Deutschland gültigen Zulassungsstatus erfolgen.

Die in den Studien zum Therapiebeginn angewendeten Diagnoseverfahren müssen auf die Situation des Neugeborenencreenings im Rahmen der Kinder-Richtlinie in Deutschland übertragbar sein.

A2.2.3 Patientenrelevante Endpunkte

Für die Untersuchung werden die unter Abschnitt A2.1.3 genannten patientenrelevanten Endpunkte betrachtet.

A2.2.4 Studientypen

RCTs sind, sofern sie methodisch adäquat und der jeweiligen Fragestellung angemessen durchgeführt wurden, mit der geringsten Ergebnisunsicherheit behaftet. Sie liefern daher die zuverlässigsten Ergebnisse für die Bewertung des Nutzens einer medizinischen Intervention.

Kann keine RCT identifiziert werden, wird auf nicht randomisierte vergleichende Interventionsstudien zurückgegriffen. Ist die auf RCTs und nicht randomisierten vergleichenden Interventionsstudien basierende Datenlage nicht hinreichend, um den patientenrelevanten Nutzen und Schaden der Therapievorverlagerung mit ausreichender Ergebnissicherheit schätzen zu können, werden vergleichende Kohortenstudien (auch retrospektive oder mit historischem Vergleich) als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen, sofern das Problem einer möglichen Strukturungleichheit (unfairer Vergleich) der Beobachtungsgruppen adäquat in der Planung und Auswertung der Studien berücksichtigt wurde (siehe Abschnitt A2.1.4). Solche Studien können zwar die Aussage von aggregierten Ergebnissen aus qualitativ belastbaren RCTs in der Regel nicht qualitativ ändern, diese aber

gegebenenfalls bestärken. Liegen keine RCTs oder prospektiv geplanten Kohortenstudien mit zeitlich paralleler Kontrollgruppe vor, so sind auf Basis vergleichender Kohortenstudien (retrospektive oder mit historischem Vergleich) Nutzaussagen nur möglich, wenn die vorliegenden Effekte so groß sind, dass sie sich nicht allein durch Verzerrung erklären lassen (dramatischer Effekt).

A2.2.5 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.2.6 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfügbar sein.

A2.2.7 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 3: Übersicht über die Kriterien für den Studieneinschluss (vergleichende Interventionsstudien zum Therapiebeginn)

Einschlusskriterien	
INT1	Patientinnen und Patienten mit 5q-assoziiertes SMA (siehe auch Abschnitt A2.2.1)
INT2	Prüfintervention: frühere Behandlung (siehe auch Abschnitt A2.2.2)
INT3	Vergleichsintervention: spätere Behandlung (siehe auch Abschnitt A2.2.2)
INT4	patientenrelevante Endpunkte wie in Abschnitt A2.1.3 formuliert
INT5	Studientypen: RCTs, nicht randomisierte vergleichende Interventionsstudien, vergleichende Kohortenstudien (auch retrospektiv oder mit historischem Vergleich) (siehe auch Abschnitt A2.2.4)
INT6	Die Publikationssprache ist Deutsch oder Englisch.
INT7	Vollpublikation verfügbar ^a
<p>a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Studienbericht gemäß ICH E3 [60] oder ein Bericht über die Studie, der den Kriterien des CONSORT- [61], TREND- [62] oder STROBE-Statements [63] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind.</p> <p>CONSORT: Consolidated Standards of Reporting Trials; ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; RCT: randomisierte kontrollierte Studie; SMA: spinale Muskelatrophie; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology; TREND: Transparent Reporting of Evaluations with Nonrandomized Designs</p>	

A2.3 Kriterien für den Einschluss von Studien zur diagnostischen Güte in die Untersuchung

Sollte sich aus der Vorverlegung des Therapiebeginns eine positive Aussage zum Nutzen ergeben, werden im Rahmen des vorliegenden Berichts auch Studien zur diagnostischen Güte

zur Nutzenbewertung herangezogen. Die Charakteristika und Einschlusskriterien für diese Studien werden im Folgenden beschrieben.

A2.3.1 Population

In die Bewertung werden Studien mit Neugeborenen aufgenommen.

A2.3.2 Indextest

Als Indextest betrachtet werden alle in den Studien verwendeten diagnostischen Testverfahren oder Kombinationen von Testverfahren zur Testung auf 5q-assoziierte SMA unter Verwendung von Trockenblut der Filterpapierkarten. Der Zeitpunkt der Probenentnahme soll auf den in der Kinder-Richtlinie des G-BA genannten Zeitrahmen für das erweiterte Neugeborenencreening übertragbar sein. Die laboranalytische Methodik und die dazugehörigen Spezifikationen für den Test zur Unterscheidung positiver und negativer Ergebnisse müssen prospektiv festgelegt worden sein. Die Testentwicklung und -validierung müssen an voneinander unabhängigen Stichproben durchgeführt worden sein.

A2.3.3 Referenztest

Der Referenztest sind genetische Analysen. Bei unauffälligem Befund im Indextest kann alternativ auch die Nachbeobachtung akzeptiert werden.

A2.3.4 Zielgrößen

Eingeschlossen werden Studien, aus denen Daten zur Berechnung der diagnostischen Güte im Hinblick auf die Entdeckung von 5q-assoziiertes SMA ableitbar sind.

A2.3.5 Studientypen

Um die diagnostische Güte des Indextests zur Erkennung von 5q-assoziiertes SMA bei Neugeborenen möglichst unverzerrt bestimmen zu können, soll eine Gruppe von Neugeborenen, die zu einem bestimmten Zeitpunkt prospektiv rekrutiert und auf 5q-assoziiertes SMA gescreent wurde, zeitnah (Querschnittsdesign) mit dem Referenztest (nach)untersucht beziehungsweise bei unauffälligem Befund im Indextest nachbeobachtet werden. Dabei sind ein konsekutiver, das heißt nicht selektiver Einschluss der Neugeborenen und die Dokumentation der fehlenden Werte notwendig.

Ist die Datenlage aus solchen Studien unzureichend, werden in die vorliegende Bewertung aufgrund der Seltenheit von 5q-assoziiertes SMA auch diagnostische retrospektive Kohortenstudien aufgenommen.

Ist die Datenlage aus Studien, die sowohl positive als auch negative Ergebnisse im Indextest mit dem Referenztest direkt überprüfen (komplette Verifikation), unzureichend, können Studien im VOPT-Design herangezogen werden. Dabei werden alle positiven Ergebnisse im Indextest mit dem Referenztest untersucht [64] und es können Aussagen zum PPV als Maß der

diagnostischen Güte getroffen werden. Eine Bewertung der testnegativen Fälle und damit eine Bestimmung der Sensitivität oder Spezifität des Tests ist mit solchen Studien nicht möglich.

A2.3.6 Studiendauer

Hinsichtlich der Studiendauer besteht keine Einschränkung.

A2.3.7 Publikationssprache

Die Publikation muss in deutscher oder englischer Sprache verfügbar sein.

A2.3.8 Tabellarische Darstellung der Kriterien für den Studieneinschluss

In der folgenden Tabelle sind die Kriterien aufgelistet, die Studien erfüllen müssen, um in die Bewertung eingeschlossen zu werden.

Tabelle 4: Übersicht über die Kriterien für den Studieneinschluss (Studien zur diagnostischen Güte)

Einschlusskriterien	
D1	Neugeborene (siehe auch Abschnitt A2.3.1)
D2	Indextest: Screening auf 5q-assozierte SMA unter Verwendung von Filterpapierkarten (siehe auch Abschnitt A2.3.2)
D3	Referenztest: genetische Analyse, Nachbeobachtung (siehe auch Abschnitt A2.3.3)
D4	Zielgrößen: personenbezogene Vierfeldertafel-Daten zur diagnostischen Güte (siehe auch Abschnitt A2.3.4)
D5	diagnostische Querschnitt- und Kohortenstudien (siehe auch Abschnitt A2.3.5)
D6	Die Publikationssprache ist Deutsch oder Englisch.
D7	Vollpublikation verfügbar ^a
a: Als Vollpublikation gilt in diesem Zusammenhang auch ein Bericht über die Studie, der den Kriterien des STARD- [65] oder STROBE-Statements [63] genügt und eine Bewertung der Studie ermöglicht, sofern die in diesen Dokumenten enthaltenen Informationen zur Studienmethodik und zu den Studienergebnissen nicht vertraulich sind. SMA: spinale Muskelatrophie; STARD: Standards for the Reporting of Diagnostic Accuracy Studies; STROBE: Strengthening the Reporting of Observational Studies in Epidemiology	

A2.4 Vorgehen im Falle einer Zulassungsänderung im Projektverlauf

Sofern sich im Projektverlauf Änderungen im Zulassungsstatus der zu bewertenden Interventionen ergeben, werden die Kriterien für den Studieneinschluss gegebenenfalls an die neuen Zulassungsbedingungen angepasst. Die jeweils vorgenommenen Änderungen werden im Vorbericht beziehungsweise im Abschlussbericht explizit vermerkt. In der vorliegenden Nutzenbewertung werden nur solche Wirkstoffe berücksichtigt, die am 14.01.2019 (Stichtag) in Deutschland zugelassen und im Handel waren.

A2.5 Einschluss von Studien, die die vorgenannten Kriterien nicht vollständig erfüllen

Für die Einschlusskriterien INS1, INT1 und D1 (Population) reicht es aus, wenn bei mindestens 80 % der eingeschlossenen Patientinnen und Patienten dieses Kriterium erfüllt ist. Liegen für solche Studien entsprechende Subgruppenanalysen vor, wird auf diese Analysen zurückgegriffen. Studien, bei denen das Einschlusskriterium INS1, INT1 beziehungsweise D1 bei weniger als 80 % erfüllt ist, werden nur dann eingeschlossen, wenn entsprechende Subgruppenanalysen vorliegen.

Ebenfalls eingeschlossen werden Studien, die zu mindestens 80 % die Einschlusskriterien INS2, INT2 und D2 erfüllen (Prüfintervention, bezogen auf die Interventionsgruppe der Studie, beziehungsweise Indextest bei Diagnosestudien) und zu mindestens 80 % die Einschlusskriterien INS3 und INT3 (Vergleichsintervention, bezogen auf die Vergleichsgruppe der Studie).

A2.6 Umfassende Informationsbeschaffung

A2.6.1 Informationsquellen

Für die umfassende Informationsbeschaffung wird eine systematische Recherche nach relevanten Studien beziehungsweise Dokumenten durchgeführt. Folgende primäre und weitere Informationsquellen sowie Suchtechniken werden dabei berücksichtigt:

Primäre Informationsquellen

- bibliografische Datenbanken
 - MEDLINE
 - Embase
 - Cochrane Central Register of Controlled Trials
 - Cochrane Database of Systematic Reviews
 - HTA Database
- Studienregister
 - U.S. National Institutes of Health. ClinicalTrials.gov
 - World Health Organization. International Clinical Trials Registry Platform Search Portal
 - European Medicines Agency. EU Clinical Trials Register (für Studien zum Therapiebeginn)
 - Biogen Clinical Trial Results

Darüber hinaus wird im folgenden Studienregister nach Einträgen mit Ergebnisberichten (für Studien zum Therapiebeginn) gesucht:

- PharmNet.Bund. Arzneimittel-Informationssystem

- Herstelleranfragen

Die Herstelleranfragen ermöglichen den Überblick über alle von Herstellern durchgeführten Studien unabhängig vom Publikationsstatus. Für potenziell relevante Studien aus dieser Übersicht werden vollständige Studienunterlagen (i. d. R. vollständige Studienberichte) angefordert und in die Bewertung einbezogen. Alle Informationen zur Methodik und zu den Ergebnissen, die in die Nutzenbewertung eingehen, werden im Bericht des IQWiG veröffentlicht. Angefragt wird bei folgendem Hersteller:

- Biogen GmbH

Weitere Informationsquellen und Suchtechniken

- Zulassungsbehörden
 - European Medicines Agency
- durch den G-BA übermittelte Dokumente
- G-BA-Website und IQWiG-Website
- Anwendung weiterer Suchtechniken
 - Sichten von Referenzlisten identifizierter systematischer Übersichten
- Anhörungen zum Berichtsplan und Vorbericht
- Autorenanfragen

Zu diesem Vorgehen gab es eine Spezifizierung im Projektverlauf, siehe Abschnitt A1.2.

A2.6.2 Selektion relevanter Studien

Selektion relevanter Studien bzw. Dokumente aus den Ergebnissen der bibliografischen Recherche

Die in bibliografischen Datenbanken identifizierten Treffer werden in einem 1. Schritt anhand ihres Titels und, sofern vorhanden, Abstracts in Bezug auf ihre potenzielle Relevanz bezüglich der Einschlusskriterien (siehe Tabelle 2, Tabelle 3, Tabelle 4) bewertet. Als potenziell relevant erachtete Dokumente werden in einem 2. Schritt anhand ihres Volltextes auf Relevanz geprüft. Beide Schritte erfolgen durch 2 Personen unabhängig voneinander. Diskrepanzen werden durch Diskussion zwischen beiden aufgelöst.

Es ist vorgesehen, einzelne Recherchen für eine Studie heranzuziehen, in der die Effizienz der Studienselktion untersucht wird [30]. In diesen Fällen erfolgen beide Schritte durch 3 Personen unabhängig voneinander in 3 verschiedenen Screeningtools. Die Ergebnisse der Selektion werden nach der Volltextbewertung zusammengefasst.

Selektion relevanter Studien bzw. Dokumente aus weiteren Informationsquellen

Rechercheergebnisse aus den folgenden Informationsquellen werden von 2 Personen unabhängig voneinander in Bezug auf ihre Relevanz bewertet:

- Studienregister,
- durch den G-BA übermittelte Dokumente.

Rechercheergebnisse aus den darüber hinaus berücksichtigten Informationsquellen werden von 1 Person auf Studien gesichtet. Die identifizierten Studien werden dann auf ihre Relevanz geprüft. Der gesamte Prozess wird anschließend von einer 2. Person überprüft. Sofern in einem der genannten Selektionsschritte Diskrepanzen auftreten, werden diese jeweils durch Diskussion zwischen den beiden aufgelöst.

A2.7 Informationsbewertung und Synthese

Alle für die Nutzenbewertung notwendigen Informationen werden aus den Unterlagen zu den eingeschlossenen Studien in standardisierte Tabellen extrahiert. Ergeben sich im Abgleich der Informationen aus unterschiedlichen Dokumenten zu einer Studie (aber auch aus multiplen Angaben zu einem Aspekt innerhalb eines Dokuments selbst) Diskrepanzen, die auf die Interpretation der Ergebnisse erheblichen Einfluss haben könnten, wird dies an den entsprechenden Stellen im Ergebnisteil des Berichts dargestellt.

A2.7.1 Darstellung der Einzelstudien

A2.7.1.1 Darstellung der Einzelstudien (vergleichende Interventionsstudien)

Die Ergebnisse zu den in den Studien berichteten patientenrelevanten Endpunkten werden im Bericht vergleichend beschrieben.

Die relevanten Ergebnisse werden endpunktspezifisch pro Studie auf ihr jeweiliges Verzerrungspotenzial überprüft. Anschließend werden die Informationen zusammengeführt und analysiert. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unter Abschnitt A2.7.3 bis A2.7.5 beschriebenen Verfahren eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen erfolgt in jedem Fall.

Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patientinnen und Patienten basieren, das heißt, wenn der Anteil der Patientinnen und Patienten, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist.

Die Ergebnisse werden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Anteile nicht berücksichtigter Patientinnen und Patienten zwischen den Gruppen größer als 15 Prozentpunkte ist.

A2.7.1.2 Darstellung der Einzelstudien (Studien zur diagnostischen Güte)

Die Ergebnisse zu den in den Studien berichteten Zielgrößen werden im Bericht vergleichend beschrieben.

Die relevanten Ergebnisse werden auf ihr jeweiliges Verzerrungspotenzial und Übertragbarkeit überprüft. Anschließend werden die Informationen zusammengeführt und analysiert. Wenn möglich werden über die Gegenüberstellung der Ergebnisse der Einzelstudien hinaus die unter Abschnitt A2.7.3 bis A2.7.5 beschriebenen Verfahren eingesetzt. Eine abschließende zusammenfassende Bewertung der Informationen erfolgt in jedem Fall.

Ergebnisse fließen in der Regel nicht in die Nutzenbewertung ein, wenn diese auf weniger als 70 % der in die Auswertung einzuschließenden Patientinnen und Patienten basieren, das heißt, wenn der Anteil der Patientinnen und Patienten, die nicht in der Auswertung berücksichtigt werden, größer als 30 % ist.

Die Ergebnisse werden auch dann nicht in die Nutzenbewertung einbezogen, wenn der Unterschied der Anteile nicht berücksichtigter Patientinnen und Patienten zwischen den Gruppen größer als 15 Prozentpunkte ist.

A2.7.2 Bewertung des Verzerrungspotenzials der Ergebnisse

A2.7.2.1 Bewertung des Verzerrungspotenzials der Ergebnisse (vergleichende Interventionsstudien)

Das Verzerrungspotenzial der Ergebnisse wird endpunktspezifisch für jede in die Nutzenbewertung eingeschlossene Studie bewertet. Dazu werden insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Kriterien systematisch extrahiert und bewertet:

A: Kriterien zur endpunktübergreifenden Bewertung des Verzerrungspotenzials der Ergebnisse

- Erzeugung der Randomisierungssequenz (bei randomisierten Studien)
- Verdeckung der Gruppenzuteilung (bei randomisierten Studien)
- zeitliche Parallelität der Gruppen (bei nicht randomisierten kontrollierten Studien)
- Vergleichbarkeit der Gruppen beziehungsweise Berücksichtigung prognostisch relevanter Faktoren (bei nicht randomisierten kontrollierten Studien)
- Verblindung der Patientin oder des Patienten sowie der behandelnden Personen (bei randomisierten Studien)
- ergebnisunabhängige Berichterstattung

B: Kriterien zur endpunktspezifischen Bewertung des Verzerrungspotenzials der Ergebnisse

- Verblindung der Endpunkterheber
- Umsetzung des Intention-to-treat(ITT)-Prinzips
- ergebnisunabhängige Berichterstattung

Für die Ergebnisse randomisierter Studien wird das Verzerrungspotenzial zusammenfassend als niedrig oder hoch eingestuft. Wird bereits hinsichtlich der unter (A) aufgeführten Kriterien ein endpunktübergreifend hohes Verzerrungspotenzial festgestellt, gilt dieses damit für alle Ergebnisse aller Endpunkte als hoch, unabhängig von der Bewertung endpunktspezifischer Aspekte. Andernfalls finden anschließend die unter (B) genannten Kriterien pro Endpunkt Berücksichtigung.

Das Verzerrungspotenzial der Ergebnisse nicht randomisierter vergleichender Studien wird aufgrund der fehlenden Randomisierung zusammenfassend grundsätzlich als hoch bewertet.

A2.7.2.2 Bewertung des Verzerrungspotenzials der Ergebnisse (Studien zur diagnostischen Güte)

Die Bewertung des Verzerrungspotenzials und der Übertragbarkeit der Studien zur diagnostischen Güte erfolgt auf Basis des QUADAS-2-Instruments [66]. Das Verzerrungspotenzial von Studien zur diagnostischen Güte wird als niedrig oder hoch eingestuft.

A2.7.3 Metaanalysen

A2.7.3.1 Metaanalysen (vergleichende Interventionsstudien)

Die geschätzten Effekte und Konfidenzintervalle aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Anschließend wird die Heterogenität des Studienpools anhand des statistischen Tests auf Vorliegen von Heterogenität [67] untersucht. Ergibt der Heterogenitätstest ein statistisch nicht signifikantes Ergebnis ($p \geq 0,05$), wird davon ausgegangen, dass die Schätzung eines gemeinsamen (gepoolten) Effekts sinnvoll ist. Im Fall von mindestens 5 Studien erfolgt die Metaanalyse mithilfe des Modells mit zufälligen Effekten nach der Methode von Knapp und Hartung unter Verwendung des Heterogenitätsschätzers nach Paule und Mandel [68]. Als Ergebnis wird der gemeinsame Effekt inklusive Konfidenzintervall dargestellt. Weil die Heterogenität im Fall weniger Studien nicht verlässlich geschätzt werden kann, werden bei 4 oder weniger Studien gegebenenfalls Modelle mit festem Effekt verwendet. Dazu müssen die Studien ausreichend ähnlich sein und es darf keine Gründe geben, die gegen die Anwendung eines Modells mit festem Effekt sprechen. Ist ein Modell mit festem Effekt nicht vertretbar, kann eine qualitative Zusammenfassung erfolgen.

Ergibt der Heterogenitätstest ein statistisch signifikantes Ergebnis ($p < 0,05$), wird im Fall von mindestens 5 Studien nur das Prädiktionsintervall dargestellt. Bei 4 oder weniger Studien erfolgt eine qualitative Zusammenfassung. In beiden Fällen wird außerdem untersucht, welche Faktoren diese Heterogenität möglicherweise verursachen. Dazu zählen methodische Faktoren (siehe Abschnitt A2.7.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt A2.7.5).

Abgesehen von den genannten Modellen können in bestimmten Situationen und mit besonderer Begründung Alternativen wie z. B. das Betabinomialmodell bei binären Daten [69] angewendet werden.

A2.7.3.2 Metaanalysen (Studien zur diagnostischen Güte)

Die Punktschätzungen und dazugehörigen univariaten 95 %-Konfidenzintervalle [70] aus den Studien werden mittels Forest Plots zusammenfassend dargestellt. Außerdem wird, sofern die dafür nötigen Anforderungen erfüllt sind, für die Testgütekriterien eine Metaanalyse anhand der Sensitivität und Spezifität in einem bivariaten Modell durchgeführt [71]. Die Schätzung der Modellparameter erfolgt über ein generalisiertes lineares gemischtes Modell [72,73]. Der Algorithmus zum Schätzen der Parameter im bivariaten Modell kann zu unpräzisen Schätzungen führen, das heißt zu Schätzungen mit zu großen Standardfehlern und entsprechenden Konfidenzregionen. Auch kann der Algorithmus gegebenenfalls keine Schätzungen liefern, wenn das Maximum-Likelihood-Verfahren nicht konvergiert. In beiden Fällen fehlen brauchbare Schätzungen. Die Gründe hierfür können beispielsweise sein, dass zu wenige Studien vorliegen oder dass einzelne Studien extreme Werte aufweisen. Sind die resultierenden Schätzungen unpräzise, werden die Ergebnisse der bivariaten Metaanalysen in der Regel nicht dargestellt.

Falls die bivariate Metaanalyse präzise Schätzungen liefert, so werden bei diagnostischen Studien die beobachteten Paare aus Sensitivität und Spezifität 2-dimensional grafisch dargestellt. Des Weiteren werden die aus der bivariaten Metaanalyse gewonnenen Schätzungen für die Erwartungswerte als gepooltes Paar der Sensitivität und der Spezifität mit der dazugehörigen 95 %-Konfidenzregion dargestellt [74].

In Ausnahmefällen, wie beispielsweise beim Vorliegen von mehreren großen Studien mit niedrigem Verzerrungspotenzial, werden die Ergebnisse geeigneter univariater statistischer Tests, das heißt für die Sensitivität und Spezifität getrennt, dargestellt.

Sollten die Sensitivität und Spezifität nicht berechenbar sein – zum Beispiel, weil nur Studien im VOPT-Design eingeschlossen wurden –, wird der PPV dargestellt und metaanalytisch zusammengefasst. Hierbei kommen Likelihood-basierte Verfahren auf Basis der individuellen Patientendaten zum Einsatz. Die Schätzung der Modellparameter erfolgt über ein generalisiertes lineares gemischtes Modell [75].

Das Vorliegen von Heterogenität wird anhand von Sensitivitätsanalysen untersucht.

A2.7.4 Sensitivitätsanalysen

Bestehen Zweifel an der Robustheit von Ergebnissen wegen methodischer Faktoren, die beispielsweise durch die Wahl bestimmter Cut-off-Werte, Ersetzungsstrategien für fehlende Werte, Erhebungszeitpunkte oder Effektmaße begründet sein können, ist geplant, den Einfluss solcher Faktoren in Sensitivitätsanalysen zu untersuchen. Das Ergebnis solcher Sensitivitätsanalysen kann die Sicherheit der aus den beobachteten Effekten abgeleiteten Aussagen beeinflussen. Ein als nicht robust eingestuftes Ergebnis kann zum Beispiel dazu führen, dass nur ein Hinweis auf anstelle eines Belegs für einen (höheren) Nutzen attestiert wird (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt A2.7.6).

A2.7.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse werden hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht. Ziel ist es, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Für einen Nachweis unterschiedlicher Effekte ist die auf einem Homogenitäts- beziehungsweise Interaktionstest basierende statistische Signifikanz Voraussetzung. In die Untersuchung werden die vorliegenden Ergebnisse aus Regressionsanalysen, die Interaktionsterme beinhalten, und aus Subgruppenanalysen einbezogen. Außerdem erfolgen eigene Analysen in Form von Meta-regressionen oder Metaanalysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren. Subgruppenanalysen werden nur durchgeführt, falls jede Subgruppe mindestens 10 Personen umfasst und bei binären Daten mindestens 10 Ereignisse in 1 der Subgruppen aufgetreten sind. Es ist vorgesehen, folgende Faktoren bezüglich einer möglichen Effektmodifikation in die Analysen einzubeziehen:

- Geschlecht,
- Alter,
- Dauer zwischen Symptom- und Therapiebeginn,
- diagnostische Testverfahren.

Sollten sich aus den verfügbaren Informationen weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden.

Bei Identifizierung möglicher Effektmodifikatoren erfolgt gegebenenfalls eine Präzisierung der aus den beobachteten Effekten abgeleiteten Aussagen. Beispielsweise kann der Beleg eines (höheren) Nutzens auf eine spezielle Subgruppe von Patientinnen und Patienten eingeschränkt werden (zur Ableitung von Aussagen zur Beleglage siehe Abschnitt A2.7.6).

A2.7.6 Aussagen zur Beleglage

Für jeden Endpunkt wird eine Aussage zur Beleglage des (höheren) Nutzens und (höheren) Schadens getroffen. Dabei sind 4 Abstufungen der Aussagesicherheit möglich: Es liegt entweder ein Beleg (höchste Aussagesicherheit), ein Hinweis (mittlere Aussagesicherheit), ein Anhaltspunkt (schwächste Aussagesicherheit) oder keine dieser 3 Situationen vor. Der letzte Fall tritt ein, wenn keine Daten vorliegen oder die vorliegenden Daten keine der 3 übrigen Aussagen zulassen. In diesem Fall wird die Aussage „Es liegt kein Anhaltspunkt für einen (höheren) Nutzen oder (höheren) Schaden vor“ getroffen.

Die regelhaft abzuleitende Aussagesicherheit ist von den in Tabelle 5 dargestellten Kriterien abhängig. Die qualitative Ergebnissicherheit ist abhängig vom Design der Studie. Ergebnisse randomisierter Studien mit niedrigem Verzerrungspotenzial haben eine hohe, Ergebnisse

randomisierter Studien mit hohem Verzerrungspotenzial eine mäßige qualitative Ergebnissicherheit. Ergebnisse nicht randomisierter vergleichender Studien haben eine geringe qualitative Ergebnissicherheit.

Aussagen zum Nutzen auf Basis von Studien mit niedrigerer Evidenzstufe sind nur in Verbindung mit einem dramatischen Effekt möglich.

Der Nutzen des Screenings kann durch die Gegenüberstellung der gesundheitsbezogenen Konsequenzen der möglichen Testergebnisse und ihrer Wahrscheinlichkeiten zusammen mit einer Aussage zum Nutzen eines früheren Therapiebeginns abgeleitet werden. Die Aussagesicherheit bezüglich des Nutzens des Screenings berücksichtigt dann sowohl die Aussagesicherheit bezüglich des Nutzens eines früheren Therapiebeginns als auch das Verzerrungspotenzial bezüglich der diagnostischen Güte.

Tabelle 5: Regelhaft abgeleitete Aussagesicherheiten für verschiedene Evidenzsituationen beim Vorliegen von Studien derselben qualitativen Ergebnissicherheit

		Anzahl Studien				
		1 (mit statistisch signifikantem Effekt)	≥ 2			
			homogen	heterogen		
			Metaanalyse statistisch signifikant	gleichgerichtete Effekte ^a		
deutlich	mäßig	nein				
Qualitative Ergebnis- sicherheit	hoch	Hinweis	Beleg	Beleg	Hinweis	–
	mäßig	Anhaltspunkt	Hinweis	Hinweis	Anhaltspunkt	–
	gering	–	Anhaltspunkt	Anhaltspunkt	–	–
a: Gleichgerichtete Effekte liegen vor, wenn trotz Heterogenität eine deutliche oder mäßige Richtung der Effekte erkennbar ist.						

A3 Details der Ergebnisse

A3.1 Umfassende Informationsbeschaffung

A3.1.1 Primäre Informationsquellen

A3.1.1.1 Bibliografische Datenbanken

Abbildung 1 zeigt das Ergebnis der systematischen Literaturrecherche in den bibliografischen Datenbanken und der Studienselektion gemäß den Kriterien für den Studieneinschluss von vergleichenden Interventionsstudien der Screeningkette sowie von Studien zur diagnostischen Güte. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Abschnitt A7.1.1. Die letzte Suche fand am 22.10.2019 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A6.2.1.

Abbildung 2 zeigt das Ergebnis der systematischen Literaturrecherche in den bibliografischen Datenbanken und der Studienselektion gemäß den Kriterien für den Studieneinschluss von vergleichenden Interventionsstudien zum Therapiebeginn. Die Suchstrategien für die Suche in bibliografischen Datenbanken finden sich in Abschnitt A7.1.2. Die letzte Suche fand am 22.10.2019 statt.

Die Referenzen der als Volltexte geprüften, aber ausgeschlossenen Treffer finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A6.2.2.

Unter den ausgeschlossenen Dokumenten wurde 1 Designpublikation identifiziert, die 1 Studie ohne bisher berichtete Ergebnisse beschreibt (siehe Abschnitt A3.1.4).

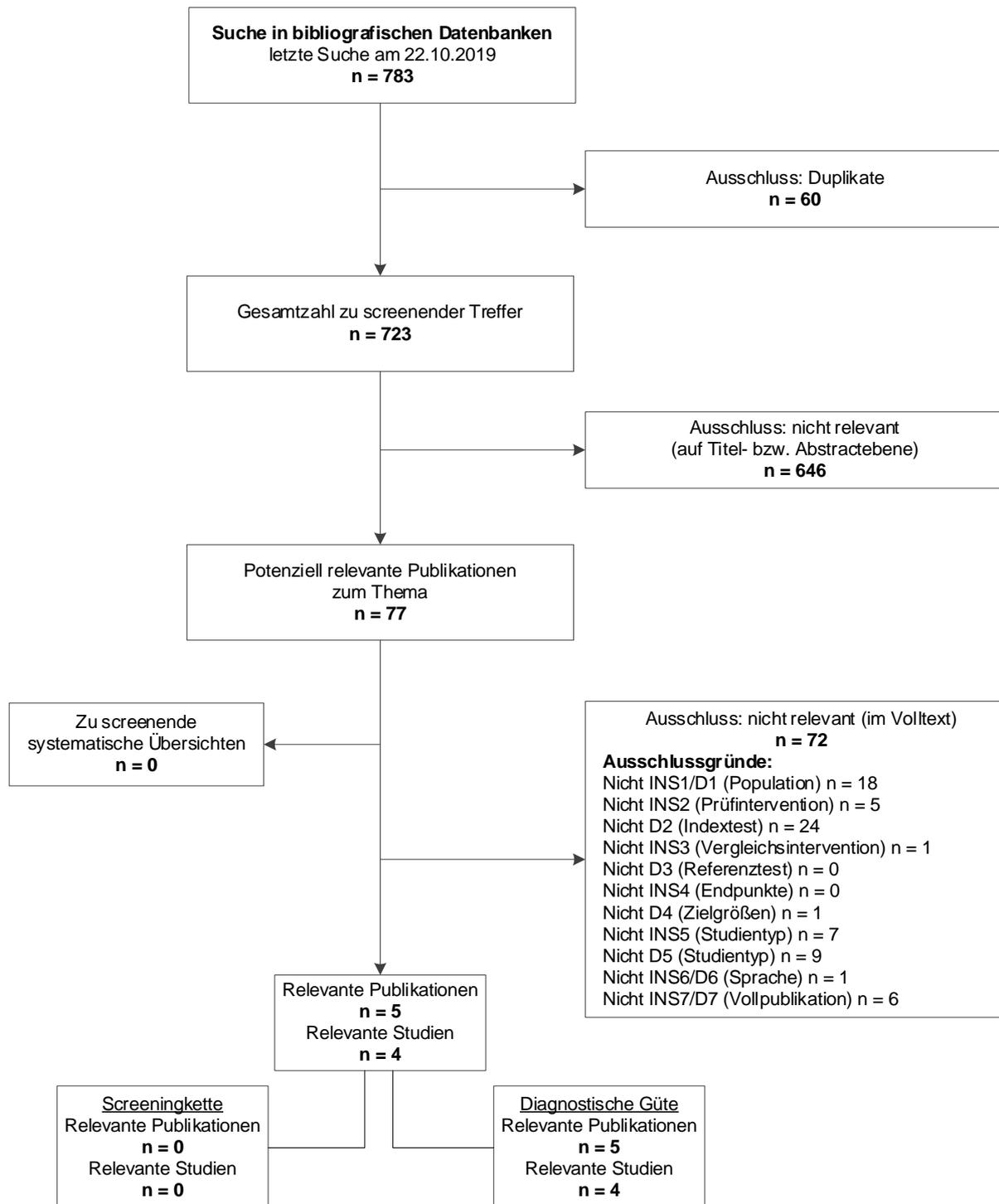


Abbildung 1: Ergebnis der bibliografischen Recherche und der Studienselektion für die Suche nach vergleichenden Interventionsstudien der Screeningkette sowie nach Studien zur diagnostischen Güte

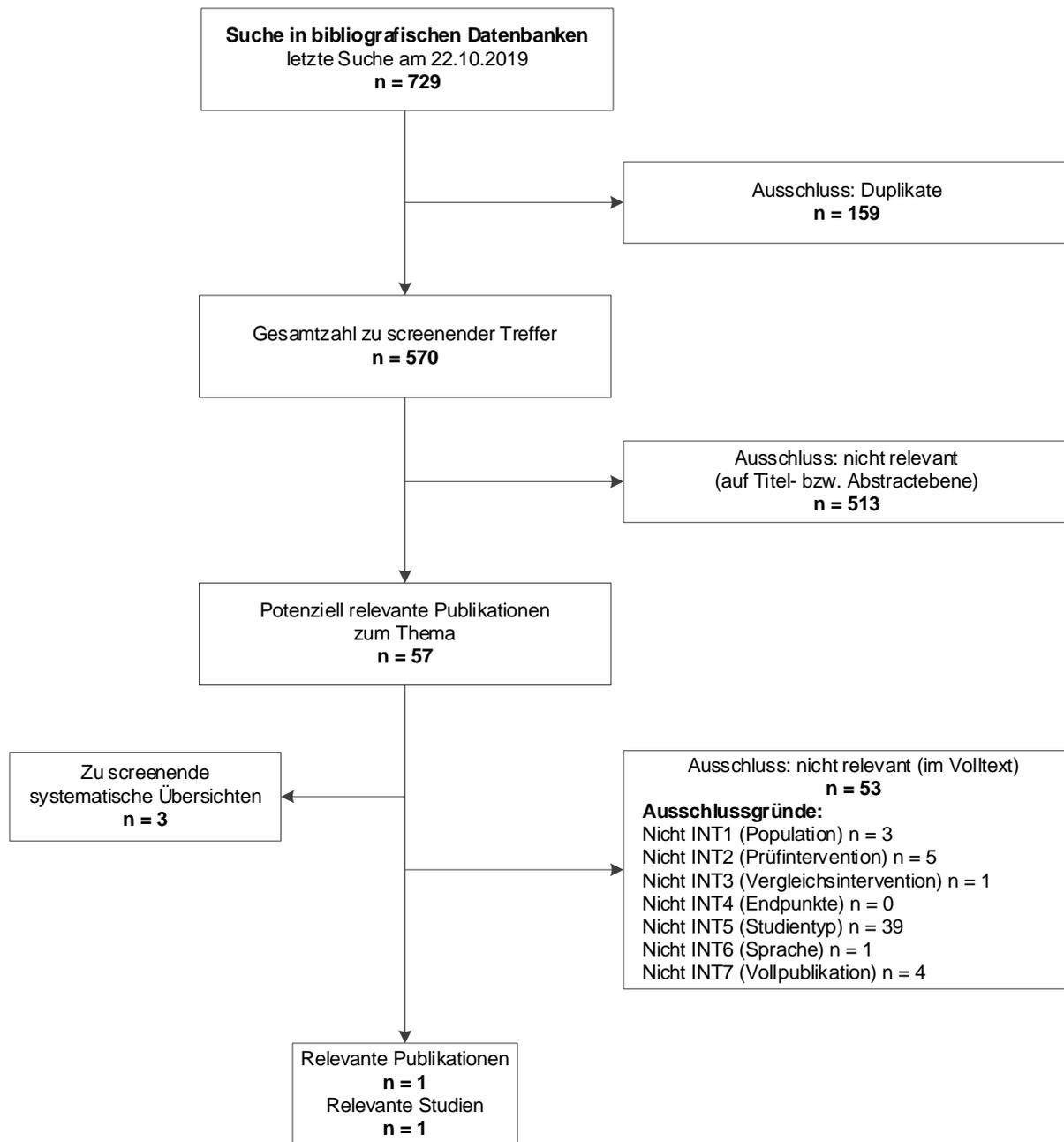


Abbildung 2: Ergebnis der bibliografischen Recherche und der Studienselektion für die Suche nach vergleichenden Interventionsstudien zum Therapiebeginn

A3.1.1.2 Studienregister

Durch die Suche in Studienregistern wurden folgende relevante Studien beziehungsweise Dokumente identifiziert (Tabelle 6):

Tabelle 6: In Studienregistern identifizierte relevante Studien bzw. Dokumente

Studie	Studienregister-ID	Studienregister	Ergebnisbericht in Studienregister vorhanden
Vergleichende Interventionsstudien der Screeningkette			
keine Studien oder Dokumente identifiziert			
Vergleichende Interventionsstudien zum Therapiebeginn			
ENDEAR	NCT02193074	ClinicalTrials.gov [76]	ja [80]
	EudraCT 2013-004422-29	EU Clinical Trials Register [77]	ja [81]
	DRKS00009141	Deutsches Register Klinischer Studien [78]	nein
	ISIS396443-CS3B	Biogen Clinical Trial Results [79]	nein
	NCT02594124 ^a	ClinicalTrials.gov [38]	nein
Studien zur diagnostischen Güte			
Chien 2017	NCT02123186	ClinicalTrials.gov [82]	nein
a: Bei diesem Eintrag handelt es sich um die laufende SHINE-Studie; es ist zu erwarten, dass dieser Studie Langzeitdaten zum Interventionsarm der Studie ENDEAR zu entnehmen sind.			

In den Studienregistern wurden zur diagnostischen Güte 3 laufende Studien ohne berichtete Ergebnisse identifiziert (siehe Abschnitt A3.1.4).

Die Suchstrategien für die Suche in Studienregistern finden sich in Abschnitt A7.2. Die letzte Suche in Studienregistern fand am 23.10.2019 statt.

A3.1.1.3 Herstelleranfragen

Der Hersteller unterschrieb die Vereinbarung zur Regelung der vollständigen Informationsübermittlung.

Es wurden folgende relevante Studien beziehungsweise Dokumente übermittelt (Tabelle 7):

Tabelle 7: Durch Hersteller übermittelte relevante Studien bzw. Dokumente

Studie	Hersteller	Verfügbare Dokumente
Vergleich früh- vs. spätsymptomatischen Therapiebeginn		
ENDEAR (Subgruppen)	Biogen GmbH	ENDEAR-Studienbericht [25], Errata [83,84], Dokument [85], Zusatzanalysen [86] ^a
Vergleich prä- vs. frühsymptomatischen Therapiebeginn		
Biogen 2019	Biogen GmbH	Anhang zur Stellungnahme [27] ^{a, b}
a: Darüber hinaus wurden vom Hersteller weitere Dokumente zur Verfügung gestellt, aus denen keine weiteren Daten herangezogen wurden.		
b: Des Weiteren wurden Publikationen und Studienberichte zu den Studien ENDEAR [25,87] und NURTURE [28,29] zur Verfügung gestellt, die ergänzend herangezogen wurden, um z. B. Details zur Intervention darstellen zu können.		

Darüber hinaus wurde vom Hersteller 1 Studienbericht zu einer weiteren Studie (CS3A) zur Verfügung gestellt [88]. Bei dieser handelt es sich um eine Dosisesskalationsstudie, die nicht zur Beantwortung der Berichtsfragestellung beitragen kann.

Für die vorliegende Bewertung wurden zum einen für die Studie ENDEAR Anfragen zu Subgruppen an den Hersteller gestellt. Zum anderen wurde die Erstellung der Studie Biogen 2019 beim Hersteller angefragt. Eine Übersicht über alle Anfragen befindet sich in Tabelle 8. Die Informationen aus den eingegangenen Antworten sind in die Studienbewertung eingeflossen.

Tabelle 8: Übersicht über Herstelleranfragen

Studie	Inhalt der Anfrage	Antwort eingegangen ja / nein	Inhalt der Antwort
Vergleich früh- vs. spätsymptomatischen Therapiebeginn			
ENDEAR (Subgruppen)	<ul style="list-style-type: none"> ▪ Daten für die Subgruppen Krankheitsdauer ≤ 12 / > 12 Wochen: <ul style="list-style-type: none"> ▫ Patientencharakteristika ▫ Ergebnisse zu dem Endpunkt HINE-2 ▫ Ergebnisse zu (S)UEs ▫ Ergebnisse zu dem Endpunkt Tod ▫ Ergebnisse zu dem Endpunkt dauerhafte Beatmung 	ja	<ul style="list-style-type: none"> ▪ Daten für die Subgruppen Krankheitsdauer ≤ 12 / > 12 Wochen: <ul style="list-style-type: none"> ▫ Patientencharakteristika ▫ Ergebnisse zu dem Endpunkt HINE-2 ▫ Ergebnisse zu dem Endpunkt Tod ▫ Ergebnisse zu dem Endpunkt Zeit bis Tod oder dauerhafte Beatmung
	<ul style="list-style-type: none"> ▪ Nachforderung von unvollständigen Daten (Appendizes zum Studienbericht, Ergebnisse zu [S]UEs) 	ja	<ul style="list-style-type: none"> ▪ Daten zu den Endpunkten SUEs, schwere UEs, Therapieabbrüche wegen UEs
Vergleich prä- vs. frühsymptomatischen Therapiebeginn			
Biogen 2019	<ul style="list-style-type: none"> ▪ Daten zu dem Vergleich präsymptomatischer Therapiebeginn vs. Therapiebeginn ab Symptomen bei Kindern mit 2^a bzw. 3^b SMN2-Kopien 	ja	<ul style="list-style-type: none"> ▪ Ergebnisse für den Endpunkt Erreichen motorischer Meilensteine in Form einer Grafik für Kinder mit 2 und 3 SMN2-Kopien
	<ul style="list-style-type: none"> ▪ Nachforderung: Spezifizierung des Vergleichs^{a, b}, Tabellenvorlagen für Patientencharakteristika und Ergebnisse zu patientenrelevanten Endpunkten übermittelt 	ja	<ul style="list-style-type: none"> ▪ Daten in Tabellenvorlagen (im Rahmen der Stellungnahme zum Vorbericht)

(Fortsetzung)

Tabelle 8: Übersicht über Herstelleranfragen (Fortsetzung)

Studie	Inhalt der Anfrage	Antwort eingegangen ja / nein	Inhalt der Antwort
	<ul style="list-style-type: none"> Nachforderung: Abbildungen zum Erreichen des höchsten Levels der HINE-2-Einzelitems (1) „sitting pivots (rotates)“, (2) „Crawling on hands and knees“, (3) „stands unaided“ und (4) „walking independently“ in Analogie zu figure 3 (S. 46) in De Vivo 2019 [29] zum Erreichen der WHO-Meilensteine jeweils für Kinder mit 2 bzw. 3 <i>SMN2</i>-Kopien 		<ul style="list-style-type: none"> Abbildungen zum Erreichen des höchsten Levels der HINE-2-Einzelitems (1) „sitting pivots (rotates)“, (2) „Crawling on hands and knees“, (3) „stands unaided“ und (4) „walking independently“ (im Rahmen der Stellungnahme zum Vorbericht)
	<ul style="list-style-type: none"> Nachforderung: Informationen zu Geschwisterkindern (<i>SMA</i>-Typ, Anzahl an <i>SMN2</i>-Kopien, Alter bei Symptombeginn) 		<ul style="list-style-type: none"> Vertrauliche, patientenindividuelle Daten zur Verfügung gestellt
<p>a: Angefragt wurden Daten zu dem Vergleich präsymptomatischer Therapiebeginn versus Therapiebeginn ab Symptomen bei Kindern mit 2 <i>SMN2</i>-Kopien mit folgenden Eckpunkten:</p> <ul style="list-style-type: none"> Population der präsymptomatischen Kohorte: ausschließlich Kinder aus NURTURE, bei denen die Therapie präsymptomatisch begonnen wurde Population der Vergleichskohorte: Kinder aus ENDEAR (und ggf. vergleichbar geeigneten Studien), bei denen die Therapie mit Symptombeginn oder bis zu 12 Wochen danach begonnen wurde Ergebnisse für die Endpunkte Tod, dauerhafte Beatmung, schwere UEs und SUEs, Rate der Studienabbrucher aufgrund von UEs, Rückenschmerzen (als SUE und schweres UEs), HINE-2-Veränderung, jeweils zum Lebensalter von 1 Jahr (Unterschied zwischen den Kohorten bez. Alter maximal 10 %) und zum spätesten gemeinsamen Lebensalter (Unterschied zwischen den Kohorten bez. Alter maximal 10 %) <p>b: Angefragt wurden Daten zu dem Vergleich präsymptomatischer Therapiebeginn versus Therapiebeginn ab Symptomen bei Kindern mit 3 <i>SMN2</i>-Kopien mit folgenden Eckpunkten:</p> <ul style="list-style-type: none"> Population der präsymptomatischen Kohorte: ausschließlich Kinder aus NURTURE, bei denen die Therapie präsymptomatisch begonnen wurde Population der Vergleichskohorte: Kinder aus CHERISH (und ggf. vergleichbar geeigneten Studien), bei denen die Therapie mit Symptombeginn oder bis zu 14 Monate danach begonnen wurde Ergebnisse für die Endpunkte Tod, dauerhafte Beatmung, schwere UEs und SUEs, Rate der Studienabbrucher aufgrund von UEs, Rückenschmerzen (als SUE und schweres UE), HFMSE-Veränderung, jeweils zu einem möglichst frühen gemeinsamen Lebensalter (mindestens aber 2,5 Jahre; Unterschied zwischen den Kohorten bez. Alter maximal 10 %) und zum spätesten gemeinsamen Lebensalter (Unterschied zwischen den Kohorten bez. Alter maximal 10 %) <p>HFMSE: Hammersmith Functional Motor Scale Expanded; HINE-2: Hammersmith Infant Neurological Examination – Subscale 2; <i>SMN2</i>: Survival Motor Neuron 2; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis; vs.: versus</p>			

A3.1.2 Weitere Informationsquellen und Suchtechniken

Über weitere Informationsquellen und Suchtechniken identifizierte relevante Studien beziehungsweise Dokumente werden nachfolgend nur dargestellt, wenn sie nicht bereits über die primären Informationsquellen gefunden wurden.

A3.1.2.1 Zulassungsbehörden

Auf der Website der EMA und der FDA wurden keine relevanten Studien beziehungsweise Dokumente identifiziert, die nicht über andere Rechenschritte gefunden werden konnten.

A3.1.2.2 Durch den G-BA übermittelte Dokumente

Im Rahmen der Auftragsbearbeitung wurden Dokumente vom G-BA an das IQWiG weitergeleitet. Diese wurden auf Duplikate zur bibliografischen Recherche überprüft. Die im Rahmen der Volltextsichtung als nicht relevant ausgeschlossenen Dokumente finden sich mit Angabe des jeweiligen Ausschlussgrundes in Abschnitt A6.3.

Es wurden folgende relevante Studien beziehungsweise Dokumente identifiziert, die nicht über andere Rechenschritte gefunden werden konnten (Tabelle 9):

Tabelle 9: In vom G-BA übermittelten Dokumenten identifizierte relevante Studien bzw. Dokumente

Studie	Verfügbare Dokumente [Zitat]
Czibere/Vill 2019	Informationsblatt [58]

A3.1.2.3 G-BA-Website und IQWiG-Website

Auf der Website des G-BA wurde 1 relevantes Dokument identifiziert, das nicht über andere Rechenschritte gefunden werden konnte. Auf der Website des IQWiG wurden keine relevanten Studien beziehungsweise Dokumente identifiziert, die nicht über andere Rechenschritte gefunden werden konnten.

Tabelle 10: In IQWiG-Produkten, öffentlich zugänglichen Herstellerunterlagen oder Beschlüssen des G-BA identifizierte relevante Studien bzw. Dokumente

Studie	Verfügbare Dokumente [Zitat]
ENDEAR	Dossier Modul 4 [26]

A3.1.2.4 Anwendung weiterer Suchtechniken

Im Rahmen der Informationsbeschaffung wurden systematische Übersichten identifiziert – die entsprechenden Referenzen finden sich in Abschnitt A6.1.1. Die Referenzlisten dieser systematischen Übersichten wurden gesichtet.

Es fanden sich keine relevanten Studien beziehungsweise Dokumente, die nicht über andere Rechenschritte identifiziert werden konnten.

A3.1.2.5 Anhörung

Im Rahmen der Anhörung wurden keine relevanten Studien beziehungsweise Dokumente genannt, die nicht über andere Rechenschritte identifiziert werden konnten.

A3.1.2.6 Autorenanfragen

Autorenanfragen bezüglich zusätzlicher Informationen zu relevanten Studien waren nicht erforderlich, da davon auszugehen war, dass solche Informationen keinen relevanten Einfluss auf die Bewertung haben würden.

A3.1.3 Resultierender Studienpool

Durch die verschiedenen Rechenschritte konnten insgesamt 6 relevante Studien identifiziert werden (siehe auch Tabelle 11).

Tabelle 11: Studienpool der Nutzenbewertung

Studie	Verfügbare Dokumente		
	Vollpublikation (in Fachzeitschriften)	Registereintrag / Ergebnisbericht aus Studienregistern	Studienbericht und sonstige Dokumente aus Hersteller-unterlagen (nicht öffentlich zugänglich)
Vergleichende Interventionsstudien der Screeningkette			
keine Studien identifiziert			
Vergleichende Interventionsstudien zum Therapiebeginn			
Früh- versus spätsymptomatischen Therapiebeginn			
ENDEAR	ja [87]	ja [38,76-79] ^a / ja [80,81]	ja [25,83-86] [26] ^b
Prä- versus frühsymptomatischen Therapiebeginn			
Biogen 2019	nein	nein	ja [27] ^c
Studien zur diagnostischen Güte			
Chien 2017	ja [14]	ja [82] / nein	nein
Czibere/Vill 2019	ja [31,32] [58] ^d	nein / nein	nein
Kariyawasam 2019	ja [33]	nein / nein	nein
Kraszewski 2018	ja [34]	nein / nein	nein
a: In der laufenden Studie SHINE wurden u. a. Kinder, die in ENDEAR behandelt wurden, weiter beobachtet. Daher wurde der SHINE-Registereintrag [38] hier der Studie ENDEAR zugeordnet. b: öffentlich zugänglich auf der Website des G-BA c: Keine Vollpublikation für den angefragten Vergleich; auf Anfrage wurden Daten im Rahmen des Stellungnahmeverfahrens zum Vorbericht geliefert, die als Studie Biogen 2019 bezeichnet werden. Ergänzend wurden Publikationen und Studienberichte zu den Studien ENDEAR [25,87] und NURTURE [28,29] herangezogen, um z. B. Details zur Intervention darstellen zu können. d: keine Vollpublikation (Informationsblatt ohne Ergebnisdarstellung) G-BA: Gemeinsamer Bundesausschuss			

A3.1.4 Studien ohne berichtete Ergebnisse

In Tabelle 12 sind alle durch die Informationsbeschaffung identifizierten Studien ohne bisher berichtete Ergebnisse dargestellt.

Tabelle 12: In der Informationsbeschaffung identifizierte Studien ohne berichtete Ergebnisse

Studie	Dokumentart, ggf. Studienregister-ID [Zitat]	Studientyp	Geplante Fallzahl	Status (ggf. geplantes Studienende)
Vergleichende Interventionsstudien der Screeningkette				
keine Studien identifiziert				
Vergleichende Interventionsstudien zum Therapiebeginn				
keine Studien identifiziert				
Studien zur diagnostischen Güte				
NCT03217578	Studienregistereintrag, NCT03217578 [89]	prospektive diagnostische Kohortenstudie, VOPT	250 000	laufend (31.12.2019)
NCT03554343	Studienregistereintrag, NCT03554343 [90]; Designpublikation [91]	prospektive diagnostische Kohortenstudie, VOPT	180 000	laufend (31.07.2021)
NCT03655223	Studienregistereintrag, NCT03655223 [92]	prospektive diagnostische Kohortenstudie, VOPT	400 000	laufend (12/2022)
SMA: spinale Muskelatrophie; VOPT: Verification of only positive Testers				

A3.2 Vergleichende Interventionsstudien zur Screeningkette

Vergleichende Interventionsstudien der Screeningkette konnten nicht identifiziert werden. Daher erfolgte eine Bewertung der einzelnen Bausteine der Screeningkette – zum einen anhand von vergleichenden Interventionsstudien zum Therapiebeginn (Abschnitt A3.3), zum anderen anhand von Studien zur diagnostischen Güte (Abschnitt A3.4).

A3.3 Vergleichende Interventionsstudien zum Therapiebeginn

Für die Untersuchung des Vergleichs früherer versus späteren Therapiebeginn wurden Studien bzw. Daten aus Studien zu 2 Vergleichen herangezogen – zum einen zum Vergleich präsymptomatischer versus symptomatischen Therapiebeginn, zum anderen zum Vergleich frühsymptomatischer versus spätsymptomatischen Therapiebeginn. Für den Vergleich früherer versus späteren Therapiebeginn bei Symptomatik konnten Daten aus 1 RCT in die Nutzenbewertung eingeschlossen werden. Zum Vergleich eines präsymptomatischen versus einen frühsymptomatischen Therapiebeginn konnten Daten aus 1 retrospektiven vergleichenden Studie herangezogen werden.

A3.3.1 Charakteristika der in die Bewertung eingeschlossenen Studie

A3.3.1.1 Studiendesign und Studienpopulationen

A3.3.1.1.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Im Folgenden wird die Studie ENDEAR mit Daten zum Vergleich früh- versus spätsymptomatischen Therapiebeginn anhand der Studiencharakteristika (Tabelle 13), Details der Prüf- und Vergleichsintervention (Tabelle 14) sowie Ein- und Ausschlusskriterien (Tabelle 15) dargestellt. Eine Übersicht über wesentliche Patientencharakteristika erfolgt speziell für die

relevanten Subgruppen (Krankheitsdauer ≤ 12 Wochen oder > 12 Wochen) in Tabelle 16, da primär die Ergebnisse dieser Subgruppen für die vorliegende Nutzenbewertung relevant sind. Die Charakteristika für das Gesamtstudienkollektiv können öffentlich zugänglichen Dokumenten entnommen werden [26,87].

Tabelle 13: Charakterisierung der eingeschlossenen Studie (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Studiendesign	Patientenzahl (randomisiert) N	Intervention	Vergleich	Ort und Zeitraum der Durchführung	Relevante Endpunkte ^a
ENDEAR	RCT doppelblind multizentrisch (31 Zentren)	122	Nusinersen (intrathekale Injektion) + BSC	Scheinbehandlung (Sham-Injektion) + BSC	Nordamerika, Europa, Asien, Australien 08/2014 bis 11/2016	2 koprimary: Erreichen motorischer Meilensteine (HINE-2-Responder), Zeit bis Tod oder dauerhafte Beatmung ^b sekundäre: Gesamtüberleben, dauerhafte Beatmung ^b , (S)UEs
<p>a: Primäre Endpunkte beinhalten alle verfügbaren Angaben ohne Berücksichtigung der Relevanz für diese Nutzenbewertung. Sekundäre Endpunkte beinhalten ausschließlich Angaben zu für diese Nutzenbewertung relevanten verfügbaren Endpunkten.</p> <p>b: ≥ 16 Stunden Beatmung/Tag für > 21 aufeinanderfolgende Tage ohne akutes reversibles Ereignis oder Tracheotomie</p> <p>BSC: Best supportive Care; HINE-2: Hammersmith Infant Neurological Examination – Subscale 2; N: Patientenzahl; RCT: randomisierte kontrollierte Studie; (S)UE: (schwerwiegendes) unerwünschtes Ereignis</p>						

Tabelle 14: Charakterisierung der Interventionen in der eingeschlossenen Studie (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Intervention	Vergleich	Begleittherapie
ENDEAR	intrathekale Injektion von Nusinersen an Tag 1, 15, 29, 64, 183 und 302 Die Dosis von 12 mg bei einem Alter ≥ 2 Jahre wurde anhand des geschätzten altersentsprechenden CSF-Volumens angepasst, sodass kleineren Kindern eine entsprechend geringere Dosis appliziert wurde.	Scheinbehandlung: kleiner Einstich am unteren Rücken, der verbunden wurde an Tag 1, 15, 29, 64, 183 und 302	Best supportive Care; dies umfasste standardisierte medikamentöse und nicht medikamentöse supportive Maßnahmen, z. B. Behandlung von UEs, Maßnahmen im respiratorischen und gastrointestinalen Bereich sowie Ernährung
CSF: Cerebrospinalflüssigkeit; k. A.: keine Angabe; UE: unerwünschtes Ereignis			

Tabelle 15: Ein- / Ausschlusskriterien für Patientinnen und Patienten in der Studie (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
ENDEAR	<ul style="list-style-type: none"> ▪ Alter ≤ 7 Monate ▪ 5q-assoziierte spinale Muskelatrophie aufgrund einer Gendeletion oder Compound-heterozygoten Deletion oder Mutation im <i>SMN1</i>-Gen ▪ 2 Kopien des <i>SMN2</i>-Gens ▪ Krankheitsbeginn im Alter von ≤ 6 Monaten ▪ adäquater Ernährungsstatus ▪ Körpergewicht ≥ 3. Perzentil nach länderspezifischem Richtwert ▪ Gestationsalter 37 bis 42 Wochen 	<ul style="list-style-type: none"> ▪ Hypoxämie (Sauerstoffsättigung < 96 %, ohne Beatmung) ▪ SMA-Symptome bei Geburt oder innerhalb der 1. Lebenswoche ▪ Infekte mit notwendiger systemischer antiviraler oder antimikrobieller Therapie ▪ Hirn- oder Rückenmarkserkrankungen, z. B. Tumoren oder Erkrankungen mit Einfluss auf die Zirkulation der cerebrospinalen Flüssigkeit ▪ Implantat eines ZNS-Katheters ▪ medikamentöse Vorbehandlung der SMA
HFMSE: Hammersmith Functional Motor Scale Expanded; SMA: spinale Muskelatrophie; <i>SMN1</i> : Survival Motor Neuron 1; <i>SMN2</i> : Survival Motor Neuron 2; WHO: World Health Organization; ZNS: zentrales Nervensystem		

Tabelle 16: Charakterisierung der Studienpopulationen (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie Charakteristika	Krankheitsdauer ≤ 12 Wochen (frühsymptomatischer Therapiebeginn)		Krankheitsdauer > 12 Wochen (spätsymptomatischer Therapiebeginn)	
	Nusinersen	Scheinbehandlung	Nusinersen	Scheinbehandlung
ENDEAR				
N	34	18	46	23
Geschlecht [w], n (%)	18 (53)	7 (39)	25 (54)	17 (74)
Alter bei Symptombeginn [Wochen], Median (Spannweite)	6 (3–18)	8 (1–20)	8 (2–16)	8 (4–16)
Alter bei Diagnose [Wochen], Median (Spannweite)	9 (0–22)	10 (2–25)	12 (2–29)	20 (12–30)
Erkrankungsdauer [Wochen], Median (Spannweite) ^a	8 (0–12)	9 (0–12)	16 (12–26)	18 (13–23)
Alter bei 1. Dosis [Wochen], Median (Spannweite) ^b	16 (7–34)	19 (4–33)	28 (18–35)	30 (20–37)
a: Differenz zwischen Alter bei Studieneinschluss und Alter bei Symptombeginn				
b: In den Herstellerunterlagen finden sich Angaben in unterschiedlichen Zeiteinheiten. Zur Vereinheitlichung wurden alle in Wochen ausgedrückt.				
MW: Mittelwert; N: Anzahl eingeschlossener Kinder; n: Anzahl Kinder in Kategorie; w: weiblich				

A3.3.1.1.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Im Folgenden wird die Studie Biogen 2019 mit Daten zum Vergleich prä- versus frühsymptomatischen Therapiebeginn anhand der Studiencharakteristika (Tabelle 17), Details der Interventionen (Tabelle 18) sowie Ein- und Ausschlusskriterien (Tabelle 19) dargestellt. Eine Übersicht über wesentliche Patientencharakteristika erfolgt in Tabelle 20.

Tabelle 17: Charakterisierung der eingeschlossenen Studie (Vergleich prä- vs. fröhsymptomatischen Therapiebeginn)

Studie	Studien-design	Patientenzahl N	Präsymptomatischer Therapiebeginn	Fröhsymptomatischer Therapiebeginn	Ort und Zeitraum der Durchführung	Studiendauer	Relevante Endpunkte
Biogen 2019 ^a	retrospektive vergleichende Studie	49	Nusinersen (intrathekale Injektion) + BSC	Nusinersen (intrathekale Injektion) + BSC	präsymptomatischer Therapiebeginn: ▪ Nordamerika, Europa, Asien, Australien ▪ laufend seit 05/2015 symptomatischer Therapiebeginn: ▪ Nordamerika, Europa, Asien, Australien ▪ 08/2014 bis 11/2016	präsymptomatischer Therapiebeginn: ▪ Behandlungsdauer: bis zu 5 Jahre und anschließende Nachbeobachtungs- dauer: 3 Monate symptomatischer Therapiebeginn: ▪ geplante Behandlungsdauer: 10 Monate ^b ▪ Nachbeobachtungs- dauer: 3 Monate	Gesamtüberleben, dauerhafte Beatmung, Erreichen motorischer Meilensteine (HINE-2), (S)UE
<p>a: präsymptomatischer Therapiebeginn: Kinder aus der Studie NURTURE, die 2 <i>SMN2</i>-Kopien aufwiesen; fröhsymptomatischer Therapiebeginn: Kinder aus der Studie ENDEAR, die 2 <i>SMN2</i>-Kopien aufwiesen und deren Therapie mit Nusinersen innerhalb ≤ 12 Wochen nach Symptombeginn begann</p> <p>b: Die Studie wurde nach einer geplanten Interimsanalyse vorzeitig beendet aufgrund eines Wirksamkeitsnachweises von Nusinersen. Die mittlere Beobachtungsdauer zum finalen Datenschnitt am 16.12.2016 betrug 40 (Minimum 1; Maximum 63) Wochen. Die Kinder werden auf Wunsch in der Studie SHINE weiterbehandelt und -beobachtet.</p> <p>BSC: Best supportive Care; HINE-2: Hammersmith Infant Neurological Examination – Subscale 2; N: Patientenzahl; (S)UE: (schwerwiegendes) unerwünschtes Ereignis</p>							

Tabelle 18: Charakterisierung der Interventionen in der eingeschlossenen Studie (Vergleich prä- vs. fröhsymptomatischen Therapiebeginn)

Studie	Präsymptomatischer Therapiebeginn	Fröhsymptomatischer Therapiebeginn
Biogen 2019 ^a	<u>Therapieschema der Studie NURTURE:</u> intrathekale Injektion von Nusinersen an Tag 1, 15, 29, 64, 183 und dann alle 4 Monate altersunabhängige Dosis von 12 mg ^b	<u>Therapieschema der Studie ENDEAR:</u> intrathekale Injektion von Nusinersen an Tag 1, 15, 29, 64, 183 und 302 Die Dosis von 12 mg bei einem Alter ≥ 2 Jahre wurde anhand des geschätzten altersentsprechenden CSF-Volumens angepasst, sodass kleineren Kindern eine entsprechend geringere Dosis appliziert wurde.
	+ Best supportive Care; dies umfasste medikamentöse und nicht medikamentöse Maßnahmen. zur Behandlung von UEs oder zur Gewährleistung einer adäquaten supportiven Behandlung	+ Best supportive Care; dies umfasste standardisierte medikamentöse und nicht medikamentöse supportive Maßnahmen, z. B. Behandlung von UEs, Maßnahmen im respiratorischen und gastrointestinalen Bereich sowie Ernährung
a: präsymptomatischer Therapiebeginn: Kinder aus der Studie NURTURE, die 2 <i>SMN2</i> -Kopien aufwiesen; fröhsymptomatischer Therapiebeginn: Kinder aus der Studie ENDEAR, die 2 <i>SMN2</i> -Kopien aufwiesen und deren Therapie mit Nusinersen innerhalb ≤ 12 Wochen nach Symptombeginn begann b: Bis zum März 2017 (Protokollversion 6) wurde die Dosis anhand des geschätzten altersentsprechenden CSF-Volumens angepasst, sodass kleineren Kindern eine entsprechend geringere Dosis appliziert wurde.		
CSF: Cerebrospinalflüssigkeit; UE: unerwünschtes Ereignis		

Tabelle 19: Ein- / Ausschlusskriterien für Patientinnen und Patienten in der Studie (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Biogen 2019 ^a	Präsymptomatischer Therapiebeginn:	
	<u>Wesentliche Kriterien der Studie NURTURE:</u>	
	<ul style="list-style-type: none"> ▪ Alter ≤ 6 Wochen bei 1. Dosis ▪ 5q-assoziierte SMA aufgrund einer homozygoten Gendeletion oder Mutation oder Compound-heterozygoten Mutation ▪ 2 oder 3 Kopien des <i>SMN2</i>-Gens ▪ adäquater Ernährungsstatus ▪ Körpergewicht ≥ 3. Perzentil nach länderspezifischem Richtwert ▪ Gestationsalter 37 bis 42 Wochen bei Geburt (34 bis 42 Wochen bei Mehrlingsschwangerschaft) 	<ul style="list-style-type: none"> ▪ Hypoxämie (Sauerstoffsättigung < 96 %, ohne Beatmung) ▪ SMA-Symptome bei Screening oder unmittelbar vor der 1. Dosis ▪ Infekte mit notwendiger systemischer antiviraler oder antimikrobieller Therapie ▪ akutes Lungenversagen neonatal ▪ Hirn- oder Rückenmarkserkrankungen, die das Durchführen der Lumbalpunktionen, die Zirkulation der cerebrospinalen Flüssigkeit oder Beurteilung der Sicherheit beeinflusst hätten ▪ Implantat eines ZNS-Katheters oder eines Shunts zur Drainage der cerebrospinalen Flüssigkeit ▪ medikamentöse Vorbehandlung der SMA
	<u>Zusätzliche Kriterien für die Studie Biogen 2019:</u>	
<ul style="list-style-type: none"> ▪ Kinder mit 2 <i>SMN2</i>-Kopien 	-	
Frühsymptomatischer Therapiebeginn:		
<u>Wesentliche Kriterien der Studie ENDEAR:</u>		
<ul style="list-style-type: none"> ▪ Alter ≤ 7 Monate ▪ 5q-assoziierte spinale Muskelatrophie aufgrund einer Gendeletion oder Compound-heterozygoten Deletion oder Mutation im <i>SMN1</i>-Gen ▪ 2 Kopien des <i>SMN2</i>-Gens ▪ Krankheitsbeginn im Alter von ≤ 6 Monaten ▪ adäquater Ernährungsstatus ▪ Körpergewicht ≥ 3. Perzentil nach länderspezifischem Richtwert ▪ Gestationsalter 37 bis 42 Wochen 	<ul style="list-style-type: none"> ▪ Hypoxämie (Sauerstoffsättigung < 96 %, ohne Beatmung) ▪ SMA-Symptome bei Geburt oder innerhalb der 1. Lebenswoche ▪ Infekte mit notwendiger systemischer antiviraler oder antimikrobieller Therapie ▪ Hirn- oder Rückenmarkserkrankungen, z. B. Tumoren oder Erkrankungen mit Einfluss auf die Zirkulation der cerebrospinalen Flüssigkeit ▪ Implantat eines ZNS-Katheters ▪ medikamentöse Vorbehandlung der SMA 	
<u>Zusätzliche Kriterien für die Studie Biogen 2019:</u>		
<ul style="list-style-type: none"> ▪ Dauer zwischen Symptombeginn und Therapiebeginn ≤ 12 Wochen 	-	
<p>a: präsymptomatischer Therapiebeginn: Kinder aus der Studie NURTURE, die 2 <i>SMN2</i>-Kopien aufwiesen; frühsymptomatischer Therapiebeginn: Kinder aus der Studie ENDEAR, die 2 <i>SMN2</i>-Kopien aufwiesen und deren Therapie mit Nusinersen innerhalb ≤ 12 Wochen nach Symptombeginn begann</p> <p>SMA: spinale Muskelatrophie; <i>SMN1</i>: Survival Motor Neuron 1; <i>SMN2</i>: Survival Motor Neuron 2; ZNS: zentrales Nervensystem</p>		

Tabelle 20: Charakterisierung der Studienpopulation (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)

Studie	Kohorte	Präsymptomatischer Therapiebeginn	Frühsymptomatischer Therapiebeginn
Charakteristika			
Biogen 2019^a			
N		15	34
Geschlecht [w / m], %		47 / 53 ^b	53 / 47 ^b
Alter bei Studieneinschluss [Wochen], MW (SD) ^c		2 (1)	16 (6)
Alter bei Diagnose [Wochen], MW (SD)		2 (1) ^{c, d}	11 (5) ^e
Alter bei Symptombeginn [Wochen], MW (SD)		- ^f	8 (4) ^e
Alter bei Therapiebeginn [Wochen], MW (SD) ^c		3 (1)	18 (6)
Alter bei letzter Therapiegabe [Wochen], MW (SD) ^c		111 (28)	46 (14)
Anzahl der Therapiegaben, MW (SD)		9,80 (1,61)	4,94 (1,2)
Alter bei letztem Follow-up [Wochen], MW (SD) ^c		110 (24)	57 (16)
Therapieabbrucher ^g , n (%)		0 (0)	0 (0) ^h
<p>a: präsymptomatischer Therapiebeginn: Kinder aus der Studie NURTURE, die 2 <i>SMN2</i>-Kopien aufwiesen; frühsymptomatischer Therapiebeginn: Kinder aus der Studie ENDEAR, die 2 <i>SMN2</i>-Kopien aufwiesen und deren Therapie mit Nusinersen innerhalb ≤ 12 Wochen nach Symptombeginn begann</p> <p>b: eigene Berechnung</p> <p>c: In den Herstellerunterlagen finden sich Angaben in unterschiedlichen Zeiteinheiten. Zur Vereinheitlichung wurden alle in Wochen ausgedrückt.</p> <p>d: Alter bei Diagnose basiert bei präsymptomatischen Kindern auf dem Datum bei Studieneinschluss.</p> <p>e: Laut Herstellerunterlagen ist das Alter in Monaten angegeben. Aus Plausibilitätsgründen (vorliegende Daten zur ENDEAR-Studie) wird davon ausgegangen, dass es sich stattdessen um Wochen handelt.</p> <p>f: Alter bei Symptombeginn wurde im Rahmen der Studie NURTURE nicht zu Baseline erhoben.</p> <p>g: ohne Kinder, die im Studienverlauf verstarben</p> <p>h: Die Studie ENDEAR wurde aufgrund eines frühzeitigen Wirksamkeitsnachweises abgebrochen. Dadurch konnte bei 18 Kindern (53 %) die vorgeschriebene Studienmedikation nicht zu Ende gegeben werden.</p> <p>m: männlich; MW: Mittelwert; N: Anzahl eingeschlossener Kinder; n: Anzahl Kinder in Kategorie; SD: Standardabweichung; w: weiblich</p>			

A3.3.1.2 Bewertung endpunktübergreifender Kriterien des Verzerrungspotenzials

A3.3.1.2.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Die Einschätzung endpunktübergreifender Kriterien des Verzerrungspotenzials der Studie zum früh- versus spätsymptomatischen Therapiebeginn ist in der folgenden Tabelle 21 dargestellt.

Tabelle 21: Endpunktübergreifendes Verzerrungspotenzial (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktübergreifendes Verzerrungspotenzial
			Patientin / Patient	Behandelnde Personen			
ENDEAR	ja	ja	ja	nein ^a	unklar ^b	ja	niedrig
a: Die Therapie wurde unverblindet verabreicht. Keine Angaben zu weiterbehandelnden Personen b: Anhänge und Kapitel des Studienberichts wurden nicht übermittelt.							

A3.3.1.2.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Auf eine tabellarische Einschätzung des Verzerrungspotenzials anhand verschiedener Kriterien wurde bei dieser Teil-Fragestellung verzichtet, da ausschließlich Daten aus 1 retrospektiven vergleichenden Studie herangezogen werden konnten. Aufgrund des Studiendesigns ist das Verzerrungspotenzial endpunktübergreifend als hoch anzusehen.

A3.3.2 Patientenrelevante Endpunkte

Im Folgenden werden für den Vergleich frühsymptomatischer versus spätsymptomatischen Therapiebeginn je Endpunkt jeweils die Bewertung des endpunktspezifischen Verzerrungspotenzials der Ergebnisse sowie die Ergebnisse der Subgruppen tabellarisch dargestellt. Für den Vergleich präsymptomatischer versus frühsymptomatischen Therapiebeginn erfolgt ausschließlich eine tabellarische Darstellung der Ergebnisse. Das endpunktspezifische Verzerrungspotenzial wurde für alle berichteten Endpunkte aufgrund des retrospektiven Studiendesigns und damit einhergehend eines hoch einzuschätzenden, endpunktübergreifenden Verzerrungspotenzials ebenfalls als hoch bewertet (siehe Abschnitt A3.3.1.2).

A3.3.2.1 Kombiniertes Endpunkt Zeit bis Tod oder dauerhafte Beatmung

A3.3.2.1.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Der kombinierte Endpunkt Zeit bis Tod oder dauerhafte Beatmung umfasste die Ereignisse Tod und eine Beatmung über ≥ 16 Stunden pro Tag durchgehend für > 21 Tage bei gleichzeitiger Abwesenheit von akuten reversiblen Ereignissen oder Tracheotomie. Akute reversible

Ereignisse waren prädefiniert und jeder Fall wurde durch ein verblindetes, zentrales und unabhängiges Gremium (Endpoint Adjudication Committee) geprüft.

Verzerrungspotenzial der Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung

Tabelle 22: Bewertung des endpunktspezifischen Verzerrungspotenzials: kombinierter Endpunkt Zeit bis Tod oder dauerhafte Beatmung (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Endpunkübergreifendes Verzerrungspotenzial	Verblindung Endpunkterheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial
ENDEAR	niedrig	ja	ja	ja	nein ^a	hoch

a: keine Angaben zur Verteilung der Begleitbehandlung in den Subgruppen; aus den Angaben für die Gesamtpopulation geht hervor, dass vielfältige Begleitbehandlungen erfolgten
ITT: Intention to treat

Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung

Tabelle 23: Ergebnisse – kombinierter Endpunkt Zeit bis Tod oder dauerhafte Beatmung (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Subgruppe	Krankheitsdauer ≤ 12 Wochen (frühsymptomatischer Therapiebeginn)				Krankheitsdauer > 12 Wochen (spätsymptomatischer Therapiebeginn)				Interaktionstest		
	Nusinersen		Scheinbehandlung		Nusinersen vs. Scheinbehandlung		Nusinersen vs. Scheinbehandlung				
Studie	N	10 %-Quantil der Zeit bis zum Ereignis in Wochen [95 %-KI] ^a	N	10 %-Quantil der Zeit bis zum Ereignis in Wochen [95 %-KI] ^a	HR [95 %-KI]; p-Wert	N	10 %-Quantil der Zeit bis zum Ereignis in Wochen [95 %-KI] ^a	N	10 %-Quantil der Zeit bis zum Ereignis in Wochen [95 %-KI] ^a	HR [95 %-KI]; p-Wert	p-Wert
Endpunkt	Kinder mit Ereignis n (%)		Kinder mit Ereignis n (%)			Kinder mit Ereignis n (%)		Kinder mit Ereignis n (%)			
ENDEAR											
kombinierter Endpunkt	34	29,0 [k. A.] ^b 6 (18)	18	6,0 [k. A.] ^b 12 (67)	0,16 [0,06; 0,44]; < 0,001 ^{b, c}	46	7,9 [k. A.] ^b 25 (54)	23	4,3 [k. A.] ^b 16 (70)	0,82 [0,43; 1,54]; 0,533 ^{b, c}	0,008 ^d
Komponente Zeit bis Tod ^{b, e}	34	n. e. 3 (9)	18	10 [k. A.] 7 (39)	0,22 [0,06; 0,86]; 0,030 ^c	46	12,0 [k. A.] 10 (22)	23	5,1 [k. A.] 9 (39)	0,46 [0,18; 1,12]; 0,088 ^c	0,383 ^d
Komponente Zeit bis dauerhafte Beatmung ^e	34	k. A. 3 (9)	18	k. A. 6 (33)	0,16 [0,04; 0,63]; 0,009 ^f	46	k. A. 15 (33)	23	k. A. 7 (30)	1,00 [0,41; 2,45]; 0,997 ^f	0,040
<p>a: Die mediane Zeit und das 25 %-Quantil der Zeit bis zum Ereignis wurden in mindestens 1 Behandlungsgruppe nicht erreicht, daher wurden die Ergebnisse zum 10 %-Quantil dargestellt.</p> <p>b: Die Daten stammen aus unveröffentlichten Dokumenten des Herstellers.</p> <p>c: Cox-Regression adjustiert für die Krankheitsdauer zum Studieneinschluss</p> <p>d: eigene Berechnung</p> <p>e: Gesamtzahl der Ereignisse im Studienverlauf</p> <p>f: Cox-Regression</p> <p>HR: Hazard Ratio; KI: Konfidenzintervall; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; n. e.: nicht erreicht; vs.: versus</p>											

Der Interaktionstest zeigt eine statistisch signifikante Interaktion zwischen den Subgruppen früh- und spätsymptomatischer Therapiebeginn (Krankheitsdauer ≤ 12 / > 12 Wochen) bei Kindern mit infantiler SMA ($p = 0,008$). Der Effekt zugunsten von Nusinersen beim früh-symptomatischen Therapiebeginn (HR: 0,16; 95 %-KI: [0,06; 0,44]; $p < 0,001$) war dabei statistisch signifikant im Gegensatz zum nicht statistisch signifikanten Unterschied beim spätsymptomatischen Therapiebeginn (HR: 0,82; 95 %-KI: [0,43; 1,54]; $p = 0,533$). Die HRs der Einzelkomponenten und des kombinierten Endpunkts wurden in den Dokumenten unterschiedlich berechnet. Die Ergebnisse der Einzelkomponenten zeigen, dass der Effekt des kombinierten Endpunkts nicht nur auf 1 der Einzelkomponenten basiert.

A3.3.2.1.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Der kombinierte Endpunkt Zeit bis Tod oder dauerhafte Beatmung wurde beim Hersteller nicht angefragt. Es lagen somit keine Daten vor.

A3.3.2.2 Gesamtüberleben

A3.3.2.2.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Für den Endpunkt Gesamtüberleben lagen verwertbare Ergebnisse aus der Studie ENDEAR für die berichtsrelevanten Subgruppen vor. Dieser Endpunkt wird als Teilkomponente des kombinierten Endpunkts Zeit bis Tod oder dauerhafte Beatmung berücksichtigt und dargestellt (siehe Abschnitt A3.3.2.1).

A3.3.2.2.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Im Folgenden werden die Ergebnisse zu dem Endpunkt Gesamtüberleben dargestellt.

Tabelle 24: Ergebnisse – Mortalität (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)

Kohorte	Präsymptomatischer Therapiebeginn		Frühsymptomatischer Therapiebeginn		Prä- vs. frühsymptomatisch		
	N	Mediane Zeit bis zum Ereignis in Monaten [95 %-KI]; Kinder mit Ereignis n (%)	N	Mediane Zeit bis zum Ereignis in Monaten [95 %-KI]; Kinder mit Ereignis n (%)	HR ^a	[95 %-KI]	p-Wert
Biogen 2019^b							
Im Lebensalter von 1 Jahr ^c	15	n. e. [n. e.; n. e.]; 0 (0)	34	n. e. [n. e.; n. e.]; 3 (9)	0	[0; ∞)	> 0,999
a: Cox-Modell							
b: präsymptomatischer Therapiebeginn: Kinder aus der Studie NURTURE, die 2 <i>SMN2</i> -Kopien aufwiesen; frühsymptomatischer Therapiebeginn: Kinder aus der Studie ENDEAR, die 2 <i>SMN2</i> -Kopien aufwiesen und deren Therapie mit Nusinersen innerhalb ≤ 12 Wochen nach Symptombeginn begann							
c: Unterschied in den Erhebungszeitpunkten zwischen den Kohorten maximal 10 %							
HR: Hazard Ratio; KI: Konfidenzintervall; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; n. e.: nicht erreicht							

A3.3.2.3 Krankheitsspezifisches Überleben

A3.3.2.3.1 Frühsymptomatischer versus spät-symptomatischen Therapiebeginn

Zum Endpunkt krankheitsspezifisches Überleben lagen keine Ergebnisse für die berichtsrelevanten Subgruppen zur Krankheitsdauer vor.

A3.3.2.3.2 Präsymptomatischer versus früh-symptomatischen Therapiebeginn

Zum Endpunkt krankheitsspezifisches Überleben erfolgte beim Hersteller keine Anfrage, da aus den zugrunde liegenden Studien ENDEAR und NURTURE keine entsprechenden Ergebnisse zu erwarten waren.

A3.3.2.4 Erreichen motorischer Meilensteine

A3.3.2.4.1 Frühsymptomatischer versus spät-symptomatischen Therapiebeginn

Zu dem Endpunkt Erreichen motorischer Meilensteine wurde von den 2 in der Studie verwendeten Instrumenten der HINE-2 in Form von Responderanalysen in der vorliegenden Nutzenbewertung herangezogen. Die Definition der Gesamtscore-Responder basierte auf 7 von 8 einzelnen Meilensteinkategorien des HINE-2, die jeweils anhand von Skalen von 3 bis 5 möglichen Entwicklungsstufen gemessen werden. Das bewusste Greifen wurde von der Studiengruppe aus den Analysen ausgeschlossen, da dieser motorische Meilenstein auch von Kindern mit SMA ohne Behandlung zu großen Teilen erreicht werden kann. Als Gesamtscore-Responder wurden Kinder gewertet, die die folgenden Kriterien erfüllten: (1) ≥ 1 -Punkt-Verbesserung in den Kategorien Kopfkontrolle, Drehen, Sitzen, Robben / Krabbeln, Stehen oder Laufen oder ≥ 2 -Punkt-Verbesserung in der Kategorie Strampeln und / oder Erreichen des Maximalscores in der Kategorie Strampeln, (2) mehr Kategorien mit einer Verbesserung als Kategorien mit einer Verschlechterung. Verstorbene Kinder sowie Studienabbrecher wurden dabei als Non-Responder gewertet.

Diese Form der Gesamtscore-Responder-Auswertung erlaubt keine eindeutige Interpretation der Ergebnisse, da Kinder, die als Responder definiert waren, zeitgleich Verbesserungen und Verschlechterungen in den einzelnen Kategorien erfahren konnten und hier zudem die Art der in den Kategorien untersuchten motorischen Fähigkeiten von unterschiedlich hoher Bedeutung für die Entwicklung sein kann. Gemeinsam mit den Ergebnissen zu den einzelnen motorischen Meilensteinkategorien (Items) konnten diese Daten jedoch interpretiert werden (siehe Abschnitt 4.3.4.4 und Tabelle 26).

Responder bei den einzelnen Items waren analog zu den Einzelteilen beim Gesamtscore definiert (≥ 1 -Punkt-Verbesserung in allen Kategorien mit Ausnahme der Kategorie Strampeln mit ≥ 2 -Punkt-Verbesserung; verstorbene Kinder sowie Studienabbrecher wurden als Non-Responder gewertet).

Verzerrungspotenzial der Ergebnisse zum Endpunkt Erreichen motorischer Meilensteine

Tabelle 25: Bewertung des endpunktspezifischen Verzerrungspotenzials: Erreichen motorischer Meilensteine (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Endpunkübergreifendes Verzerrungspotenzial	Verblindung Endpunkterheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial
ENDEAR	niedrig	ja	ja	ja	nein ^a	hoch
<p>a: keine Angaben zur Verteilung der Begleitbehandlung in den Subgruppen; aus den Angaben für die Gesamtpopulation geht hervor, dass vielfältige Begleitbehandlungen erfolgten</p> <p>ITT: Intention to treat</p>						

Ergebnisse zum Endpunkt Erreichen motorischer Meilensteine

Tabelle 26: Ergebnisse – Responder Erreichen motorischer Meilensteine HINE-2 (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Subgruppe	Erkrankungsdauer ≤ 12 Wochen (frühsymptomatischer Therapiebeginn)							Erkrankungsdauer > 12 Wochen (spätsymptomatischer Therapiebeginn)						Interaktionstest p-Wert	
	Nusinersen		Scheinbehandlung		Scheinbehandlung vs. Nusinersen			Nusinersen		Scheinbehandlung		Scheinbehandlung vs. Nusinersen			
	N	n (%)	N	n (%)	RD ^a	[95 %-KI] ^b	p-Wert ^c	N	n (%)	N	n (%)	RD ^a	[95 %-KI] ^b		p-Wert ^c
ENDEAR															
HINE-2 gesamt	34	24 (71)	18	0 (0)	0,71	[0,55; 0,86]	< 0,001	46	17 (37)	23	0 (0)	0,37	[0,23; 0,51]	< 0,001	0,003
Strampeln ^d	32 ^e	10 (31)	16 ^e	0 (0)	–	–	–	41 ^e	2 (5)	21 ^e	0 (0)	–	–	–	–
Kopfkontrolle ^f	32 ^e	21 (66)	16 ^e	0 (0)	–	–	–	41 ^e	12 (29)	21 ^e	0 (0)	–	–	–	–
Drehen ^f	32 ^e	18 (56)	16 ^e	0 (0)	–	–	–	41 ^e	7 (17)	21 ^e	1 (5)	–	–	–	–
Sitzen ^f	32 ^e	16 (50)	16 ^e	0 (0)	–	–	–	41 ^e	5 (12)	21 ^e	0 (0)	–	–	–	–
Krabbeln ^f	32 ^e	6 (19)	16 ^e	0 (0)	–	–	–	41 ^e	2 (5)	21 ^e	0 (0)	–	–	–	–
Stehen ^f	32 ^e	5 (16)	16 ^e	0 (0)	–	–	–	41 ^e	2 (5)	21 ^e	0 (0)	–	–	–	–
Gehen ^f	32 ^e	0 (0)	16 ^e	0 (0)	–	–	–	41 ^e	0 (0)	21 ^e	0 (0)	–	–	–	–
<p>a: Da keine Ereignisse in der Kontrollgruppe auftraten, wurde auf eine Berechnung eines relativen Effektmaßes (OR) verzichtet und stattdessen die Risikodifferenz (RD) herangezogen; eigene Berechnung.</p> <p>b: eigene Berechnung, asymptotisch</p> <p>c: eigene Berechnung (unbedingter exakter Test [CSZ-Methode nach [93]])</p> <p>d: Responder: Verbesserung um ≥ 2 Punkte auf der Skala von insgesamt 5 Entwicklungsstufen; verstorbene Kinder sowie Studienabbrecher wurden als Non-Responder bewertet</p> <p>e: Anzahl ausgewerteter Kinder im Efficacy Set mit mindestens 6-monatiger (Studientag 183) Nachbeobachtung</p> <p>f: Responder: Verbesserung um ≥ 1 Punkt auf der Skala von insgesamt 3 bis 5 Entwicklungsstufen; verstorbene Kinder sowie Studienabbrecher wurden als Non-Responder bewertet</p> <p>HINE-2: Hammersmith Infant Neurological Examination – Subscale 2; KI: Konfidenzintervall; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; RD: Risikodifferenz; vs.: versus</p>															

A3.3.2.4.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Im Folgenden werden die Ergebnisse zu dem Endpunkt Erreichen motorischer Meilensteine dargestellt.

Tabelle 27: Ergebnisse – HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)

Kohorte Zeitpunkt (Lebensalter)	Präsymptomatischer Therapiebeginn		Frühsymptomatischer Therapiebeginn		Prä- vs. frühsymptomatisch				
	N ^a	Messwerte zum Zeitpunkt		N ^a	Messwerte zum Zeitpunkt		MD ^b	[95 %-KI]	p-Wert
		MW	SD		MW	SD			
Biogen 2019^c									
Zu Studienbeginn	15	2,67	1,59	34	1,15	1,10	1,52	[0,58; 2,46]	0,003
Im Lebensalter von 1 Jahr ^d	15	19,67	4,15	14	5,86	3,16	13,81	[11,01; 16,61]	< 0,001
Hedges' g ^e :							3,62	[2,38; 4,86]	
<p>a: Anzahl der Kinder, die in der Auswertung zur Berechnung des Effektschätzers berücksichtigt wurden. Die Angaben zu Studienende und -beginn (Baseline) können auf anderen Anzahlen basieren.</p> <p>b: ITT-Auswertung</p> <p>c: präsymptomatischer Therapiebeginn: Kinder aus der Studie NURTURE, die 2 <i>SMN2</i>-Kopien aufwiesen; frühsymptomatischer Therapiebeginn: Kinder aus der Studie ENDEAR, die 2 <i>SMN2</i>-Kopien aufwiesen und deren Therapie mit Nusinersen innerhalb ≤ 12 Wochen nach Symptombeginn begann</p> <p>d: Unterschied in den Erhebungszeitpunkten zwischen den Kohorten maximal 10 %</p> <p>e: eigene Berechnung</p> <p>HINE-2: Hammersmith Infant Neurological Examination – Subscale 2; k. A.: keine Angabe; KI: Konfidenzintervall; MW: Mittelwert; MD: Mittelwertdifferenz; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; n. b.: nicht berechnet; SD: Standardabweichung</p>									

Zu den HINE-2-Einzelitems Drehen, Krabbeln, Stehen und Gehen wurden beim Hersteller Abbildungen zum Erreichen des höchsten Levels („dreht sich im Sitzen“, „krabbelt auf Händen und Knien“, „steht ohne Hilfe“ und „geht selbstständig“) angefragt. Abbildung 3, Abbildung 4, Abbildung 5 und Abbildung 6 zeigen die Anzahl der Kinder, die den Maximalscore des jeweiligen Einzelitems des HINE-2 erreichen beziehungsweise nicht erreichen.

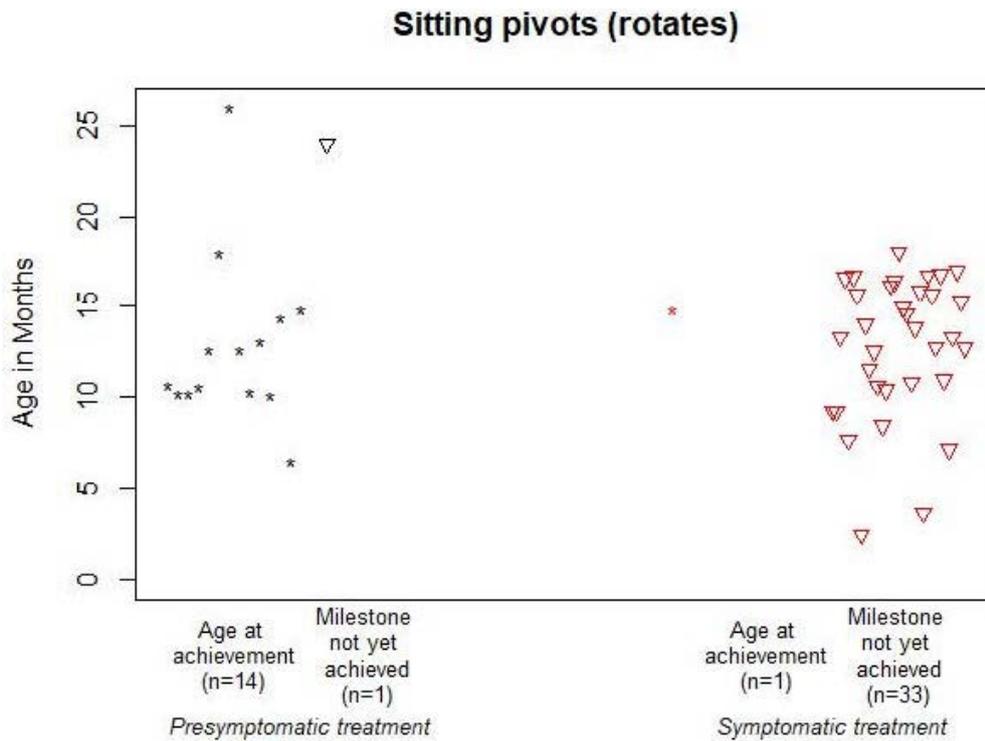


Abbildung 3: Responder Erreichen des Maximalscores beim Einzelitem Drehen („dreht sich im Sitzen“) des HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn [27])

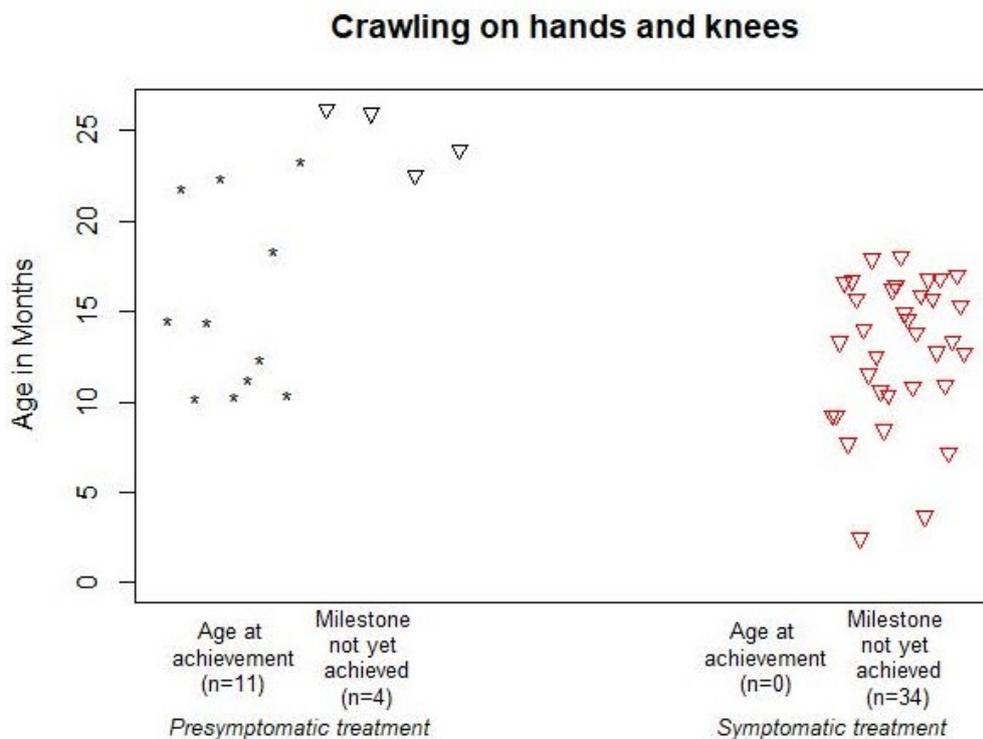


Abbildung 4: Responder Erreichen des Maximalscores beim Einzelitem Krabbeln („krabbelt auf Händen und Knien“) des HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn [27])

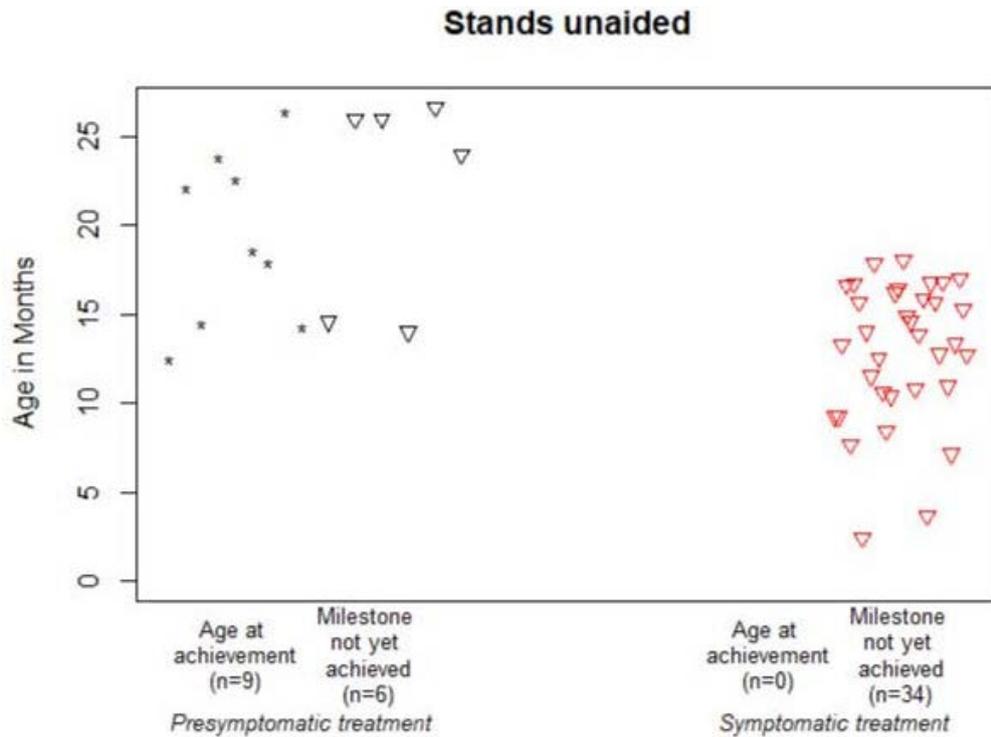


Abbildung 5: Responder Erreichen des Maximalscores beim Einzelitem Stehen („steht ohne Hilfe“) des HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn [27])

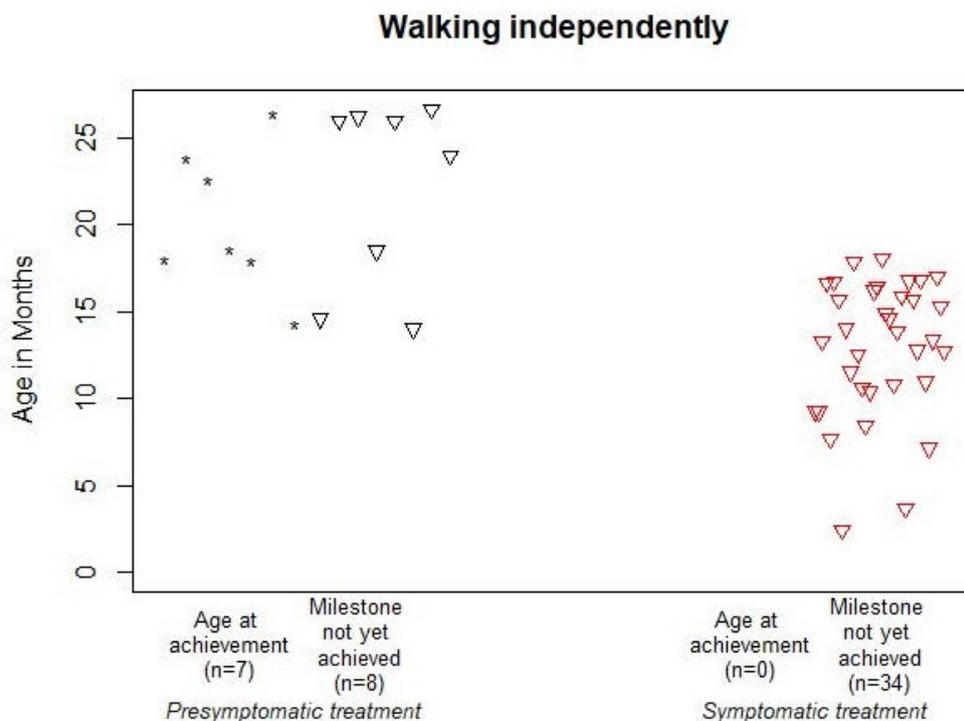


Abbildung 6: Responder Erreichen des Maximalscores beim Einzelitem Gehen („geht selbstständig“) des HINE-2 (Vergleich prä- vs. frühsymptomatischen Therapiebeginn [27])

A3.3.2.5 Entwicklungs- und Wachstumsstörungen

A3.3.2.5.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Neben dem Endpunkt Erreichen motorischer Meilensteine wurden in der Studie ENDEAR weitere Entwicklungs- und Wachstumsstörungen mittels HINE (Subskalen 1 und 3) erhoben. Für die berichtsrelevanten Subgruppen zur Krankheitsdauer lagen keine verwertbaren Daten vor.

A3.3.2.5.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Der Endpunkt Entwicklungs- und Wachstumsstörungen wurde beim Hersteller nicht angefragt. Es lagen somit keine Daten vor.

A3.3.2.6 Krankenhausaufenthalte

A3.3.2.6.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Für den Endpunkt Krankenhausaufenthalte lagen Daten für die berichtsrelevanten Subgruppen vor [26]. Berichtet wurde die Rate von Hospitalisierungen pro Jahr. Dazu wurde die Gesamtzahl der Ereignisse in Bezug zu den Patientenjahren mittels Rate Ratio verglichen, ohne die Abhängigkeit von wiederholten Ereignissen, wie z. B. mehrfachen Krankenhausaufenthalten eines Kindes zu berücksichtigen. Diese Auswertung wurde daher nicht für die vorliegende Nutzenbewertung herangezogen. Angaben, für wie viele Kinder in den jeweiligen Subgruppen überhaupt eine Hospitalisierung erfolgte, lagen nicht vor.

A3.3.2.6.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Der Endpunkt Krankenhausaufenthalte wurde beim Hersteller nicht angefragt. Es lagen somit keine Daten vor.

A3.3.2.7 Atemnot

A3.3.2.7.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Atemnot wurde im Rahmen der Studie als UE (Preferred Term) erfasst. Für die berichtsrelevanten Subgruppen zur Krankheitsdauer lagen keine Daten vor.

A3.3.2.7.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Der Endpunkt Atemnot wurde beim Hersteller nicht angefragt. Es lagen somit keine Daten vor.

A3.3.2.8 Dauerhafte Beatmung

A3.3.2.8.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Der Endpunkt Beatmung wurde in der Studie über den kombinierten Endpunkt hinaus in 5 unterschiedlichen Operationalisierungen erhoben, darunter 2 zur dauerhaften Beatmung. Für die berichtsrelevanten Subgruppen lagen Ergebnisse zu beiden Operationalisierungen zur dauerhaften Beatmung vor. Herangezogen wurden die Daten zur Operationalisierung Zeit bis

zur dauerhaften Beatmung, die als Teilkomponente des kombinierten Endpunkts Zeit bis Tod oder dauerhafte Beatmung berücksichtigt wurde. Die Ergebnisse der Teilkomponente finden sich in Tabelle 23 in Abschnitt A3.3.2.1.

A3.3.2.8.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Im Folgenden werden die Ergebnisse zu dem Endpunkt dauerhafte Beatmung dargestellt.

Tabelle 28: Ergebnisse – dauerhafte Beatmung (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)

Kohorte	Präsymptomatischer Therapiebeginn		Frühsymptomatischer Therapiebeginn		Prä- vs. frühsymptomatisch		
	N	Mediane Zeit bis zum Ereignis in Monaten [95 %-KI]; Kinder mit Ereignis n (%)	N	Mediane Zeit bis zum Ereignis in Monaten [95 %-KI]; Kinder mit Ereignis n (%)	HR ^a	[95 %-KI]	p-Wert
Biogen 2019^b							
Im Lebensalter von 1 Jahr ^c	15	n. e. [n. e.; n. e.]; 0 (0)	34	n. e. [n. e.; n. e.]; 2 (6)	0	[0; ∞)	> 0,999
a: Cox-Modell b: präsymptomatischer Therapiebeginn: Kinder aus der Studie NURTURE, die 2 <i>SMN2</i> -Kopien aufwiesen; frühsymptomatischer Therapiebeginn: Kinder aus der Studie ENDEAR, die 2 <i>SMN2</i> -Kopien aufwiesen und deren Therapie mit Nusinersen innerhalb ≤ 12 Wochen nach Symptombeginn begann c: Unterschied in den Erhebungszeitpunkten zwischen den Kohorten maximal 10 % HR: Hazard Ratio; KI: Konfidenzintervall; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; n. b.: nicht berechnet; n. e.: nicht erreicht							

A3.3.2.9 Schwerwiegende respiratorische Ereignisse

A3.3.2.9.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Der Endpunkt schwerwiegende respiratorische Ereignisse wurde in der Studie ENDEAR erfasst und war definiert als alle UEs, die während des gesamten Studienzeitraums in die Systemorganklasse (SOC) der Erkrankungen der Atemwege, des Brustraums und Mediastinums als primäre SOC oder sekundäre SOC und als schwerwiegend eingestuft wurden. Seltene Ereignisse der SOC Infektionen und parasitäre Erkrankungen sowie Untersuchungen wurden dabei nicht berücksichtigt.

Für die berichtsrelevanten Subgruppen zur Krankheitsdauer lagen Daten lediglich in Form des p-Werts des Interaktionstests zum Effektmaß Rate Ratio vor. Dazu wurde die Gesamtzahl der Ereignisse in Bezug zu den Patientenjahren mittels Rate Ratio verglichen, ohne die Abhängigkeit von Ereignissen innerhalb eines Kindes zu berücksichtigen. Angaben, wie viele Kinder in den jeweiligen Subgruppen ein oder mehrere Ereignisse hatten, lagen nicht vor. Diese Auswertung wurde daher nicht für die vorliegende Nutzenbewertung herangezogen.

A3.3.2.9.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Der Endpunkt schwerwiegende respiratorische Ereignisse wurde beim Hersteller nicht angefragt. Es lagen somit keine Daten vor.

A3.3.2.10 (Schwerwiegende) unerwünschte Ereignisse

A3.3.2.10.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Im Folgenden werden die Ergebnisse zu den Endpunkten SUEs, schwere UEs sowie Therapieabbruch wegen UEs dargestellt.

Schwerwiegende unerwünschte Ereignisse

Verzerrungspotenzial zu SUEs

Tabelle 29: Bewertung des endpunktspezifischen Verzerrungspotenzials: schwerwiegende unerwünschte Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Endpunkübergreifendes Verzerrungspotenzial	Verblindung Endpunkterheber	ITT-Prinzip adäquat umgesetzt	Ergebnisunabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunktspezifisches Verzerrungspotenzial
ENDEAR	niedrig	ja	ja	ja	nein ^a	hoch
a: Keine Angaben zur Verteilung der Begleitbehandlung in den Subgruppen; aus den Angaben für die Gesamtpopulation geht hervor, dass vielfältige Begleitbehandlungen erfolgten. ITT: Intention to treat						

Ergebnisse zu SUEs

Tabelle 30: Ergebnisse – schwerwiegende unerwünschte Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Subgruppe	Erkrankungsdauer ≤ 12 Wochen (frühsymptomatischer Therapiebeginn)							Erkrankungsdauer > 12 Wochen (spätsymptomatischer Therapiebeginn)							Inter- aktions- test
	Nusinersen		Scheinbehand- lung		Scheinbehandlung vs. Nusinersen			Nusinersen		Schein- behandlung		Scheinbehandlung vs. Nusinersen			
	N	n (%)	N	n (%)	OR	[95 %-KI]	p-Wert	N	n (%)	N	n (%)	OR	[95 %-KI]	p-Wert	
ENDEAR															
mindestens 1 SUE	34	25 (74)	18	17 (94)	0,16	[0,02; 1,41]	0,100	46	36 (78)	23	22 (96)	0,16	[0,02; 1,37]	0,095	0,999
KI: Konfidenzintervall; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; OR: Odds Ratio; SUE: schwerwiegendes unerwünschtes Ereignis															

Schwere unerwünschte Ereignisse

Verzerrungspotenzial zu schweren unerwünschten Ereignissen

Tabelle 31: Bewertung des endpunktspezifischen Verzerrungspotenzials: schwere unerwünschte Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Endpunkt- übergreifendes Verzerrungs- potenzial	Verblindung Endpunkterheber	ITT-Prinzip adäquat umgesetzt	Ergebnis- unabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunkt- spezifisches Verzerrungs- potenzial
ENDEAR	niedrig	ja	ja	ja	nein ^a	hoch
a: Keine Angaben zur Verteilung der Begleitbehandlung in den Subgruppen; aus den Angaben für die Gesamtpopulation geht hervor, dass vielfältige Begleitbehandlungen erfolgten. ITT: Intention to treat						

Ergebnisse zu schweren unerwünschten Ereignissen

Tabelle 32: Ergebnisse – schwere unerwünschte Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Subgruppe	Erkrankungsdauer ≤ 12 Wochen (frühsymptomatischer Therapiebeginn)							Erkrankungsdauer > 12 Wochen (spätsymptomatischer Therapiebeginn)							Interaktions- test	
	Nusinersen		Schein- behandlung		Scheinbehandlung vs. Nusinersen			Nusinersen		Schein- behandlung		Scheinbehandlung vs. Nusinersen				p-Wert
	N	n (%)	N	n (%)	OR	[95 %-KI]	p-Wert	N	n (%)	N	n (%)	OR	[95 %-KI]	p-Wert		
ENDEAR																
mindestens 1 schweres UE	34	18 (53)	18	16 (89)	0,14	[0,03; 0,71]	0,017	46	27 (59)	23	17 (74)	0,50	[0,17; 1,51]	0,219	0,203	
KI: Konfidenzintervall; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; OR: Odds Ratio; UE: unerwünschtes Ereignis																

Therapieabbrüche wegen unerwünschter Ereignisse

Verzerrungspotenzial zu Therapieabbrüchen wegen unerwünschter Ereignisse

Tabelle 33: Bewertung des endpunktspezifischen Verzerrungspotenzials: Therapieabbrüche wegen unerwünschter Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Studie	Endpunkt- übergreifendes Verzerrungs- potenzial	Verblindung Endpunkterheber	ITT-Prinzip adäquat umgesetzt	Ergebnis- unabhängige Berichterstattung	Fehlen sonstiger Aspekte	Endpunkt- spezifisches Verzerrungs- potenzial
ENDEAR	niedrig	ja	ja	ja	nein ^a	hoch
a: Keine Angaben zur Verteilung der Begleitbehandlung in den Subgruppen; aus den Angaben für die Gesamtpopulation geht hervor, dass vielfältige Begleitbehandlungen erfolgten						
ITT: Intention to treat						

Ergebnisse zu Therapieabbrüchen wegen unerwünschter Ereignisse

Tabelle 34: Ergebnisse – Therapieabbrüche wegen unerwünschter Ereignisse (Vergleich früh- vs. spätsymptomatischen Therapiebeginn)

Subgruppe	Erkrankungsdauer ≤ 12 Wochen (frühsymptomatischer Therapiebeginn)							Erkrankungsdauer > 12 Wochen (spätsymptomatischer Therapiebeginn)							Interaktions- test p-Wert
	Nusinersen		Schein- behandlung		Scheinbehandlung vs. Nusinersen			Nusinersen		Schein- behandlung		Scheinbehandlung vs. Nusinersen			
	N	n (%)	N	n (%)	OR	[95 %-KI]	p-Wert	N	n (%)	N	n (%)	OR	[95 %-KI]	p-Wert	
ENDEAR															
Therapie- abbrüche wegen UEs	34	3 (9)	18	7 (39)	0,15	[0,03; 0,69]	0,015	46	10 (22)	23	9 (39)	0,43	[0,15; 1,29]	0,132	0,274
KI: Konfidenzintervall; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; OR: Odds Ratio; UE: unerwünschtes Ereignis															

A3.3.2.10.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Im Folgenden werden die Ergebnisse zu den Endpunkten SUEs, schwere UEs, Therapieabbruch wegen UEs sowie Rückenschmerzen (als SUE und schweres UE) dargestellt.

Tabelle 35: Ergebnisse – (schwerwiegende) unerwünschte Ereignisse (Vergleich prä- vs. frühsymptomatischen Therapiebeginn)

Kohorte	Präsymptomatischer Therapiebeginn			Frühsymptomatischer Therapiebeginn			Intervention vs. Vergleich		
	N	Kinder mit Ereignissen		N	Kinder mit Ereignissen		OR	[95 %-KI]	p-Wert
		n	%		n	%			
Biogen 2019^a									
Lebensalter von 1 Jahr^b									
Anzahl mit mindestens 1 SUE	15	6	40	34	24	71	0,28	[0,08; 0,99]	0,045
Anzahl mit mindestens 1 schweren UE	15	1	7	34	18	53	0,06	[0,01; 0,54]	0,002
Anzahl Abbruch wegen UE	15	0	0	34	3	9	0,29	[0,01; 5,98]	0,401
Anzahl mit Rückenschmerzen	15	0	0	34	1	3	0,72	[0,03; 18,71]	0,844
SUE	15	0	0	34	0	0	n. b.	n. b.	n. b.
schweres UE	15	0	0	34	0	0	n. b.	n. b.	n. b.
a: präsymptomatischer Therapiebeginn: Kinder aus der Studie NURTURE, die 2 <i>SMN2</i> -Kopien aufwiesen; frühsymptomatischer Therapiebeginn: Kinder aus der Studie ENDEAR, die 2 <i>SMN2</i> -Kopien aufwiesen und deren Therapie mit Nusinersen innerhalb ≤ 12 Wochen nach Symptombeginn begann									
b: Unterschied in den Erhebungszeitpunkten zwischen den Kohorten maximal 10 %									
KI: Konfidenzintervall; n: Anzahl Kinder mit Ereignis; N: Anzahl ausgewerteter Kinder; n. b.: nicht berechnet; OR: Odds Ratio; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis									

A3.3.2.11 Gesundheitsbezogene Lebensqualität des Kindes

A3.3.2.11.1 Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Der Endpunkt gesundheitsbezogene Lebensqualität des Kindes wurde in der eingeschlossenen Studie nicht erhoben.

A3.3.2.11.2 Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Der Endpunkt gesundheitsbezogene Lebensqualität des Kindes wurde in der Studie ENDEAR, die Biogen 2019 unter anderem zugrunde liegt, nicht erhoben, womit keine Daten für das Teilkollektiv mit frühsymptomatischem Therapiebeginn verfügbar waren. Ergebnisse zu diesem Endpunkt lagen folglich nicht vor.

A3.3.2.12 Metaanalysen

Es wurde keine Metaanalyse durchgeführt.

A3.3.2.13 Sensitivitätsanalysen

Es wurde keine Sensitivitätsanalyse durchgeführt.

A3.3.2.14 Subgruppenmerkmale und andere Effektmodifikatoren

Es wurde keine Subgruppenanalyse und / oder Interaktionstests über die in den Abschnitten A3.3.2.1, A3.3.2.4 und A3.3.2.10 genannten hinaus durchgeführt.

A3.4 Studien zur diagnostischen Güte

A3.4.1 Charakteristika der in die Bewertung eingeschlossenen Studien

A3.4.1.1 Studiendesign und Studienpopulationen

Im Folgenden werden die Studien zur diagnostischen Güte anhand der Studiencharakteristika (Tabelle 36) und Ein- und Ausschlusskriterien (Tabelle 37) dargestellt. Zur Studienpopulation wurden in den eingeschlossenen Studien keine Angaben gemacht, die über die Anzahl ausgewerteter Neugeborener hinausgehen. Es wurde weder das Alter noch die Geschlechtsverteilung beschrieben. Eine Übersicht über Details des Indextests und des Referenzstandards gibt Tabelle 38.

Tabelle 36: Charakterisierung der eingeschlossenen Studien zur diagnostischen Güte

Studie	Studiendesign	Anzahl ausgewerteter Neugeborener n	Studienziel	Ort und Zeitraum der Durchführung	Zielgrößen
Identifikation aller Erkrankten (5q-assozierte SMA unabhängig von der Anzahl der SMN2-Kopien)					
Chien 2017	prospektive diagnostische Kohortenstudie, VOPT	120 267	Beschreibung der Ergebnisse eines 5q-assozierte-SMA-Screenings in Taiwan	Taiwan: 11/2014–09/2016	richtig-positive, falsch-positive Testergebnisse, PPV ^a
Czibere/Vill 2019	prospektive diagnostische Kohortenstudie, VOPT ^b	213 279 ^c	Beschreibung der Ergebnisse eines 5q-assozierte-SMA-Screenings in Deutschland	Deutschland: Bayern und NRW 01/2018– Mitte/2019	richtig-positive, falsch-positive Testergebnisse, PPV ^a
Kraszewski 2018	prospektive diagnostische Kohortenstudie, VOPT	3826	Beschreibung der Ergebnisse eines 5q-assozierte-SMA-Screenings in New York State	USA, New York State: 01/2016–01/2017	richtig-positive, falsch-positive Testergebnisse, PPV ^a
Partielle Identifikation von Erkrankten (5q-assozierte SMA mit < 4 SMN2-Kopien)					
Kariyawasam 2019	prospektive diagnostische Kohortenstudie, VOPT	103 903	Beschreibung des klinischen Wegs von screeningpositiven Neugeborenen in Australien	Australien: New South Wales und Australian Capital Territory 08/2018–07/2019	richtig-positive, falsch-positive Testergebnisse, PPV ^a
<p>a: In der Publikation werden keine Zielgrößen a priori definiert, aber richtig-positive Testergebnisse angegeben oder aber es liegen Angaben vor, die eine entsprechende Einordnung erlauben. Der PPV kann selbst berechnet werden.</p> <p>b: Bei der Studie handelt es sich um eine 1-armige Interventionsstudie der Screeningkette, womit sie nicht die Anforderungen des Berichts für Studien zur Screeningkette erfüllt, denen ein vergleichendes Studiendesign zugrunde liegen muss. Für die Fragestellung der diagnostischen Güte hingegen erfüllt sie die Einschlusskriterien als prospektive diagnostische Kohortenstudie.</p> <p>c: Die Daten entstammen der Publikation Czibere 2019; jüngere Daten aus Vill 2019 wurden nicht herangezogen, da diese von den Autorinnen und Autoren zum einen als vorläufig bezeichnet wurden; zum anderen waren der Publikation keine Angaben zur genetischen Überprüfung dieser vorläufigen Ergebnisse mittels MLPA zu entnehmen.</p> <p>MLPA: multiplexe ligationsabhängige Sondenamplifikation; NRW: Nordrhein-Westfalen; PPV: positiver prädiktiver Wert; SMA: spinale Muskelatrophie; VOPT: Verification of only positive Testers</p>					

Tabelle 37: Ein- / Ausschlusskriterien für Neugeborene in den Studien zur diagnostischen Güte

Studie	Wesentliche Einschlusskriterien	Wesentliche Ausschlusskriterien
Identifikation aller Erkrankten (5q-assozierte SMA unabhängig von der Anzahl der SMN2-Kopien)		
Chien 2017	▪ k. A.	▪ k. A.
Czibere/Vill 2019	▪ k. A.	▪ k. A.
Kraszewski 2018	▪ k. A.	▪ Neugeborene mit akuter Erkrankung oder angeborenen Anomalien auf der neonatalen Intensivstation
Partielle Identifikation von Erkrankten (5q-assozierte SMA mit < 4 SMN2-Kopien)		
Kariyawasam 2019	▪ k. A.	▪ k. A.
k. A.: keine Angabe		

Tabelle 38: Indextest und Referenzstandard – Studien zur diagnostischen Güte

Studie	Indextests	Referenzstandard
Identifikation aller Erkrankten (5q-assozierte SMA unabhängig von der Anzahl der SMN2-Kopien)		
Chien 2017	2-stufiger Nachweis der homozygoten <i>SMN1</i> -Deletion mit Filterkartenblut 1. RT-PCR mit TaqMan SNP Genotypisierung 2. ddPCR von derselben Probe bei positivem Screeningergebnis zwecks Absicherung und Bestimmung der Anzahl der <i>SMN2</i> -Kopien	Nachweis der homozygoten <i>SMN1</i> -Exon-7-Deletion mittels MLPA mit frischem Vollblut und Bestimmung der Anzahl der <i>SMN2</i> -Kopien
Czibere/Vill 2019	Nachweis der homozygoten <i>SMN1</i> Exon 7 Deletion mittels Multiplex TaqMan qPCR mit Filterkartenblut (Bestimmung durch die Second-Derivate-Maximum-Analyse); bei unzureichenden Proben und bei positiven Testergebnissen wurde ein neues Stück Filterkartenblut ausgestanzt und zwecks interner Überprüfung erneut getestet.	Nachweis der homozygoten <i>SMN1</i> Exon 7 Deletion mittels MLPA mit frischem Vollblut in einem externen Labor und Bestimmung der Anzahl der <i>SMN2</i> -Kopien
Kraszewski 2018	Nachweis der homozygoten <i>SMN1</i> -Exon-7-Deletion mittels Multiplex TaqMan qPCR mit Filterkartenblut (Bestimmung der <i>SMN1</i> -Exon-7-Kopienzahl durch die 2- $\Delta\Delta$ Ct-Methode); nachgeschaltete Sequenzierung zur Absicherung	genetische Analyse in einem externen Labor in Bezug auf homozygote <i>SMN1</i> -Exon-7-Deletion und Bestimmung der Anzahl der <i>SMN2</i> -Kopien
Partielle Identifikation von Erkrankten (5q-assozierte SMA mit < 4 SMN2-Kopien)		
Kariyawasam 2019	2-stufiger Nachweis der homozygoten <i>SMN1</i> -Exon-7-Deletion mit Filterkartenblut 1. Perkin Elmer 4-plex RT-PCR 2. ddPCR von einer neuen Probe bei positivem Screeningergebnis zwecks Absicherung; zudem Bestimmung der Anzahl der <i>SMN2</i> -Kopien in der 1. und 2. Probe (mit < 4 <i>SMN2</i> -Kopien als screeningpositivem Befund und \geq 4 <i>SMN2</i> -Kopien als negativem Testergebnis)	Genetischer Nachweis in einer Klinik für neuromuskuläre Erkrankungen
ddPCR: digitale Tröpfchen-Polymerase-Kettenreaktion; MLPA: multiplexe ligationsabhängige Sondenamplifikation; qPCR: quantitative Echtzeit-Polymerase-Kettenreaktion; RT-PCT: Echtzeit-Polymerase-Kettenreaktion; <i>SMN1</i> : Survival Motor Neuron 1; <i>SMN2</i> : Survival Motor Neuron 2; SNP: Einzelnukleotid-Polymorphismus		

A3.4.1.2 Einschätzung des Verzerrungspotenzials

Die Bestimmung des Verzerrungspotenzials und der Übertragbarkeit der Ergebnisse der Primärstudien auf die Fragestellung des Berichts erfolgt auf Basis des Instruments QUADAS 2 [66] (siehe Abschnitt A2.7.2.2).

Verzerrungspotenzial nach QUADAS 2

In der folgenden Tabelle 39 ist die Bewertung des Verzerrungspotenzials der 3 Studien zur diagnostischen Güte dargestellt.

Tabelle 39: Verzerrungspotenzial nach QUADAS 2

Studie	Patientenselektion (Domäne 1)	Indextest I (Domäne 2)	Referenzstandard (Domäne 3)	Patientenfluss und zeitlicher Ablauf (Domäne 4)	Zusammenfassende Einschätzung
Identifikation aller Erkrankten (5q-assozierte SMA unabhängig von der Anzahl der SMN2-Kopien)					
Chien 2017	niedrig	niedrig	niedrig	niedrig	niedrig
Czibere/Vill 2019	unklar	niedrig	niedrig	niedrig	niedrig
Kraszewski 2018	unklar	unklar	niedrig	niedrig	hoch
Partielle Identifikation von Erkrankten (5q-assozierte SMA mit < 4 SMN2-Kopien)					
Kariyawasam 2019	niedrig	niedrig	niedrig	niedrig	niedrig

Bedenken bezüglich der Übertragbarkeit nach QUADAS 2

In der folgenden Tabelle 40 ist die Bewertung zu den Bedenken bezüglich der Übertragbarkeit der Ergebnisse der Primärstudien auf die Fragestellung des Berichts dargestellt.

Tabelle 40: Bedenken bezüglich der Übertragbarkeit QUADAS 2

Studie	Patientenselektion (Domäne 1)	Indextest I (Domäne 2)	Referenzstandard (Domäne 3)	Zusammenfassende Einschätzung
Identifikation aller Erkrankten (5q-assozierte SMA unabhängig von der Anzahl der SMN2-Kopien)				
Chien 2017	gering	gering	gering	gering
Czibere/Vill 2019	gering	gering	gering	gering
Kraszewski 2018	gering	unklar	gering	gering
Partielle Identifikation von Erkrankten (5q-assozierte SMA mit < 4 SMN2-Kopien)				
Kariyawasam 2019	gering	gering	gering	gering

A3.4.2 Ergebnisse zu den Zielgrößen

A3.4.2.1 Ergebnisse zum positiven prädiktiven Wert

Die nachfolgende Tabelle 41 stellt die Ergebnisse der 4 eingeschlossenen Studien zur diagnostischen Güte des Neugeborenentests dar. Eine metaanalytische Zusammenfassung der Ergebnisse zum PPV erfolgte nicht, da die verwendeten Testverfahren und die Zielerkrankungsdefinition in den Studien differieren. Die rechte Spalte zeigt den PPV für jede einzelne Studie.

Tabelle 41: Ergebnisse der Studien zur diagnostischen Güte

Studie	n	Test-positive im Indextext	Indextest	Trennwert	Referenzstandard	Trennwert	RP	FN	FP	RN	PPV	[95 %-KI]
Identifikation aller Erkrankten (5q-assoziierte SMA unabhängig von der Anzahl der SMN2-Kopien)												
Chien 2017	120 267	7	RT-PCR, ddPCR	$\Delta Rn < 1$, k. A.	MLPA	k. A.	7	k. A. ^a	0	k. A.	100	[64,6; 100] ^b
Czibere/Vill 2019	213 279	30	qPCR	Cq < 36	MLPA	k. A.	30	k. A. ^c	0	k. A.	100	[88,6; 100] ^b
Kraszewski 2018	3826	1	qPCR	RQ = 0	genetische Analyse in externem Labor	k. A.	1	k. A.	0	k. A.	100	[20,7; 100] ^b
Partielle Identifikation von Erkrankten (5q-assoziierte SMA mit < 4 SMN2-Kopien)												
Kariyawasam 2019	103 903	10	RT-PCR, ddPCR	k. A., k. A.	genetischer Nachweis	k. A.	9	k. A.	1 ^d	k. A.	90	[59,6; 98,2] ^b
<p>a: Der Publikation waren keine Angaben zu systematisch erfassten falsch-negativen Ergebnissen zu entnehmen; es wurde aber beschrieben, dass den Autorinnen und Autoren keine falsch-negativen Fälle bekannt waren; zudem wurde in der Publikation beschrieben, dass sie – anhand eines SMA-Registers – fortführen, nach übersehenen Fällen spät auftretender SMA zu suchen.</p> <p>b: eigene Berechnung; KI nach Wilson</p> <p>c: Den Publikationen waren keine Angaben zu systematisch erfassten falsch-negativen Ergebnissen zu entnehmen; es wurde aber beschrieben, dass bis zum jeweiligen Publikationszeitpunkt keine falsch-negativen Fälle an die Autorinnen und Autoren berichtet wurden.</p> <p>d: Bei diesem Neugeborenen handelt es sich um ein Kind mit konsanguinem Hintergrund, das Träger der Erkrankung war.</p> <p>Cq: Quantifizierungszyklus; ddPCR: digitale Tröpfchen-Polymerase-Kettenreaktion; FN: falsch-negativ; FP: falsch-positiv; k. A.: keine Angabe; KI: Konfidenzintervall; MLPA: multiplexe ligationsabhängige Sondenamplifikation; n: Anzahl ausgewerteter Neugeborener; PPV: positiver prädiktiver Wert; qPCR: quantitative Echtzeit-Polymerase-Kettenreaktion; Rn: Normalized Reporter Fluorescence Intensity; RN: richtig-negativ; RP: richtig-positiv; RQ: relative Häufigkeit der Anzahl an Genkopien; RT-PCR: Echtzeit-Polymerase-Kettenreaktion</p>												

A3.4.2.2 Sensitivitätsanalysen

Es wurde keine Sensitivitätsanalyse durchgeführt.

A3.4.2.3 Subgruppenmerkmale und andere Effektmodifikatoren

Es wurde keine Subgruppenanalyse durchgeführt.

A4 Kommentare

Nachfolgend werden die Ergebnisse der vorliegenden Nutzenbewertung kommentiert. Sofern thematisch zutreffend, werden dabei Aspekte aus der Anhörung zum Vorbericht gewürdigt. In Abschnitt A4.3 werden alle wesentlichen Aspekte gewürdigt, die in Kapitel 5 und in den Abschnitten A4.1 bis A4.2 noch nicht adressiert wurden.

A4.1 Bericht im Vergleich zu anderen systematischen Übersichten

Im Rahmen der systematischen Literaturrecherche wurden 3 systematische Übersichten identifiziert.

Dabei handelte es sich um 1 Cochrane-Review [94] sowie je 1 systematische Übersichtsarbeit belgischer Autoren [95] und der American Academy of Neurology [9,95]. Alle 3 systematischen Übersichtsarbeiten hatten nicht die Bewertung eines Neugeborenencreenings auf SMA im Fokus, sondern die Behandlung der SMA. Im Cochrane-Review wurde körperliches Training beim SMA-Typ III untersucht. Die Teilnehmerinnen und Teilnehmer waren zwischen 10 und 48 Jahre alt. Eine Vergleichbarkeit mit dem vorliegenden Bericht war somit nicht gegeben. Die Übersichtsarbeiten aus Belgien und den USA unterzogen Nusinersen einer systematischen Bewertung. Die in dieser Übersichtsarbeit herangezogenen Studien wurden auch im Rahmen der Informationsbeschaffung für den vorliegenden Bericht identifiziert. Aufgrund der anderen Ausrichtung (Nutzenbewertung eines Neugeborenencreenings auf SMA) wurden im vorliegenden Bericht nicht alle dieser Studien in die Nutzenbewertung eingeschlossen (siehe Kapitel 5).

A4.2 Kritische Reflexion des Vorgehens

Unterscheidung der klinischen Formen der SMA

Im Stellungnahmeverfahren zum Berichtsplan wurde Bezug genommen auf die im Hintergrund des Berichtsplans beschriebene Unterteilung der 5q-assozierten SMA in 4 verschiedene Typen. Es wurde darauf hingewiesen, dass die Typenunterteilung mittlerweile nicht mehr trennscharf vorgenommen werden könne und die SMA als Krankheitskontinuum zu betrachten sei. Die in Deutschland angewandte ICD-10-Codierung zur Diagnoseklassifikation nehme diese Tatsache auf, indem eine Einteilung in eine infantile SMA und eine sonstige vererbte SMA erfolge. Der Hintergrund des vorliegenden Berichts wurde zum Vorbericht um diese Unterteilung der ICD-10 ergänzt.

Frühsymptomatischer versus spätsymptomatischen Therapiebeginn

Nutzenbewertung auf Basis von Subgruppenanalysen

Für die Bewertung eines früheren gegenüber einem späteren Therapiebeginn bei bereits vorliegender Symptomatik wurden ausschließlich Subgruppenanalysen aus 1 RCT herangezogen. Diese stellen die auf absehbare Zeit beste wissenschaftliche Evidenz für die Bewertung dieser Fragestellung dar, wie im Folgenden dargelegt wird:

Bei der RCT ENDEAR handelt es sich um eine zentrale Zulassungsstudie zu Nusinersen, in der die herangezogenen Subgruppenanalysen zur Krankheitsdauer a priori geplant waren und die Randomisierung nach dem Subgruppenmerkmal Krankheitsdauer stratifiziert wurde. In einem solchen Fall kann die Vergleichbarkeit der Behandlungsgruppen innerhalb der Subgruppen angenommen werden [30]. Es ist darüber hinaus nicht zu erwarten, dass im Nachgang zu dieser Zulassungsstudie weitere prospektive Studien durchgeführt werden, die den Vergleich eines frühsymptomatischen mit einem spätsymptomatischen Therapiebeginn bei bekannter Diagnose zum Zuteilungszeitpunkt untersuchen. Gegen diese Verifizierung der beobachteten Effekte sprechen ethische Überlegungen, denn die aktuelle Versorgungssituation sieht vor, den betroffenen symptomatischen Kindern die verfügbare Therapie unmittelbar nach der Diagnose zukommen zu lassen. Ungeachtet dessen war den Herstellerunterlagen kein Hinweis auf eine solche prospektive nicht randomisierte vergleichende Studie zum Therapiebeginn zu entnehmen.

Grenze für einen früh- und spätsymptomatischen Therapiebeginn

In der vorliegenden Nutzenbewertung wurde zur Unterscheidung eines früh- und spätsymptomatischen Therapiebeginns die zeitliche Grenze von 12 Wochen ab Symptombeginn als angemessen bewertet. Dieses Vorgehen basierte auf folgenden Annahmen und Hintergründen:

Die tatsächliche derzeitige Zeitspanne zwischen Symptombeginn und genetischer Diagnosestellung bzw. Therapiebeginn in Deutschland ist unbekannt. Einer systematischen Übersicht auf Basis internationaler Studien kann jedoch entnommen werden, dass diese Spanne sich zwischen den unterschiedlichen Typen der SMA stark unterscheidet und bei der infantilen Form der SMA durchschnittlich 3,6 Monate beträgt [15]. Als gegeben wurde angenommen, dass in Deutschland seit der Zulassung von Nusinersen die Behandlung zumindest von Neugeborenen bzw. Kindern mit 5q-assoziierter SMA i. d. R. unmittelbar nach der Diagnosestellung beginnt und dass eine genetische Diagnosesicherung auch vollumfänglich möglich ist. Die aus der Studie ENDEAR herangezogenen Daten der geplanten Subgruppenanalysen zur Krankheitsdauer (\leq bzw. $>$ 12 Wochen zwischen Alter des Kindes bei Symptombeginn und Alter bei Studieneinschluss) wurden demnach im Sinne eines Vergleichs eines früh- versus einen spätsymptomatischen Therapiebeginn als übertragbar auf den deutschen Versorgungskontext bewertet. Die Subgruppenergebnisse dieser Studie gewinnen zusätzlich dadurch an Bedeutung, dass auch die Randomisierung in der Studie stratifiziert nach der Krankheitsdauer erfolgte, d. h., es wurde die Voraussetzung geschaffen, dass sich die Kinder mit einer Krankheitsdauer \leq bzw. $>$ 12 Wochen zu Studienbeginn etwa gleichmäßig auf die beiden Behandlungsgruppen verteilen konnten.

Präsymptomatischer versus frühsymptomatischen Therapiebeginn

Beschränkung der Vergleichsgruppe auf Kinder mit frühsymptomatischem Therapiebeginn

In der Anfrage an den Hersteller erfolgte für die Vergleichsgruppe der Studie (symptomatischer Therapiebeginn) explizit eine Beschränkung auf diejenigen Kinder, die frühsymptomatisch behandelt wurden, d. h., die Therapie wurde bei diesen Kindern innerhalb von 12 Wochen nach

Symptomeintritt begonnen. Die Begründung für diese Vorgehensweise liegt 1.) in der oben beschriebenen Annahme zur derzeitig anzunehmenden Zeitspanne zwischen Symptom- und Therapiebeginn (diagnostischer Delay von etwa 3 Monaten) und 2.) in dem bereits im Vorbericht gezeigten Vorteil eines früheren Therapiebeginns (innerhalb von 12 Wochen nach Symptomeintritt) im Vergleich zu einem späteren Therapiebeginn (> 12 Wochen seit Symptomeintritt) bei symptomatischen Kindern mit infantiler SMA. Eine möglichst frühe Behandlung von erkrankten Kindern, insbesondere mit erwartbar schweren Krankheitsverläufen (infantile SMA), ist entsprechend als Versorgungsstandard anzusetzen und stellt daher den angemessenen Komparator für den angefragten Vergleich zu einer noch weiteren Therapievorverlegung auf die präsymptomatische Situation dar.

Auswertungszeitpunkte der Studie Biogen 2019

Um die Ergebnisse beider Teilkollektive, die sich hinsichtlich des Alters bei Studieneinschluss unterschieden, angemessen miteinander vergleichen zu können, wurden Daten nicht zu Nachbeobachtungszeitpunkten ab Studienbeginn angefragt, sondern zeitlich am Lebensalter der Kinder verankert (vgl. Abschnitt A3.1.1.3). Die Teilkollektive wurden ab dem Zeitpunkt, ab dem die Kinder pro Kollektiv durchschnittlich etwa gleich alt waren, als sinnvoll vergleichbar hinsichtlich der angefragten Endpunkte angenommen. Dies unterlag 2 Annahmen: 1.) dass sich die Prognose der Kinder mit 2 *SMN2*-Kopien unbehandelt erwartbar homogen deterministisch und mit frühem Krankheitsbeginn (infantile SMA) äußern würde [96]; 2.) dass sich unter Therapie bei sonst ähnlichen Bedingungen deutliche Unterschiede zwischen den Vergleichsgruppen in einem bestimmten, gleichen Lebensalter primär auf den Zeitpunkt des Therapiebeginns zurückführen lassen würden.

Angefragt wurden beim Hersteller Ergebnisse zu 2 Zeitpunkten: 1.) eine Auswertung im Alter von 1 Jahr und 2.) eine Auswertung für das späteste gemeinsame Alter der beiden Studienkollektive im Beobachtungszeitraum. Zu Letzterem lieferte der Hersteller Ergebnisse zum Lebensalter von 148 Tagen mit der Begründung, dies sei der letzte Auswertungszeitpunkt, zu dem alle Kinder lebten. Die Daten zu diesem frühen Zeitpunkt waren für die Nutzenbewertung nicht geeignet, allein weil das mittlere Alter der frühsymptomatisch behandelten Kinder zu Studieneinschluss nur unwesentlich niedriger lag. Ungeachtet dessen lieferte der Hersteller jedoch auch Daten zum angefragten Zeitpunkt im Lebensalter von 1 Jahr, die für den Abschlussbericht herangezogen werden konnten.

Präsymptomatischer Therapiebeginn mit Nusinersen versus natürlichen Verlauf

Die Ergebnisse der Studie NURTURE [29], in der Kinder bereits vor Symptomeintritt (präsymptomatisch) mit Nusinersen behandelt werden, unterscheiden sich zweifellos dramatisch vom natürlichen Verlauf. Allerdings stellt der natürliche Verlauf lediglich den angemessenen Komparator für die Fragestellung dar, *ob* eine Behandlung nützt oder nicht. Zur Klärung der berichtsrelevanten Teilfrage, *wann* eine Therapie beginnen sollte beziehungsweise, ob eine Therapievorverlegung von Nutzen ist, kann ein Vergleich mit dem natürlichen Verlauf nicht beitragen.

Die 1. Frage, ob eine Behandlung nützt, wurde bereits mit der Zulassung von Nusinersen 2017 geklärt. Denn der medizinische Zusatznutzen von Nusinersen als Arzneimittel mit dem Status „Orphan Drug“ gilt gemäß § 35a Abs. 1 Satz 10 SGB V bereits durch die Zulassung als belegt. Bei Vorliegen einer 5q-assozierten SMA kann die Behandlung mit Nusinersen somit aktuell als Standardtherapie betrachtet werden.

Die 2. Frage, wann eine Therapie beginnen sollte, konnte in Teilen bereits im Vorbericht beantwortet werden. Demnach profitieren Kinder mit infantiler SMA mehr von einem frühsymptomatischen als von einem spätsymptomatischen Therapiebeginn. Um zu prüfen, ob eine weitere Therapievorverlegung – mit Beginn vor Einsetzen erster Symptome – nützt, war aufgrund der gezeigten Effektunterschiede einzig der Vergleich zu einem frühsymptomatischen Therapiebeginn zweckmäßig. Der natürliche Verlauf stellte in diesem Zusammenhang nicht den angemessenen Komparator dar.

A4.3 Würdigung der Anhörung

Insgesamt wurden 6 Stellungnahmen zum Vorbericht frist- und formgerecht eingereicht.

Die im Rahmen der Anhörung vorgebrachten Aspekte wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des Vorberichts überprüft. Die wesentlichen Argumente werden im Folgenden diskutiert.

In den eingereichten Stellungnahmen wurden folgende Aspekte angesprochen, die bereits in Kapitel 5 und in Abschnitt A4.2 adressiert wurden:

- der angefragte Vergleich zu einem präsymptomatischen versus einen frühsymptomatischen Therapiebeginn (Kapitel 5 und Abschnitt A4.2)
- Nutzenbewertung auf Basis von Subgruppenanalysen (Kapitel 5 und Abschnitt A4.2)
- der Vergleich zum natürlichen Verlauf (Abschnitt A4.2)
- der Ausschluss der Studie CHERISH (Kapitel 5)
- geplanter Vergleich der Daten von Kindern aus dem deutschen Neugeborenencreening-Programm mit Daten von Kindern, die im SMARtCARE Register dokumentiert werden (Kapitel 5)

Die Stellungnahmen werden weitergehend in der nachfolgenden Tabelle 42 gewürdigt. Literaturzitate aus dem Wortlaut der Stellungnahmen finden sich im separaten Dokument „Dokumentation der Anhörung“.

Die Zusammenfassung aller Änderungen des Abschlussberichts gegenüber dem Vorbericht, die sich u. a. durch die Anhörung zum Vorbericht ergeben haben, ist in Abschnitt A1.2 dargestellt.

Tabelle 42: Würdigung der Stellungnahmen

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
3 Methoden			
Avexis	Abschnitt A2.1.4 (S. 29)	<p><u>Anmerkung:</u> Das IQWiG führt aus, dass vorrangig RCTs in die Nutzenbewertung eingeschlossen werden. Auf Studien niedrigerer Evidenzstufe wird nur dann zurückgegriffen, wenn weder RCTs oder prospektiv geplante Kohortenstudien mit zeitlich paralleler Kontrollgruppe vorliegen. Grundsätzlich wird dieser Ansatz von AveXis unterstützt und wird als adäquat betrachtet, spezifische Fallkonstellationen wie bei der spinalen Muskelatrophie sollten jedoch zusätzlich Berücksichtigung finden. Diese Berücksichtigung muss dabei über den Einschluss von prospektiv geplanten Kohortenstudien mit zeitlich paralleler Kontrollgruppe hinausgehen. Nach der Durchführung der ersten Studien mit wirksamen Therapien zur Behandlung der spinalen Muskelatrophie, d. h. nach Durchführung der vom IQWiG herangezogenen Studie ENDEAR als auch der von AveXis durchgeführten Studie CL-101, ist die Durchführung von randomisierten kontrollierten Studien gegenüber Placebo (bzw. Scheinintervention) bzw. die zeitlich parallele Untersuchung unbehandelter Patienten ethisch nicht mehr vertretbar. Grund dafür ist die ungewöhnliche Schwere der Erkrankung und die dramatische Verbesserung des Verlaufs durch die Behandlung mit den neuen Therapien. Aus diesem Grund wurde auch die Studie ENDEAR frühzeitig beendet und die Patienten im Placebo-Arm bekamen die Möglichkeit, in den Nusinersen-Arm zu wechseln. Diese Entscheidung war alternativlos, da die Ergebnisse der Studie zeigen, dass bei den Patienten in der Placebo-Gruppe bereits ein irreversibler Schaden eingetreten war und Nusinersen gleichzeitig eine wirksame Therapieoption darstellt. Diese Beobachtung hat zum einen zur Folge, dass die Beobachtungsdauer in der ENDEAR-Studie relativ kurz ist, zum anderen aber eben auch, dass für jedwede nachfolgende Therapie die</p>	<p>Keine Änderung.</p> <p>Begründung: Berücksichtigung von Studien niedrigerer Evidenzstufe bereits in Abschnitt A2.1.4 thematisiert.</p> <p>Für Teilfragestellungen, zu denen noch keine Ergebnisse herangezogen werden konnten, werden Studien niedrigerer Evidenz berücksichtigt (z. B für den Vergleich präsymptomatische Behandlung versus symptomatische Behandlung).</p> <p>Die Problematik der Verifizierung der – mittels Subgruppenanalysen aus ENDEAR – beobachteten Effekte wird geteilt und wurde bereits in Abschnitt A4.2 des Vorberichts adressiert. Placebo (beziehungsweise Scheinintervention) beziehungsweise die zeitlich parallele Untersuchung unbehandelter Patientinnen und Patienten ist mit der Zulassung von Nusinersen nicht länger der geeignete Komparator. Vielmehr müssten neue Therapien im Rahmen qualitativ hochwertiger Studien mit Nusinersen verglichen werden.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Durchführung einer vergleichbaren Studie nicht mehr möglich ist. Für die Bewertung der Fragestellung des Vorberichts zieht das IQWiG nun ausschließlich eine randomisierte Studie heran, in der Patienten, die früh behandelt wurden, mit Patienten verglichen werden, die erst später behandelt wurden – in beiden Subgruppen im Vergleich mit keiner Behandlung. Vergleichbare Studien sind nach aktuellem Stand des medizinischen Wissens weder für Nusinersen noch für neuere Therapien mehr durchführbar. Vor diesem Hintergrund führt eine rigorose Beschränkung auf RCTs und nicht-randomisierte vergleichende Studien, sofern mindestens eine solche Studie identifiziert worden ist, dazu, dass jegliche weitere Evidenz, die bereits vorliegt oder in der Zukunft generiert wird und den Vorteil einer frühen Behandlung demonstriert, per se ausgeschlossen wird. Diese Methodik ist nicht geeignet, die Datenlage in einer zukünftigen Neubewertung/Aktualisierung vollumfänglich und sachgerecht abzubilden. Mit der Diskussion der Studie NURTURE (S. 19/20) signalisiert das IQWiG bereits, dass auch Studien, die ohne zeitlich parallele Kontrollgruppe geplant sind, für die Bewertung der Fragestellung relevant sind.</p> <p><u>Vorgeschlagene Änderung:</u> Es sollte im Methodik-Teil des Bewertungsberichts klargestellt werden, dass zusätzlich zu RCTs und Studien mit zeitlich parallel geplanter Kontrollgruppe unter Würdigung der besonderen Umstände auch Studien niedrigerer Evidenzstufe herangezogen werden, wenn diese für die Beantwortung der wissenschaftlichen Fragestellung relevant sind.</p>	
4 Ergebnisse			
Avevis	Allgemeine Aspekte zu den Ergebnissen	Das IQWiG sieht als Ergebnis seiner Analyse einen Anhaltspunkt für einen Nutzen des Neugeborenencreenings auf 5q-assozierte SMA (spinale Muskelatrophie) im	Kenntnisnahme.

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Vergleich zu keinem entsprechenden Screening. Der Nutzen ergibt sich aus einer früheren Diagnose, die auch eine frühere Behandlung ermöglicht. Neben der Eignung der vorhandenen diagnostischer Testverfahren leitet das IQWiG diesen Nutzen aus dem nachgewiesenen Vorteil einer früheren Behandlung ab. Für diesen Nachweis zieht das IQWiG die Studie ENDEAR heran, die den Wirkstoff Nusinersen gegenüber einer Scheinbehandlung („Sham-control“) untersucht. AveXis teilt die Interpretation der Studienergebnisse und die daraus abgeleitete Empfehlung. Wie von den Autoren des Vorberichts ausgeführt, befindet sich das Produkt AVXS-101 (Onasemnogen abeparvovec, ZOLGENSMA®) von AveXis derzeit im Zulassungsverfahren, die Zulassung wird Q1/2020 erwartet. Die Phase-III-Studien sind derzeit noch nicht abgeschlossen, es liegen finale Ergebnisse der Studie CL-101 vor [1, 2]. Die derzeit vorliegenden (auch noch unveröffentlichten) Ergebnisse von AveXis bestätigen das Ergebnis des IQWiG-Vorberichts: Bei Patienten mit SMA Typ 1, die nach Symptombeginn, d. h. klinischer Diagnose, in die Studien eingeschlossen wurden, zeigt sich ein dramatischer Vorteil der Gentherapie gegenüber dem natürlichen Verlauf der Erkrankung bzw. ausschließlich unterstützenden Maßnahmen (v. a. Beatmung und künstliche Ernährung). Dieser Vorteil tritt bei der Behandlung präsymptomatischer Patienten (genetische Diagnose) noch deutlicher zutage [3, 4].</p> <p>Die Einführung neuer Behandlungsmethoden für eine schwerwiegende, früh zum Tod führende Erkrankung wie SMA Typ 1, für die es vormals keine Behandlungsmöglichkeiten gab, bietet Patienten erstmals die Chance auf einen günstigeren Krankheitsverlauf im Sinne einer langfristigen Vermeidung von schweren Krankheitssymptomen und Tod. Voraussetzung dafür, die neuen Chancen bestmöglich nutzen zu können, ist eine möglichst frühe Diagnose und somit eine Identifizierung der Patienten bevor ein großer irreversibler Schaden eingetreten</p>	<p>Keine Berücksichtigung der Ergebnisse der Studie CL-101 in der vorliegenden Nutzenbewertung, da dieses Medikament derzeit nicht zugelassen ist.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>ist. Die Innovationen im Bereich der spinalen Muskelatrophie haben dazu geführt, dass auf internationaler Ebene (Niederlande, USA) die Einführung eines Neugeborenencreenings bereits empfohlen wird [5, 6]. Die sachgerechte Analyse der vorliegenden Evidenz und die korrekte Schlussfolgerung und Empfehlung zugunsten eines Neugeborenencreenings auf 5q-assozierte spinale Muskelatrophie wird von AveXis daher ausdrücklich begrüßt.</p>	
AveXis	Allgemeine Aspekte zu den Ergebnissen	<p>Eine Stellungnahme zum Berichtsplan, der am 27.02.2019 veröffentlicht wurde, wurde von AveXis nicht eingereicht. Die im Berichtsplan beschriebenen Methoden sind nach Ansicht von AveXis als adäquat einzuschätzen. Die Ausführung des dargelegten Berichtsplans umfasst nun jedoch einige Konkretisierungen, vor allem in Hinblick auf eingeschlossene patientenrelevante Endpunkte, die einer Kommentierung und – nach Einschätzung von AveXis – einer nachfolgenden Änderung bedürfen. Dies wird in der folgenden Stellungnahme zu spezifischen Aspekten ausgeführt.</p>	Kenntnisnahme.
MDS	Allgemeine Aspekte zu den Ergebnissen	<p>1. Der vorliegende Vorbericht leitet einen Anhaltspunkt für einen Nutzen des Neugeborenen-Screenings (NBS) auf 5q-assozierte SMA im Vergleich zu keinem Screening ab. Der festgestellte Anhaltspunkt eines Nutzens eines Neugeborenencreenings ist mit den verwendeten Methoden nicht ableitbar. Aus der positiven Interaktion für das Subgruppen bildende Merkmal „Krankheitsdauer“ und den beobachteten Effektunterschieden in der herangezogenen ENDEAR-Studie allein kann keine Schlussfolgerung auf die Kausalität dieses Merkmals gezogen werden – auch wenn dieses Merkmal die einzige Stratifizierungsvariable für die Randomisierung war. Andere potenzielle Effektmodifikatoren müssen zwingend berücksichtigt und als kausal einflussnehmende Merkmale ausgeschlossen werden. Eine solche Untersuchung ist offenbar unterblieben, obwohl in ENDEAR auch andere subgruppenbildende Merkmale mit positiven Interaktionen beobachtet wurden. Insbesondere die</p>	<p>Keine Änderung. Begründung: Die aus der Studie ENDEAR herangezogenen Subgruppenanalysen zur Krankheitsdauer waren zum einen a priori geplant. Zum anderen wurde in der Studie die Randomisierung nach dem Merkmal Krankheitsdauer stratifiziert durchgeführt. In einem solchen Fall kann die Vergleichbarkeit der Behandlungsgruppen innerhalb der Subgruppen angenommen werden. Ungeachtet dessen zeigen die für den Vorbericht beim Hersteller nachgeforderten Baselinedaten der berichtsgegenständlichen Subgruppen, dass das mediane Alter bei Symptombeginn und auch die angegebene Spannweite in den beiden Subgruppen</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		ebenfalls feststellbare Effektmodifikation durch das Merkmal „Alter bei Symptombeginn“ könnte als Faktor der eigentlich wirksame sein.	<p>vergleichbar sind (siehe Tabelle 15 im Vorbericht).</p> <p>Unabhängig davon, dass die Subgruppenanalyse nach Krankheitsdauer als einzige den zu bewertenden Vergleich eines frühen versus einen späteren Behandlungsbeginn widerspiegelt, wurden andere potenzielle Effektmodifikatoren, zu denen Subgruppenanalysen aus ENDEAR verfügbar waren, auf Gesamtstudienenebene nochmals näher betrachtet. Diese Prüfung führte zu keiner Änderung der Vorgehensweise im Bericht.</p> <p>Ethische Überlegungen sprechen darüber hinaus gegen eine Verifizierung der beobachteten Effekte in zukünftigen Studien. Dies wurde bereits im Vorbericht ausführlich adressiert (siehe Abschnitt A.4.2).</p> <p>Schließlich verdeutlichen die vom Hersteller nachgelieferten Daten zum Vergleich prä- versus symptomatischen Therapiebeginn die Tatsache, dass eine Vorverlegung der Therapie grundlegend nützt und die Kinder umso mehr von der Therapie profitieren, je früher sie behandelt werden.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
MDS	Allgemeine Aspekte zu den Ergebnissen	<p>Ad 1. Fehlende Zuschreibbarkeit des Merkmals Krankheitsdauer als Ursache der beobachteten Interaktion: Um eine hinreichend sichere Aussage zum Nutzen eines NBS aussprechen zu können, die auf Subgruppenanalysen von Therapiestudien mit nicht per NBS identifizierten, symptomatischen Patienten beruhen sollen, muss das Merkmal Krankheitsdauer weitgehend zweifelsfrei als das für die beobachteten Effektunterschiede ursächliche Merkmal identifiziert werden.</p> <p>Es gibt nun aber auch andere univariate Subgruppeneffekte basierend auf anderen Merkmalen in der herangezogenen Studien ENDEAR, die in S18-02 nicht besprochen wurden, insbesondere das Alter bei Symptombeginn (aber häufiger auch für das Geschlecht der Kinder oder die Region der Studiendurchführung [1]). Nur weil die Krankheitsdauer Stratifikationsvariable in ENDEAR war, bedeutet nicht, dass diese der einzige Effektmodifikator sein kann.</p> <p>Die SMA ist gekennzeichnet von einem progredienten Verlauf, dessen Geschwindigkeit mit in die Definition der verschiedenen Verlaufsformen eingeht, selbst innerhalb der klassisch mit Typ I, II, III etc. bezeichneten Klassen (Typ IA, IB und IC als Unterteilung von Typ I z. B. [4]). Es wäre daher nicht verwunderlich, wenn es eine möglicherweise sogar starke Korrelation zwischen dem Alter bei Symptombeginn und der Krankheitsdauer bei Therapiebeginn gäbe, weil ein schnellerer Verlauf in jüngerem Alter zu einer Diagnose und zu einer schnelleren Behandlung ab Diagnose führen könnte. Es käme jetzt darauf an zu zeigen, dass dennoch die Krankheitsdauer ursächlich für den Effektunterschied wäre (und damit einen Nutzen einer Therapievorverlagerung nahelegt) und nicht das Alter bei Symptombeginn, das hier als Surrogat für den Krankheitsschweregrad angesehen werden könnte. Dieses Merkmal Alter bei Symptombeginn kann genauso Plausibilität als möglicher Effektmodifikator in Anspruch nehmen wie das Merkmal Krankheitsdauer.</p>	<p>Keine Änderung.</p> <p>Begründung: <u>Siehe Würdigung in vorangehender Zeile zu 1. (ebenfalls MDS)</u></p> <p>Bzgl. des angeführten möglichen Powerproblems: Die Einteilung in die 2 Subgruppen für das Merkmal Alter bei Symptombeginn führt, wie in der Stellungnahme beschrieben, mit dem gewählten Schwellenwert von 12 Wochen zu sehr ungleichen Gruppengrößen, was die Power des Interaktionstests in der Tat negativ beeinflussen kann. Es kann nicht ausgeschlossen werden, dass mit einem anderen Schwellenwert für die Subgruppenzuteilung signifikante Unterschiede für das Merkmal Alter bei Symptombeginn hätten beobachtet werden können. Dies ist aber kein Grund, die beobachtete statistisch signifikante Interaktion für das Merkmal Krankheitsdauer nicht heranzuziehen.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz																																
		<p>Die in ENDEAR vorliegenden Interaktionen für das Merkmal Alter bei Symptombeginn sind nicht vollständig konkordant zu denen für die Krankheitsdauer. Da die getroffene Wahl des Schwellenwertes für die Subgruppenbildung nach Alter nach Symptombeginn mit 12 Wochen zu sehr ungleichen Gruppengrößen führt (104 vs. 17 Kinder in der „jüngeren“ bzw. „älteren“ Subgruppe; medianes Alter bei Symptombeginn waren 8 Wochen [1]), ist die Power des Interaktionstests im Vergleich zur Krankheitsdauer (52 vs. 69 Kinder in der „kürzeren“ bzw. „längeren“ Subgruppe) nochmals geringer, und Zufallsbefunde sind zu erwarten. Die Subgruppenanalysen zeigen, wenn positive Interaktionen vorliegen, dass die Effekte in der „jüngeren“ Subgruppe immer größer zugunsten Nusinersens sind als in der „älteren“. Anhand des Endpunktes „CHOP-INTEND-Responder“ (für das aufgrund positiver IA-Tests für beide Merkmale Subgruppenanalysen vorliegen) sei die Übereinstimmung hier beispielhaft angeführt [1]:</p>																																	
		<table border="1"> <thead> <tr> <th data-bbox="819 839 1021 946">Studie ENDEAR, CHOP-INTEND-Responder</th> <th data-bbox="1021 839 1093 946">IA</th> <th colspan="3" data-bbox="1093 839 1375 946">Kürzer bzw. jünger</th> <th colspan="3" data-bbox="1375 839 1666 946">Länger bzw. älter</th> </tr> <tr> <th data-bbox="819 946 1021 1010">Subgruppenmerkmal</th> <th data-bbox="1021 946 1093 1010">P-Wert</th> <th data-bbox="1093 946 1167 1010">Nu n/N (%)</th> <th data-bbox="1167 946 1240 1010">Sham n/N (%)</th> <th data-bbox="1240 946 1375 1010">Effekt (Risiko-differenz [95%-KI])</th> <th data-bbox="1375 946 1449 1010">Nu n/N (%)</th> <th data-bbox="1449 946 1523 1010">Sham n/N (%)</th> <th data-bbox="1523 946 1666 1010">Effekt (Risiko-differenz [95%-KI])</th> </tr> </thead> <tbody> <tr> <td data-bbox="819 1010 1021 1090">Krankheitsdauer</td> <td data-bbox="1021 1010 1093 1090">0,009</td> <td data-bbox="1093 1010 1167 1090">30/34 (88)</td> <td data-bbox="1167 1010 1240 1090">0/18 (0)</td> <td data-bbox="1240 1010 1375 1090">0,85 [0,71; 0,98]</td> <td data-bbox="1375 1010 1449 1090">28/46 (61)</td> <td data-bbox="1449 1010 1523 1090">1/23 (4)</td> <td data-bbox="1523 1010 1666 1090">0,57 [0,40; 0,73]</td> </tr> <tr> <td data-bbox="819 1090 1021 1169">Alter bei Symptombeginn</td> <td data-bbox="1021 1090 1093 1169">0,03</td> <td data-bbox="1093 1090 1167 1169">55/72 (76)</td> <td data-bbox="1167 1090 1240 1169">1/32 (3)</td> <td data-bbox="1240 1090 1375 1169">0,73 [0,62; 0,85]</td> <td data-bbox="1375 1090 1449 1169">3/8 (38)</td> <td data-bbox="1449 1090 1523 1169">0/9 (0)</td> <td data-bbox="1523 1090 1666 1169">0,34 [-0,01; 0,69]</td> </tr> </tbody> </table> <p data-bbox="819 1169 1666 1224">IA: Interaktion, KI: Konfidenzintervall, Nu: Nusinersen</p>	Studie ENDEAR, CHOP-INTEND-Responder	IA	Kürzer bzw. jünger			Länger bzw. älter			Subgruppenmerkmal	P-Wert	Nu n/N (%)	Sham n/N (%)	Effekt (Risiko-differenz [95%-KI])	Nu n/N (%)	Sham n/N (%)	Effekt (Risiko-differenz [95%-KI])	Krankheitsdauer	0,009	30/34 (88)	0/18 (0)	0,85 [0,71; 0,98]	28/46 (61)	1/23 (4)	0,57 [0,40; 0,73]	Alter bei Symptombeginn	0,03	55/72 (76)	1/32 (3)	0,73 [0,62; 0,85]	3/8 (38)	0/9 (0)	0,34 [-0,01; 0,69]	
Studie ENDEAR, CHOP-INTEND-Responder	IA	Kürzer bzw. jünger			Länger bzw. älter																														
Subgruppenmerkmal	P-Wert	Nu n/N (%)	Sham n/N (%)	Effekt (Risiko-differenz [95%-KI])	Nu n/N (%)	Sham n/N (%)	Effekt (Risiko-differenz [95%-KI])																												
Krankheitsdauer	0,009	30/34 (88)	0/18 (0)	0,85 [0,71; 0,98]	28/46 (61)	1/23 (4)	0,57 [0,40; 0,73]																												
Alter bei Symptombeginn	0,03	55/72 (76)	1/32 (3)	0,73 [0,62; 0,85]	3/8 (38)	0/9 (0)	0,34 [-0,01; 0,69]																												

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Um ein NBS zu empfehlen, müsste, wenn man diesen methodischen Weg wählt, die Interaktion eindeutig auf das Merkmal Krankheitsdauer zurückgeführt werden können. Daher sind die qua Herstelleranfragen nachgereichten Daten (Tab. 8 des Vorberichts) unzureichend: Die notwendigen Untersuchungen zum Ausschluss anderer, plausibler Effektmodifikatoren sind nicht geleistet. Insbesondere ist die Wahl eines anderen Schwellenwertes für das Merkmal Alter bei Symptombeginn, das ähnlichere Subgruppengrößen erzeugt hätte (d. h. in der Nähe des Medians von 8 Wochen) nicht nachgefordert worden, was eine bessere Beurteilung der Konkordanz der Interaktionen zwischen den beiden gleichermaßen plausiblen Merkmalen Alter bei Symptombeginn und Krankheitsdauer ermöglicht hätte.</p> <p>Da aber auch schon jetzt gleichgerichtete Effektmodifikationen zu beobachten sind, wäre es überraschend, wenn bei besser gepowerter Analyse für das Merkmal Alter bei Symptombeginn keine Interaktionen mehr aufträten und eine klare Zuschreibung auf das Merkmal Krankheitsdauer möglich wäre.</p>	
MDS	Allgemeine Aspekte zu den Ergebnissen	<p>2. Selbst wenn man den nach den oben angesprochenen Methoden geführten Nachweis als gegeben annehmen würde, sind die abgeleiteten Empfehlungen a) unklar und b) unvollständig</p> <p>a. Welches Vorgehen für die Behandlung von Screeningpositiven wird genau empfohlen? Man kann implizit an der Art der nachgeforderten Daten zum Vergleich ENDEAR-NURTURE ablesen, dass der Nutzen eines NBS wohl in einer „Watchful-Waiting“-Strategie gesehen wird, da aus ENDEAR nur Daten zu Kindern mit <12 Wochen Krankheitsdauer für diesen Vergleich nachgefordert werden. Wie der Umgang mit Screeningpositiven genau aussehen soll, bleibt aber unklar. Einerseits wird konstatiert, dass keine Übertragung der Ergebnisse auf eine mögliche Behandlung präsymptomatischer Screeningpositiver erfolgen kann.</p>	<p>zu 2. a): Keine Änderung.</p> <p>Begründung:</p> <ul style="list-style-type: none"> ▪ Empfehlungen zur konkreten Therapieplanung sind nicht Bestandteil dieses IQWiG-Auftrags. Auf bestehende Unklarheiten und mögliche nachteilige Konsequenzen des Screenings wurde bereits im Vorbericht verwiesen. Wie im Bericht beschrieben, liegen derzeit keine evidenzbasierten Leitlinien zur Behandlung der SMA vor, aber erste internationale Handlungsempfehlungen bezüglich Nusinersen [32,33,55]; Diese Empfehlungen werden u. a. derzeit auch im laufenden

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Andererseits bleibt unklar, ob eine Übertragbarkeit vorliegt – und weshalb diese denn gegeben sei – auf Screeningpositive, die z. B. der Krankheit vorausgehende Frühsymptome/Biomarker (wie veränderte Muskelaktionspotenzialmessungen) aufweisen und dann behandelt werden sollen, oder ob keine Übertragbarkeit vorliegt und nur symptomatisch werdende Kinder behandelt werden sollen. In letzterem Falle wiederum ergibt sich nicht zwingend eine Empfehlung für eine so weit reichende Maßnahme wie ein NBS, da die in ENDEAR eingeschlossene Population ausschließlich nicht per NBS identifizierte Patienten waren. Den Ergebnissen des IQWiG-Vorberichts strikt folgend wäre eine Kampagne zur Aufmerksamkeitserhöhung bzgl. SMA-(Früh-)Symptomen möglicherweise anstelle eines NBS ausreichend, um bei symptomatischen Kindern mit infantiler SMA einen früh-symptomatischen Therapiebeginn zu erreichen. Eine solche Empfehlung würde auch mögliche Nachteile eines NBS vermeiden, indem es keine falsch-negativen Ergebnisse geben könnte. Im Falle eines NBS könnte es bei den Kindern mit falsch-negativem Screeningergebnis sogar zu einer Verzögerung der Diagnose und damit des Therapiebeginns kommen, weil die Beteiligten sich fälschlich in Sicherheit wiegen können.</p>	<p>deutschen Neugeborenencreening-Pilotprojekt in Bayern und NRW [31,32] umgesetzt, an dem sich beispielsweise orientiert werden kann. Weitere Therapien befinden sich erst in der Zulassungsprüfung.</p> <ul style="list-style-type: none"> ▪ Für den Nachweis eines Nutzens eines Neugeborenencreenings (in Kombination mit einer Vorverlegung der Diagnosestellung und Behandlung) im Vergleich zu keinem Screening ist die Identifikation von erkrankten Kindern nicht zwingend mittels Neugeborenencreening nötig; wie im Berichtsplan dargestellt, müssen Ergebnisse aber übertragbar sein auf die Screeningsituation (siehe Kapitel A.2 im Vorbericht); der Linked-Evidence-Ansatz stellt in diesem Zusammenhang ein etabliertes Verfahren dar ▪ Aufgrund der unterschiedlichen Auftretenszeitpunkte und Schwere der Verläufe der unterschiedlichen SMA-Formen ist die Nutzenbewertung des Neugeborenencreenings für alle an SMA Erkrankten nicht trivial; dies wurde ausführlich im Vorbericht adressiert (siehe Kapitel 5). Insgesamt ergibt sich ein Nutzen des Screenings, da ein großer Teil der an SMA erkrankten Kinder maßgeblich von der frühzeitigen Diagnose und Therapie profitiert, wie die neuen Daten im Abschlussbericht nun zusätzlich bestätigen. ▪ Der von den Stellungnehmenden adressierte mögliche Nachteil für Kinder mit falsch-negativem Screeningergebnis scheint theoretischer Natur zu sein. In 2 der 4 für den Abschlussbericht identifizierten Studien zur

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
			<p>diagnostischen Güte wird beschrieben, dass zumindest keine falsch-negativen Fälle bekannt waren bzw. an die Studienleiter herangetragen wurden. Einzig diejenigen Kinder (ca. 5 %; siehe Kapitel 1), die eine andere Auffälligkeit als die homozygote <i>SMN1</i>-Deletion aufweisen (heterozygote Deletion und Punktmutation auf dem vorhandenen <i>SMN1</i>-Gen), werden mittels der derzeit verfügbaren diagnostischen Tests per Neugeborenencreening nicht gefunden. Das Neugeborenencreening aus diesem Grund nicht einzuführen, wird als unverhältnismäßig erachtet.</p> <p>Aufgrund des Einbezugs neuer Daten wurden die Empfehlungen für den Abschlussbericht überarbeitet.</p>
		<p>b. Gegenstand eines NBS sind genetisch diagnostizierte SMA-Patienten (mit einer homozygoten <i>SMN1</i>-Mutation) mit einer variablen Zahl an <i>SMN2</i>-Kopien (1 bis mindestens 6) und damit verbunden höchst unterschiedliche Prognosen der SMA-Verlaufsform. Der Nutzen wird jedoch anhand der ENDEAR-Studie abgeleitet, in der nur SMA-Typ I Patienten mit 2 <i>SMN2</i>-Kopien eingeschlossen worden sind. Zu allen anderen Gruppen Screeningpositiver mit teils erheblich anderen Verläufen fehlt jegliche Aussage.</p> <p>Auch bleibt unberücksichtigt, dass aufgrund des vorgesehenen Screenings auf eine homozygote <i>SMN1</i>-Deletion ca. 8% der Kinder mit 5q-assoziierter SMA, die von anderen Mutationen als den Deletionen verursacht werden [10], per se falsch negativ bzgl. einer 5q-assozierten SMA beurteilt werden. Dies kann, weil man sich auf den negativen</p>	<p>Zu 2. b): Keine Änderung</p> <p>Begründung: Einige Aspekte wurden bereits in <u>vorangehender Zeile (ebenfalls MDS)</u> zu 2. a) gewürdigt.</p> <p>Bzgl. der Studie CHERISH: Als Kriterium zum Studieneinschluss war hinsichtlich der Population gefordert, dass die Diagnosestellung bei Patientinnen und Patienten mit früherem Therapiebeginn auf die Screeningsituation bei Neugeborenen übertragbar sein sollte. Dies war bei CHERISH nicht der Fall, was im Vorbericht in Kapitel 5 detailliert erläutert wurde. Wie die Stellungnehmenden korrekt beschreiben, weichen die Schwellenwerte der a priori geplanten Subgruppenanalysen zur Krankheitsdauer stark vom deutschen und</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Befund des NBS verlässt, sogar zu Verzögerungen einer Diagnose und Behandlung führen.</p> <p>Weshalb die Studie CHERISH [8], die Aussagen für Patienten mit 3 SMN2-Kopien ermöglicht hätte, ausgeschlossen wurde, bleibt unklar. Das Argument, dass die berichteten Krankheitsdauern – anders als bei ENDEAR – stark vom deutschen Versorgungskontext abweichen, ist zunächst korrekt. Auch die berichteten Subgruppenanalysen [1] lassen aufgrund der gewählten Schwellenwerte keine Aussagen zu. Allerdings hätten Subgruppenanalysen mit angemesseneren Schwellenwerten, die dann doch für den deutschen Versorgungskontext relevant gewesen wären, nachgefordert werden können. Der Vorbericht selbst scheint von der pauschalen Ablehnung ja auch nicht auszugehen, werden doch die als versorgungsrelevant angesehenen Daten aus CHERISH (mit einer Krankheitsdauer von <14 Monaten) für einen anderen Vergleich nachgefordert (s.u.).</p>	<p>internationalen Versorgungsstandard ab (die Therapie erfolgt deutlich später als üblich nach Symptombeginn, vgl. [15]) und können so keine validen und übertragbaren Erkenntnisse zur Therapievorverlegung liefern. Auch deshalb wurde auf eine Anfrage nach Post-hoc-Subgruppenanalysen zu einem geeigneten Schwellenwert verzichtet, weil hier die Subgruppe mit früherem Therapiebeginn erwartbar eine zu kleine Patientenzahl (< 10 Personen) umfasst hätte und somit die Ergebnisse des Interaktionstests nicht verwertbar gewesen wären.</p> <p>Wie auch die beim Hersteller angefragten Daten zum Vergleich eines prä- versus symptomatischen Therapiebeginns bei Kindern mit 3 SMN2-Kopien zeigen, ist der Anteil an Kindern aus CHERISH, deren Situation am ehesten auf die Screeningsituation übertragbar ist, zu gering, als dass auf Basis dessen eine Bewertung erfolgen könnte.</p>
MDS	Allgemeine Aspekte zu den Ergebnissen	<p>3. Die nachgeforderten Daten zum Vergleich eines NBS mit präsymptomatischer Behandlung Screeningpositiver vs. NBS mit symptomatischer Behandlung Screeningpositiver, die durch den Vergleich von NURTURE und ENDEAR geliefert werden soll, werden nicht die korrekten Antworten liefern.</p> <p>Beim NBS auf 5q-assozierte SMA wird über die homozygote SMN1-Deletion die molekulargenetische Diagnose der SMA gestellt, zusätzlich wird die SMN2-Kopienanzahl bestimmt. Eine präzise Prognose der SMA-Verlaufsform über die SMN2-Kopienanzahl ist im präsymptomatischen Stadium nicht möglich. Zwar gibt es bei SMA-Patienten eine Korrelation zwischen SMN2-Kopienzahl und späterem Symptomverlauf (d. h. SMA-Typ), allerdings besteht keine 100%-ige Konkordanz. So entwickeln Patienten mit 2 SMN2-</p>	<p>Zu 3.: Keine Änderung der Vorgehensweise im Bericht</p> <p>Begründung:</p> <ul style="list-style-type: none"> ▪ Der angefragte Vergleich ist per Definition retrospektiver Natur; gemäß Berichtsmethodik kann auf Basis eines solchen Vergleichs eine Nutzensaussage nur erfolgen, wenn die vorliegenden Effekte so groß sind, dass sie sich nicht allein durch Verzerrung erklären lassen (dramatischer Effekt) (siehe Vorbericht, Abschnitt A2.2.4). ▪ Entsprechend wurde in der Anfrage an den Hersteller auf die Beachtung bestmöglicher

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Kopien in ca. 80% der Fälle den infantilen Verlauf (SMA Typ I), ein Teil entwickelt jedoch eine mildere Verlaufsform (ca. 16% Typ II, ca. 4% sogar Typ III) [3]. Für 3 SMN2-Kopien ist die Variabilität noch größer (ca. 15% Typ I, 55% Typ II, 30% Typ III [3]).</p> <p>Die Nichtbeachtung der Variabilität der Verlaufsformen bei gegebener SMN2-Kopienanzahl der präsymptomatischen Patienten führt in beiden angestrebten Vergleichen mit Patienten mit 2 bzw. 3 SMN2-Kopien zu verzerrten Ergebnissen. Im angestrebten Vergleich für Patienten mit 2 SMN2-Kopien werden ausschließlich Typ I Kinder aus ENDEAR mit einer Mischpopulation aus präsymptomatisch behandelten Patienten mit Typ I und II (und zu einem geringen Anteil auch Typ III) aus NURTURE verglichen werden. Im Vergleich von Patienten mit 3 SMN2-Kopien werden Typ II Kinder aus CHERISH mit einer Mischpopulation aus präsymptomatisch behandelten SMA-Patienten mit Typ I, II und III aus NURTURE verglichen werden. Ohne Adjustierungen wird der Vergleich der Patienten mit 2 SMN2-Kopien zugunsten der präsymptomatischen Behandlung verzerrt sein, während die Verzerrungsrichtung im Vergleich der Patienten mit 3 SMN2-Kopien unklar bleiben muss.</p> <p>Des Weiteren scheint die ausschließliche Heranziehung der Patienten aus ENDEAR, die maximal 12 Wochen Krankheitsdauer aufwiesen, nur unter der Annahme eines Nutzens einer wie auch immer gearteten Vorverlagerung bei symptomatischen Patienten sinnvoll. Da hierfür lediglich ein (hier unseres Erachtens fälschlich abgeleiteter) Anhaltspunkt festgestellt wurde, und die Studie ENDEAR mit einer medianen Krankheitsdauer von 13,5 Wochen [6] repräsentativ für die in der Population gefundene Verzögerung zwischen Symptombeginn und Diagnose bei Typ I-Patienten von 3,6 Monaten zu sein scheint [7], erscheint auch die Fragestellung zum Vergleich kein NBS vs. NBS mit präsymptomatischer Behandlung Screeningpositiver</p>	<p>Strukturgleichheit hingewiesen; darüber hinaus wurden Ergebnisse zeitlich am Lebensalter verankert (anstelle von Daten z. B. zur Dauer seit Studienbeginn) angefordert und geliefert.</p> <ul style="list-style-type: none"> ▪ Die thematisierte Korrelation zwischen <i>SMN2</i>-Kopienanzahl und Symptomverlauf (SMA-Typ) ist bekannt und wurde im Vorbericht hinreichend adressiert. Da bei einem präsymptomatischen Therapiebeginn die Verlaufsform per se nicht bekannt sein kann, wurde beim angefragten Vergleich behelfsweise die genetische Disposition berücksichtigt (separate Anfrage nach Vergleichen für die Populationen mit 2 beziehungsweise 3 <i>SMN2</i>-Kopien). ▪ Für den angefragten Vergleich zu Kindern mit 2 <i>SMN2</i>-Kopien kann eine Vergleichbarkeit der Kohorten ausreichend sicher angenommen werden, da, wie u. a. der Stellungnahme des MDS zu entnehmen ist, etwa 80 % der präsymptomatisch behandelten Kinder mit 2 <i>SMN2</i>-Kopien unbehandelt eine infantile SMA entwickeln würden (siehe auch [96]). ▪ Bei Kindern mit 3 <i>SMN2</i>-Kopien ist die Variabilität der SMA-Verläufe größer [96]. Dies war bekannt, hätte aber in bestimmten Konstellationen und Vorliegen eines dramatischen Effekts nicht per se gegen die Ableitung eines Effekts gesprochen. Aufgrund der zu geringen Fallzahlen konnten die Ergebnisse des angefragten Vergleichs zu Kindern mit 3 <i>SMN2</i>-Kopien jedoch ohnehin nicht herangezogen werden.

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		eine sinnvolle, sodass auch der vollständige aktive Arm aus ENDEAR als Vergleichsgruppe herangezogen werden sollte.	<ul style="list-style-type: none"> ▪ Angesichts des gezeigten Vorteils eines früheren Therapiebeginns (innerhalb von 12 Wochen) im Vergleich zu einem späteren bei bereits vorliegender Symptomatik ist eine möglichst frühe Behandlung der Kinder als Versorgungsstandard anzusetzen und stellt somit den angemessenen Komparator für den angefragten Vergleich zu einer noch weiteren Vorverlegung der Therapie auf die präsymptomatische Situation dar (siehe Abschnitt A4.2). Würden die präsymptomatisch behandelten Kinder mit allen symptomatisch behandelten Kindern aus ENDEAR verglichen werden, ist davon unabhängig erwartbar, dass die beobachteten Effekte noch deutlicher werden würden.
MDS	Allgemeine Aspekte zu den Ergebnissen	<p>Ad 3. Nachgeforderte Daten zum Vergleich präsymptomatische vs. symptomatische Therapie: 2 SMN2-Kopien:</p> <p>Es fehlt die Beachtung der unterschiedlichen Prognosen in den beiden Studien. Während in ENDEAR ausschließlich Typ I Kinder mit 2 SMN2-Kopien eingeschlossen wurden, sind es in NURTURE potenziell alle SMA-Verläufe, die mit 2 SMN2-Kopien vorkommen, neben den 80% Typ I v.a. also ca. 16% Typ II und 4% Typ III [3]. Die Studie ENDEAR bildet die Screeningsituation nicht gut ab. Ein naiver Vergleich würde zu verzerrten Effekten zugunsten der präsymptomatischen Behandlung führen, was dann aufgrund einer möglichen falsch positiven Nutzensaussage auch mit der Gefahr einer Übertherapie einer abgesehen vom ohnehin belastenden Applikationsweg nicht risikofreien Therapie verbunden sein könnte [2].</p> <p>Die einmalige Chance einer unverzerrteren Analyse sollte genutzt werden. Diese besteht darin, dass 18/25 NURTURE-Kinder nicht per NBS sondern über ihre älteren, an SMA erkrankten Geschwister identifiziert wurden [5]. Da eine hohe</p>	<p>Keine Änderung der Vorgehensweise im Bericht.</p> <p>Begründung: <u>Siehe Würdigung in vorangehender Zeile zu 3. (ebenfalls MDS)</u> Zu den Geschwisterkindern von der NURTURE-Population: <u>in der vorangehenden Zeile</u> wurde bereits dargelegt, dass der Großteil der Kinder mit 2 SMN2-Kopien ohne Therapie eine infantile SMA entwickeln würde; weitergehende Informationen zu Geschwisterkindern waren daher nicht erforderlich, um eine Nutzensaussage zur Therapievorverlegung treffen zu können, die ohnehin nur auf Basis dramatischer Effekte erfolgen konnte (siehe Vorbericht, Abschnitt A2.2.4).</p> <p>Der aktuellen Publikation De Vivo 2019 [29] sind dessen ungeachtet für einen Teil der</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Konkordanz zwischen Geschwistern hinsichtlich des SMA-Verlaufs besteht [5, 9], können trotz präsymptomatischer Behandlung die zugrunde liegende Verlaufsformen bei diesen Geschwistern mit einiger Sicherheit angenommen und bei den Auswertungen entsprechend berücksichtigt werden (z. B. durch Ausschluss derjenigen NURTURE-Kinder aus Analysen, deren Geschwister trotz 2 SMN2-Kopien einen Typ II Verlauf aufwiesen).</p> <p>Hierzu müsste Biogen offenlegen, wie viele und welche Kinder an SMA erkrankte Geschwister hatten, die einen Typ II oder Typ III Verlauf hatten und dies in den Analysen entsprechend berücksichtigen.</p> <p>Für die Kinder, deren Geschwister einen milderen Verlauf zeigten, wäre ein Vergleich mit den (wenigen) CHERISH-Kindern mit 2 SMN2-Kopien der Relevante.</p> <p>3 SMN2-Kopien:</p> <p>Hier gilt analog das oben Gesagte für Typ II (CHERISH) bzw. Kinder deren Geschwister Typ II entwickelt hatten (NURTURE). NURTURE-Kinder mit 3 SMN2-Kopien und Geschwistern mit Typ I oder III Verlauf müssten aus den Analysen ausgeschlossen werden.</p> <p>Des Weiteren scheint für ENDEAR die Einschränkung auf die Kinder mit einer Krankheitsdauer von <12 Wochen für den Vergleich voreilig. Da die mediane Krankheitsdauer von ca. 13,5 Wochen in ENDEAR [6] ziemlich genau dem entspricht, was an Verzögerung zwischen Symptombeginn und Diagnosestellung in der Bevölkerung für Typ I Kinder festgestellt wurde (3,6 Monate) [7], zielt diese Einschränkung offenbar auf folgenden Vergleich: NBS+Behandlung nach Watchful Waiting vs. NBS+präsymptomatischer Behandlung. Sinnvoll ist dies ohnehin nur dann, wenn man den Nutzen einer Watchful-Wating-Strategie für nachgewiesen hält, was man zum jetzigen Zeitpunkt bezweifeln muss (s. o.).</p> <p>Da auch nur ein Anhaltspunkt ausgesprochen wurde, erscheint zumindest auch der folgende Vergleich</p>	<p>NURTURE-Population, nämlich die Kinder, die im Laufe der Studie Symptome entwickelten, Informationen zu den Geschwisterkindern zu entnehmen.</p> <p>Ferner handelt es sich (ebenso wie zur Teilfragestellung früh- versus spätsymptomatischen Therapiebeginn) bezüglich dieser Teilfragestellung um die auf absehbare Zeit beste wissenschaftliche Evidenz zur Beantwortung der Frage des Nutzens einer präsymptomatischen gegenüber einer symptomatischen Behandlung.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		gerechtfertigt: Kein NBS vs. NBS mit sofortigem Therapiebeginn bei präsymptomatischen Screeningpositiven. Daher sollte auch der gesamt aktive Arm aus ENDEAR mit den alters- und prognosegematchten Kindern aus NURTURE verglichen werden.	
Novartis	Allgemeine Aspekte zu den Ergebnissen	<p>Das IQWiG kommt in seinem Vorbericht zu dem Schluss, dass sich ein Anhaltspunkt für einen Nutzen des Neugeborenencreenings auf 5q-assozierte SMA im Vergleich zu keinem Screening ergibt. Die Novartis Pharma GmbH unterstützt das Ergebnis der Nutzenbewertung ausdrücklich.</p> <p>Das IQWiG betont in seinem Vorbericht in der Zusammenfassung (S.v) weiter, dass <i>„die verfügbaren Daten keine Schlüsse erlauben, ob durch das Screening identifizierte Kinder mit SMA bereits von einem präsymptomatischen Therapiebeginn profitieren würden. Dies gilt insbesondere für diejenigen Kinder mit späterem Krankheitsbeginn. Derzeit unklar ist vor allem die Behandlung von im Screening identifizierten Neugeborenen mit erwartbar spätem Krankheitsbeginn (≥ 4 SMN2-Kopien).“</i></p> <p>Novartis veröffentlichte 2018 in einem Poster Daten aus der laufenden Phase I/II Studie CLMI070X2201 zu Branaplam, einem Wirkstoff zur Behandlung der spinalen Muskelatrophie des Typ1 (1). Die Auswertung des Endpunktes CHOP-INTEND (Beurteilung der motorischen Entwicklung) zeigte hierbei deutlich, dass ein jüngeres Lebensalter bei der Erstdosis mit einer verbesserten motorischen Entwicklung korreliert. Die Nebenwirkungen bleiben dabei sehr gering. Wie der Publikation von Charnas et al. entnommen werden kann, sank der maximal erreichte Wert im CHOP-INTEND mit steigendem Alter der Patienten bei Behandlungsbeginn.</p>	<p>Kenntnisnahme.</p> <p>Keine Berücksichtigung der Ergebnisse der Studie CLMI070X2201, da Medikament zum Zeitpunkt der Berichterstellung noch nicht zugelassen.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		Die Daten suggerieren daher, dass auch Patienten mit ggf. spätem Krankheitsbeginn, die in einem systematischen Neugeborenencreening identifiziert werden, von einer möglichst frühzeitigen medikamentösen Therapie profitieren werden. Inwieweit dies mit einer sofortigen, d.h. präsymptomatischen, Therapieinitiierung einhergehen soll, liegt im Ermessen des behandelnden Experten. Ein positiver Gentest kann im Rahmen eines Neugeborenencreenings potentiell zu einer erhöhten Vigilanz führen, sodass das Einsetzen von Symptomen frühzeitig erkannt und entsprechend eine schnelle Therapiinitiierung erfolgen kann.	
Pilotprojekt Neugeborenen-creening SMA / Deutsche Gesellschaft für Kinder- und Jugendmedizin ^a	A2.1.4 Studientypen	<u>Anmerkung:</u> Es ist geplant, die Daten der Kinder, die im Pilotprojekt detektiert wurden, mit der Kohorte von nach Symptombeginn erfassten Kindern zu vergleichen, die im SMARtCARE Register dokumentiert werden.	Kenntnisnahme und Ergänzung des Kapitels 5 um die Information, dass ein solcher Vergleich geplant ist.
4.1 Ergebnisse der umfassenden Informationsbeschaffung			
Biogen	S. 22 (5.)	<p><i>IQWiG Vorbericht Konsequenzen der Ergebnisse für ein Neugeborenencreening</i></p> <p><i>(...) weitere zentrale Studie CHERISH zu nennen...Das Zeitfenster zwischen Diagnosestellung und Therapiebeginn betrug etwa 30 Monate (Nusinersengruppe) beziehungsweise 18 Monate (Schein-behandlungsgruppe). Dies entspricht nicht der aktuellen Versorgungssituation in Deutschland, da seit der Zulassung von Nusinersen mit der Therapie heutzutage in der Regel direkt nach der Diagnosestellung begonnen wird. Aus diesem Grund wurde die Studie als nicht relevant ausgeschlossen.</i></p> <p><u>Anmerkung:</u> Wie im Dossier zur Nutzenbewertung 2017 beschrieben, zeigen die Ergebnisse der CHERISH-Studie, dass bei späterem Therapieeinsatz Patienten einen signifikanten</p>	<p>Keine Änderung.</p> <p>Begründung: Gegenstand des vorliegenden Auftrags ist die Bewertung des Nutzens eines Neugeborenencreenings auf 5q-assozierte SMA; Der Nutzen der Therapie Nusinersen für Patientinnen und Patienten mit SMA wurde bereits mit der Zulassung von Nusinersen 2017 geklärt. Denn der medizinische Zusatznutzen von Nusinersen als Arzneimittel mit dem Status „Orphan Drug“ gilt gemäß § 35a Abs. 1 Satz 10 SGB V bereits durch die Zulassung als belegt. Wie bereits bei der <u>Würdigung der Stellungnahme des MDS</u> dargelegt, kann die Studie CHERISH keine geeigneten Daten zur</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Nutzen von der Therapie erhalten — auch trotz des langen Zeitfensters zwischen Diagnosestellung und Therapiebeginn. Es zeigte sich ein Vorteil für Nusinersen im Vergleich zu BSC bei Patienten mit späterem SMA-Krankheitsbeginn in den patientenrelevanten Dimensionen Morbidität, Lebensqualität und Sicherheit. Dies konnte anhand der Endpunkte „Anzahl der HFMSE-Responder, Häufigkeit neuer motorischer Meilensteine pro Patient gemäß WHO, Veränderung des RULM-Scores, Veränderung des ACEND-Scores sowie Häufigkeit von Hospitalisierungen“ belegt werden.</p> <p>Dieser Zusatznutzen von Nusinersen wird auch durch den G-BA bestätigt, der die CHERISH-Studie in seine Bewertung mit einbezieht [3]. Der G-BA stuft die Ergebnisse als beträchtlich ein, da es sich hierbei um eine bislang nicht erreichte Abschwächung schwerwiegender Symptome und eine spürbare Linderung der Erkrankung handelt.</p> <p>Zudem belegen die mit dieser Stellungnahme eingereichten zusätzlichen Ergebnisse eines Vergleichs von präsymptomatisch behandelten NURTURE-Patienten zu CHERISH-Patienten mit Therapiebeginn ≤ 14 Monate ebenfalls eine signifikant verbesserte Motorik gemessen anhand des HFMSE (siehe Anhang ‚Tabellenvorlage zum Vergleich präsymptomatischer vs. symptomatischer Therapiebeginn‘). In der CHERISH-Studie zeigte sich im Durchschnitt unter Therapie eine Zunahme des HFMSE-Scores. Im natürlichen Verlauf der Erkrankung zeigt sich so gut wie nie eine Zunahme, bedingt durch den stetig progredienten Verlust der motorischen Nervenfunktion.</p> <p><u>Vorgeschlagene Änderung:</u> Ergebnisse der CHERISH Studie sind in vollem Umfang in der vorliegenden Bewertung zu berücksichtigen. Mit Blick auf die vorliegenden Ergebnisse und gezeigten Vorteile von</p>	<p>Bewertung des Nutzens eines Neugeborenencreenings beitragen.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		frühsymptomatischer und präsymptomatischer Therapie, sind die Ergebnisse der CHERISH Studie als eher konservativ einzustufen und lassen vermuten, dass ein kürzerer Abstand zwischen Diagnose und Therapiebeginn sich positiv auf die Ergebnisse auswirken würde.	
Pilotprojekt Neugeborenen-screening SMA / Deutsche Gesellschaft für Kinder- und Jugendmedizin ^a	Seite V der Zusammenfassung	<p><u>Anmerkung:</u> Aussage auf Seite V der Zusammenfassung, dass „die verfügbaren Daten keine Schlüsse erlauben, ob durch das Screening identifizierte Kinder mit SMA bereits von einem präsymptomatischen Therapiebeginn profitieren“ würden, kann auf Basis neuer Publikationen revidiert werden.</p> <p><u>Vorgeschlagene Änderung:</u> Die inzwischen publizierten Daten der Nurture Studie¹ zeigen klar, dass sich der Verlauf der bereits präsymptomatisch behandelten Kinder mit spinaler Muskelatrophie von dem natürlichen Verlauf unterscheidet. Während Kinder mit zwei SMN-Gen2Kopien üblicherweise spätestens bis zum 2. Lebensjahr permanent beatmungspflichtig oder verstorben sind, war eine dauerhafte Beatmung bei den behandelten Kindern in keinem Fall notwendig.</p> <p>In der in Bayern und in NRW gemeinsam durchgeführten Pilotstudie zum Neugeborenen-Screening bei spinaler Muskelatrophie, welche seit Anfang 2018 läuft und inzwischen mehr als 250.000 Kinder eingeschlossen hat, konnte gezeigt werden, dass der Verlauf der Kinder, die durch das Screening detektiert wurden, dramatisch besser ist als der natürliche Verlauf, wie er aus den bisher publizierten natural-history Daten bekannt ist².</p> <p>Im Unterschied zum Screening-Projekt in Australien³, bei welchem bei 44 % der gescreenten Kinder, die entsprechend dem dortigen Protokoll nicht von Anfang an behandelt wurden, Symptome auftraten, kam es bei keinem unserer Kinder unter Behandlung zum Neuauftreten von Symptomen.</p>	<p>Kenntnisnahme.</p> <p>Es erfolgte eine Anpassung im Bericht dahingehend, dass die Studien Vill 2019 [32] und Kariyawasam 2019 [33], die auch im Rahmen der Nachrecherche identifiziert wurden, als Studien zur diagnostischen Güte eingeschlossen wurden.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>4 Kinder in unserer Studie mit 2 bzw. 3 Kopien von SMN-2, die aus verschiedenen Gründen (1 x fehlende Kostenübernahme, 3 x Entscheidung der Eltern) nicht behandelt wurden, entwickelten Symptome einer SMA.</p>	<p>Bei der Publikation De Vivo 2019 [29], handelt es sich um die 1-armige laufende Studie NURTURE, die alleinstand nicht zur Nutzenbewertung herangezogen werden kann (siehe Kapitel 5 des Vorberichts). Die Ergebnisse dieser Studie, in der Kinder bereits präsymptomatisch mit Nusinersen behandelt werden, unterscheiden sich zweifellos dramatisch vom natürlichen Verlauf. Allerdings stellt der natürliche Verlauf lediglich den angemessenen Komparator für die Fragestellung dar, <i>ob</i> eine Behandlung nützt oder nicht (diese Frage wurde bereits mit der Nutzenbewertung von Nusinersen 2017 positiv beschieden). Zur Klärung der Frage, <i>wann</i> eine Therapie beginnen sollte, ist der natürliche Verlauf nicht der angemessene Komparator (siehe Abschnitt A4.2).</p> <p>Im Rahmen der vorliegenden Bewertung eines Neugeborenencreenings können behelfsweise vergleichende Studien zur Therapievorverlegung beitragen. Die Studie NURTURE fand hier Berücksichtigung im Rahmen des beim Hersteller angefragten Vergleichs zu einem präsymptomatischen Therapiebeginn versus einen späteren Therapiebeginn frühzeitig nach Auftreten von Symptomen (siehe Abschnitt 4.3.1.2 im Abschlussbericht)</p> <p>Die Publikation De Vivo 2019 wird somit lediglich ergänzend zu der vom Hersteller zur Verfügung gestellten retrospektiven vergleichenden Studie berücksichtigt, in der sich ein Studienarm aus Kindern der Studie NURTUTRE zusammensetzt.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
Pilotprojekt Neugeborenen-screening SMA / Deutsche Gesellschaft für Kinder- und Jugendmedizin ^a	Allgemeine Aspekte	Zum Zeitpunkt der Erstellung des Berichts lagen die inzwischen erschienenen Publikationen zu ersten Ergebnissen der in Bayern und in NRW gemeinsam durchgeführten Pilotstudie zum Neugeborenen-Screening bei spinaler Muskelatrophie, aber auch zu anderen Screening-Projekten aus Australien noch nicht vor. Wir dürfen Ihnen die inzwischen veröffentlichte Publikation über die 1-Jahresdaten des in Bayern und NRW laufenden Pilotprojekts in der Anlage beilegen.	Kenntnisnahme (Berücksichtigung der erwähnten Publikationen <u>siehe vorangegangene Zeile</u>).
4.3.3 Bewertung des Verzerrungspotenzials der Ergebnisse			
Biogen	Seite 11 (4.3.3)	<p><i>Vorbericht IQWiG</i></p> <p><i>4.3.3 Bewertung des Verzerrungspotentials der Ergebnisse</i></p> <p><i>Das endpunktspezifische Verzerrungspotenzial für die Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung, zum Endpunkt Erreichen motorischer Meilensteine sowie zu den Endpunkten SUEs, schwere UEs und Therapieabbrüche wegen UEs wurde als hoch bewertet. Dies lag daran, dass keine Angaben zur Verteilung der Begleitbehandlung in den Subgruppen vorlagen. Da in der Studie jedoch eine Begleitbehandlung nach BSC erfolgte und verfügbare Angaben zur Gesamtpopulation auch vielfältige Begleitbehandlungen dokumentieren, wären zur Einschätzung eines möglichen Kointerventionsbias entsprechende Angaben für die Subgruppen erforderlich gewesen. Damit wurde die qualitative Ergebnissicherheit für alle Endpunkte als mäßig bewertet.</i></p>	<p>Keine Änderung.</p> <p>Begründung: Dass eine Behandlung der Studienpopulation nach BSC erfolgte, war zweifelsfrei angemessen und erforderlich, dies und die verwendeten Standards werden nicht infrage gestellt.</p> <p>Für die berichtsrelevanten Subgruppen liegen allerdings keine Angaben dazu vor, welche konkreten Maßnahmen innerhalb der Subgruppen Einsatz fanden und ob diese zwischen den Subgruppen gleichverteilt waren, ist daher nicht bekannt. Es konnte somit für die Subgruppenanalyse keine Einschätzung eines möglichen Kointerventionsbias erfolgen. Daher wird das endpunktspezifische Verzerrungspotenzial für die Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung, zum Endpunkt Erreichen motorischer Meilensteine sowie zu den Endpunkten SUEs, schwere UEs und Therapieabbrüche wegen UEs als hoch bewertet.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p><u>Anmerkung:</u> Sowohl die Einschlusskriterien der ENDEAR-Studie als auch der G-BA in seiner Nutzenbewertung 2017 belegen und bestätigen das Einhalten der BSC-Maßnahmen für alle Studienpatienten und damit deren gleichmäßige Verteilung zwischen den Therapiearmen. Daraus kann kein erhöhtes Verzerrungspotential abgeleitet werden.</p> <p>Gemäß den Einschlusskriterien der Studie ENDEAR mussten alle Patienten die Richtlinien des Konsenses zu Standard of Care nach Wang et al. 2007 für die Teilnahme an der Studie erfüllen und über den Studienverlauf weiterhin einhalten [1]. Diese Richtlinien umfassen die im Studienprotokoll unter Appendix D gelisteten Maßnahmen für die Unterstützung im respiratorischen und gastrointestinalen Bereich, welche bei Kindern mit infantiler SMA meist sehr schwer betroffen sind [2]. Die im Studienprotokoll international anerkannten Maßnahmen umfassten:</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>BSC: respiratorischer Bereich</p> <p>1. Eine Beurteilung der respiratorischen Behandlung und die Diskussion verschiedener Behandlungsoptionen sollte kurz nach der Diagnose durchgeführt werden. Hauptkomponenten der Untersuchung beinhalten die Bewertung der Effizienz des Abhustens, die Beobachtung der Atmung und die Überwachung des Gasaustauschs.</p> <p>2. Das akute Management bei Atemwegserkrankungen erfordert vermehrt Techniken zum Atemwegs- und Sekretmanagement unter Verwendung von mechanischer Insufflations-Exsufflation oder manueller Hustenhilfe, erhöhter Unterstützung der Atmung (einschließlich nicht invasiver Beatmung), Ernährungs- und Hydratationsmanagement und einen frühen Beginn einer Antibiotikatherapie.</p> <p>3. Das chronische Atemwegsmanagement umfasst Methoden für die Atemwegs-Clearance, einschließlich mechanischer Insufflations-Exsufflation oder manueller Hustenhilfe und nicht invasiver Beatmung.</p> <p>4. Routineimpfungen werden empfohlen.</p> <p>5. Diskussion mit den Familien bzgl. der Optionen für die respiratorische Behandlung und Festlegen der Ziele für die chronische und akute respiratorische Behandlung sollten frühzeitig im Krankheitsverlauf geführt werden. Diese sollten in einem laufenden Dialog besprochen werden.</p> <p>6. Die perioperative Versorgung umfasst eine gründliche präoperative Bewertung des respiratorischen Status, idealerweise durch einen Pulmologen und eine vorausschauende Begleitung durch das chirurgische Team und das postoperative Management für eine optimale Versorgung.</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>BSC: gastrointestinaler Bereich und Ernährungsmanagement</p> <p>1. Ernährungsprobleme und Schluckstörungen: Bulbäre Dysfunktion besteht grundsätzlich bei Patienten mit SMA und einer schweren Muskelschwäche und kann dazu führen, dass es zu Ernährungsproblemen und Schluckstörungen kommt. Im Zusammenhang damit kann es zu einer Aspirationspneumonie kommen, die häufig zum Tod führt.</p> <p>2. Gastrointestinale Dysfunktion: Zu den Problemen durch eine gastroösophageale Dysmotilität gehören Verstopfung, verzögerte Magenentleerung und ggf. lebensbedrohlicher gastroösophagealer Reflux.</p> <p>3. Wachstumsstörungen und Probleme bzgl. Unterernährung/Überernährung: Ohne ein optimales Management sind Wachstumsstörungen bei nicht sitzfähigen Patienten üblich. Hingegen besteht bei sitzfähigen und gehfähigen Patienten die Gefahr einer übermäßigen Gewichtszunahme. Das empfohlene Ernährungsmanagement beinhaltet die orale Ernährung oder Sondenernährung, Hydratation und das medizinische oder chirurgische Management gastroösophagealer Refluxkrankheiten.</p> <p>Auch der G-BA beschreibt in den tragenden Gründen im Verfahren zur frühen Nutzenbewertung nach § 35a von Nusinersen ,Alle Patienten erhielten eine Grundversorgung nach internationalen Behandlungsstandards (Best Supportive Care, BSC)‘[3].</p> <p>Damit geht auch der G-BA davon aus, dass es keine Hinweise darauf gibt, dass die Grundversorgung sich unterschiedlich auf die Behandlungsarme verteilt. Der G-BA spricht auch nicht von einem hieraus abzuleitenden Verzerrungspotential.</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p><u>Vorgeschlagene Änderung:</u> Alle Patienten erhielten eine Grundversorgung nach internationalen Behandlungsstandards (Best Supportive Care, BSC). Das endpunktspezifische Verzerrungspotenzial für die Ergebnisse zum kombinierten Endpunkt Zeit bis Tod oder dauerhafte Beatmung, zum Endpunkt Erreichen motorischer Meilensteine sowie zu den Endpunkten SUEs, schwere UEs und Therapieabbrüche wegen UEs wird als niedrig bewertet.</p>	
4.3.4 Ergebnisse zu patientenrelevanten Endpunkten			
Biogen	S. 20 (5.)	<p>Vorbericht IQWiG 5. Einordnung des Arbeitsergebnisses <i>Um die grundsätzlich relevanten Daten aus der Studie NURTURE in der Nutzenbewertung berücksichtigen zu können, wurden beim Hersteller Daten zu einem Vergleich dieser präsymptomatisch therapierten Kinder mit symptomatisch therapierten Kindern (zum Beispiel aus den Studien ENDEAR und CHERISH) angefragt, in dem oben angeführte sowie weitere Aspekte berücksichtigt werden (siehe Details in Abschnitt A3.1.1.3). Der Hersteller hat diese Daten für November 2019 in Aussicht gestellt.</i></p> <p><u>Anmerkung:</u> Ein Vergleich präsymptomatisch therapierter Kinder aus der NURTURE Studie mit symptomatisch therapierten Kindern aus den Studien ENDEAR und CHERISH sind im Anhang ‚Tabellenvorlage zum Vergleich präsymptomatischer vs. symptomatischer Therapiebeginn‘ beigefügt. Bei dem Vergleich von Patienten der NURTURE-Studie mit präsymptomatischem Therapiebeginn und 2 SMN2-Kopien gegenüber den Patienten der Studie ENDEAR mit symptomatischem Therapiebeginn und 2 SMN2-Kopien zeigten sich signifikante Ergebnisse in der Motorik als auch bei der Sicherheit zugunsten der präsymptomatisch behandelten Kinder.</p>	<p>Kenntnisnahme. Es erfolgte eine Anpassung im Bericht dahingehend, dass die nachgelieferten Daten im Bericht dargestellt wurden und – sofern sie hierfür geeignet waren – zur Nutzenableitung herangezogen wurden.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Es zeigte sich eine signifikante Verbesserung der motorischen Funktionsfähigkeit gemessen anhand des HINE-2 (Subskala 2) ($p < 0,0001$). Eine signifikant geringere Anzahl an Kindern erlitt ≥ 1 SUE ($p = 0,045$) sowie signifikant weniger Kinder erfuhren ≥ 1 schweres UE ($p = 0,002$). Es zeigte sich somit ein Zusatznutzen einer Behandlung vor Auftreten erster SMA-Symptome.</p> <p>Vergleicht man NURTURE-Patienten mit CHERISH-Patienten, die aufgrund verspätet einsetzender SMA-Symptome später behandelt worden sind, zeigen sich signifikante Ergebnisse hinsichtlich der motorischen Funktion bei Kindern ≥ 2 Jahre.</p> <p>Der Hammersmith Functional Motor Scale Expanded (HFMSSE) erfasst mit 33 Items die Veränderung motorischer Funktionen und somit die klinische Progression der SMA. Anhand des HFMSSE zeigt sich, dass Kinder mit präsymptomatischem Therapiebeginn und 3 <i>SMN2</i>-Kopien eine signifikante Verbesserung ihrer Motorik im Vergleich zu Kindern mit symptomatischem Therapiebeginn und 3 <i>SMN2</i>-Kopien zeigen ($p < 0,05$). Die Anzahl Patienten mit mindestens 1 SUE, mindestens 1 schweren UE, Abbruch wegen UE oder mit Rückenschmerzen unterscheidet sich zwischen den Gruppen nicht.</p> <p><u>Vorgeschlagene Änderung:</u> Eine präsymptomatische Behandlung der Kinder mit 3 <i>SMN2</i>-Kopien gegenüber symptomatisch behandelten Kindern und 3 <i>SMN2</i>-Kopien führt zu einer erheblichen Verbesserung ihrer motorischen Fähigkeiten/Funktion und geht nicht mit einem erhöhten Risiko an Nebenwirkungen einher.</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
4.3.4.4 Ergebnisse zum Erreichen motorischer Meilensteine (Berücksichtigung Fragebogen CHOP INTEND)			
Avexis	Abschnitt 4.3.4.4 (S. 12)	<p><u>Anmerkung:</u> Das IQWiG führt aus, dass in der Studie ENDEAR zu dem Endpunkt „Erreichen motorischer Meilensteine“ zwei verschiedene Instrumente, der HINE-2 und der CHOP INTEND zum Einsatz kamen. Für beide Instrumente sieht das IQWiG keine abschließende Validierung und keine Validierung einer MID („minimal important difference“). Dennoch wurde der HINE-2 inklusive Responderanalysen berücksichtigt, der CHOP INTEND jedoch ausgeschlossen, da es sich laut IQWiG um eine andere Operationalisierung desselben Endpunkts handele. Zu diesen Aussagen und dem resultierenden Vorgehen des IQWiG ist auf verschiedener Ebene Stellung zu nehmen.</p> <p>1. Erreichen motorischer Meilensteine Patienten mit SMA sind in ihrer motorischen Entwicklung stark eingeschränkt. Das Spektrum reicht von Patienten, die niemals sitzen lernen, zu Patienten, die erst im Erwachsenenalter Einschränkungen ihrer motorischen Funktion bis zum Verlust der Gehfähigkeit erleiden. Dementsprechend ist über das gesamte Krankheitsspektrum hinweg das Erreichen oder auch der Verlust von motorischen Meilensteinen ein wesentliches Kriterium, das auch als Endpunkt in entsprechenden klinischen Studien abgebildet sein sollte. Für die Erfassung motorischer Meilensteine gibt es verschiedene Möglichkeiten der Operationalisierung:</p> <ul style="list-style-type: none"> • Definition der Meilensteine nach WHO (World Health Organization) • HINE-2 (Hammersmith Infant Neurological Examination, Subskala 2: Motorische Meilensteine) • Definition der Meilensteine nach Bayley-Scale 	<p>Keine Änderung. Begründung:</p> <ul style="list-style-type: none"> ▪ In Bezug auf die Studie ENDEAR liefert der CHOP INTEND nur unwesentlich zusätzliche Informationen über die des HINE-2 hinaus. Zum einen bilden die Items des CHOP INTEND eine feingliedrigere Motorik ab als der HINE-2. Es zeigten sich unter Therapie bei den im HINE-2 abgebildeten motorischen Meilensteinen jedoch bereits Effekte, sodass eine Betrachtung von detaillierteren motorischen Funktionen verzichtbar scheint. Des Weiteren liegen die Ergebnisse zum CHOP INTEND für die berichtsrelevanten Subgruppen nur in Form von Responderanalysen bzw. MW-Änderungen bezogen auf den Gesamtscore vor, nicht aber für die einzelnen Items, die den Gesamtscore weiter erklären könnten. Dass die Ergebnisse zum Gesamtscore in dieselbe Richtung weisen wie die herangezogenen Ergebnisse des HINE-2, wurde im Vorbericht bereits angemerkt. ▪ Es ist gut möglich, dass in zukünftigen Bewertungen der CHOP INTEND Berücksichtigung findet, wenn keine anderen geeigneten Daten zum Endpunkt Erreichen motorischer Meilensteine vorliegen. ▪ Die herausragende Bedeutung des Erreichens eines Meilensteins (wie z. B. freies Stehen) wird geteilt. Für den angefragten Vergleich prä- versus frühsymptomatischen Therapiebeginn wurden deshalb beim Hersteller Ergebnisse zum (vollständigen) Erreichen eines Meilensteins (wie z. B. freies Stehen) angefragt und von diesem zur Verfügung gestellt.

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p><i>WHO</i> Motorische Meilensteine sind wichtige Bewertungskriterien, nicht nur bei Patienten mit SMA, sondern auch in der normalen Entwicklung von Kindern. Dementsprechend hat die World Health Organization (WHO) Definitionen motorischer Meilenstein festgelegt, die zur Bewertung herangezogen werden können und für die aus Studien Altersspannen vorliegen, in welchem Alter diese Meilensteine im Verlauf der normalen Entwicklung erwartet werden [7-9]. Unabhängiges Sitzen wird von der WHO definiert als aufrechtes Sitzen mit aufrechter Kopfhaltung für mindestens 10 Sekunden, das Kind verwendet weder die Arme noch die Hände als Unterstützung [7].</p> <p><i>HINE-2</i> Im HINE-2 ist keine genaue Definition des für SMA Typ 1 wesentlichen Meilensteins des unabhängigen Sitzens direkt abgebildet. Das Messinstrument wurde nicht für SMA Typ 1 entwickelt und validiert, es konnte anhand des HINE-2 jedoch gezeigt werden, dass motorische Meilensteine bei Patienten mit SMA Typ 1 im Prinzip nicht erreicht werden [10]. Ein Responsekriterium kann aus diesen Beobachtungen jedoch nicht abgeleitet werden, zumal für die Subskala 2 gar keine Punktwertung vorgesehen ist (Asymmetrie). Entsprechend verzichtet auch der G-BA in der Nutzenbewertung zu Nusinersen auf eine Darstellung der Responderanalyse im Beschluss [11]. Behelfsweise wird von den Autoren der Studienpublikation das Erreichen der Kategorien „stable sit“ und „pivots“ als das Erreichen des Meilensteins „independent sitting“ angenommen [12], dieses Vorgehen wird von AveXis als adäquat eingeschätzt.</p> <p><i>Bayley-Scale</i> Ein validiertes Messinstrument zur Bewertung der motorischen Funktion ist der Bayley-Scale [13]. Im Rahmen der Bewertung der motorischen Fähigkeiten wird mit Item #26 u. a. auch das unabhängige Sitzen (≥ 30 Sekunden) als motorischer Meilenstein bewertet. Der Bayley-Scale wurde auch vom G-</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>BA in anderen Nutzenbewertung als validiertes und patientenrelevantes Messinstrument zur Bewertung der motorischen Funktion bei kleinen Kindern eingeschätzt [14, 15].</p> <p>2. CHOP INTEND</p> <p>Entgegen der Aussagen des IQWiG handelt es sich beim CHOP INTEND um ein validiertes Messinstrument zur Bewertung der motorischen Funktion bei Kindern mit SMA Typ 1 [16-18]. Es handelt sich hierbei nicht um ein Messinstrument, das das Erreichen eines weiten Spektrums motorischer Meilensteine abbildet (Rollen und Kopfkontrolle werden erfasst), sondern um ein Messinstrument zur Bewertung der motorischen Funktion von Patienten mit SMA Typ 1, die im natürlichen Verlauf weiter fortgeschrittene motorische Meilensteine wie freies Sitzen, Stehen, Gehen gar nicht erreichen. Der CHOP INTEND ist das einzige krankheitsspezifische Messinstrument für die Erfassung der motorischen Funktion bei Patienten mit SMA Typ 1 und sollte somit zusätzlich zu motorischen Meilensteinen bei Patienten mit infantiler SMA in jedem Fall berücksichtigt werden. Auch der G-BA berücksichtigt den CHOP INTEND bei der Bewertung von Nusinersen zusätzlich zur Auswertung des HINE-2 [11, 19].</p> <p>Es ist richtig, dass für den CHOP INTEND keine validierte MID vorliegt, so dass Responderanalysen mit einer Verbesserung von 4 Punkten als Responsekriterium eine eingeschränkte Aussagekraft haben, aber es ist auch hier die besondere Situation zu würdigen: Da es sich um das einzige validierte Messinstrument für Patienten mit SMA Typ 1 handelt, liegen für diese Patientengruppen Referenzwerte bzw. Beobachtungen aus dem natürlichen Verlauf vor. Im natürlichen Verlauf wird ab Eintreten der Symptomatik im Wesentlichen keine Verbesserung des CHOP INTEND mehr beobachtet [20], sondern eine kontinuierliche Abnahme. Basierend auf dieser Beobachtung ist jeder Trend in die andere Richtung, d. h. jede Verbesserung als patientenrelevant zu</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>betrachten, woraus nachvollziehbar eine Responseschwelle von 4 Punkten festgelegt wurde. Eine bessere Bewertung lässt im vorliegenden Krankheitsbild das Erreichen bestimmter Schwellenwerte zu: CHOP INTEND-Werte von 40 werden von Kindern mit SMA Typ 1 fast nie erreicht [20] und wenn, dann erfolgt unweigerlich später eine Abnahme des Wertes. Ein Erreichen von Werten, die 50 Punkte überschreiten, wird bereits mit einer mildereren Form der Erkrankung (SMA Typ 2) assoziiert [21]. Der Maximalwert beträgt 64 Punkte.</p> <p><u>Vorgeschlagene Änderung:</u> Das Heranziehen des HINE-2 wird als sachgerechtes Vorgehen eingeschätzt, wenn in einer Studie keine andere Erfassung des Erreichens motorischer Meilensteine stattgefunden hat. Entsprechend dem Vorgehen von Finkel et al. 2017 sollten statt der Responderanalysen die Anteile an Patienten mit Erreichen eines motorischen Meilensteins für die Bewertung herangezogen werden.</p> <p>Der CHOP INTEND ist ein validiertes Messinstrument, das eigens zur Bewertung der motorischen Funktion von Patienten mit SMA Typ 1 entwickelt wurde. Der CHOP INTEND sowie das Erreichen bestimmter Schwellenwerte (40 Punkte im Fall von SMA Typ 1) sollte als zusätzlicher patientenrelevanter Endpunkt (motorische Funktion) aufgenommen werden.</p>	
Biogen	S. 12 (4.3.4.4) und S. 12 (4.3.4.4)	<p>Vorbericht IQWiG 4.3.4.4. Ergebnisse zum Erreichen motorischer Meilensteine</p> <p>Zu dem Endpunkt Erreichen motorischer Meilensteine lagen aus der Studie ENDEAR für die berichtsrelevanten Subgruppen Ergebnisse zu den Instrumenten HINE-2 (als primärem Endpunkt der Studie) und CHOP (Children's Hospital of Philadelphia Infant Test of Neuro-muscular Disorders) INTEND in unterschiedlichen Operationalisierungen vor. Die Instrumente wurden bislang beide nicht abschließend validiert. Darüber hinaus gibt es keine validierte Minimal important Difference (MID) für die Instrumente.</p>	<p>Keine Änderung.</p> <p>Begründung: siehe Würdigung <u>in vorangehender Zeile</u> (Stellungnahme Avexis)</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Vorbericht IQWiG 4.3.4.4. Ergebnisse zum Erreichen motorischer Meilensteine</p> <p><u>Von einer Darstellung des CHOP INTEND wird darüber hinaus abgesehen, da es sich um eine andere Operationalisierung desselben Endpunkts handelt. Die Ergebnisse weisen in dieselbe Richtung wie die des HINE-2.</u></p> <p>Anmerkung: Der HINE-2 und der CHOP INTEND sind als unabhängige patientenrelevante Bewertungsskalen der motorischen Funktion und Fähigkeit zu betrachten. Dies wird auch in der Nutzenbewertung durch den G-BA im Jahr 2017 deutlich, der für beide Instrumente auf Basis der standardisierten Mittelwertdifferenz einen Zusatznutzen vergibt.</p> <p>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) eine neurologische Untersuchung, (2) Beurteilung der motorischen Entwicklung und (3) Beurteilung des Verhaltens.</p> <p>Die Subskala 2 zur Beurteilung der Motorik besteht aus 8 Kategorien, die verschiedene motorische Meilensteine abbilden: (1) Bewusstes Greifen, (2) Fähigkeit in Rückenlage zu treten, (3) Kopfkontrolle, (4) Drehen, (5) Sitzen, (6) Krabbeln, (7) Stehen und (8) Gehen. In der Nutzenbewertung bewertet der G-BA den HINE als patientenrelevantes Instrument, dessen Validierung allerdings noch nicht vollständig abgeschlossen ist [3].</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Der HINE ist ein von Experten entwickeltes, etabliertes und krankheitsspezifisches Instrument. Die Kategorien, die im HINE abgefragt werden, werden von Experten als inhaltlich valide (Inhaltsvalidität) erachtet [4-6]. In einer Analyse von Bishop et al. wurde anhand von Säuglingen mit SMA untersucht, ob der HINE (Subskala 2) geeignet, zuverlässig und sensitiv auf Veränderungen ist. Über den Zeitverlauf konnten Veränderungen im HINE bei 16 von 19 Patienten in allen Domänen beobachtet werden. Verbesserungen im HINE korrelierten mit Änderungen in anderen neuromuskulären Endpunktparametern verbunden mit einer sehr guten Test-Retest-Reliabilität für den HINE ($R = 0,987$; $P < 0.0001$) [7][24]. Zudem existiert ein standardisiertes Trainingsprogramm zur richtigen Anwendung des Tests [8].</p> <p>Der CHOP INTEND wurde speziell für die Beurteilung der Entwicklung der Motorik von Säuglingen mit motorischer Schwäche, einschließlich Säuglingen mit SMA, entwickelt. Die Skala besteht aus 16 Items [9]. 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. Der CHOP INTEND erfasst daher zusätzliche und andere motorische Entwicklungen als der HINE.</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Der CHOP INTEND ist ein von Experten entwickeltes, etabliertes und krankheitsspezifisches Instrument [3]. In einer Arbeit von Glanzman et al. (2010) wurde gezeigt, dass das Instrument reliabel ist. Dies wurde anhand der Reliabilitätskriterien der Interrater-Reliabilität, der Test-Retest Reliabilität sowie der internen Konsistenz untersucht [10]. Zudem wurde die Validität des CHOP INTEND durch eine weitere Arbeit von Glanzman et al. (2011) gezeigt. In dieser wurde der CHOP INTEND inhaltlich als valide bewertet. Außerdem wurde gezeigt, dass er eine Übereinstimmungsvalidität (concurrent validity) aufweist [11].</p> <p>In kürzlich aufgesetzten und laufenden nationalen und internationalen SMA-Registern unter Leitung bzw. Supervision von klinischen Experten werden sowohl der HINE-2 als auch der CHOP INTEND nebeneinander eingesetzt. Dies unterstreicht die Bedeutung beider Instrumente in der Beurteilung der motorischen Fähigkeiten in der klinischen Praxis unabhängig voneinander.</p> <p><i>Die Beurteilung des G-BA 2017 fällt auch positiv aus und er erkennt folgendes an [3]: ‚Die Operationalisierung des Endpunkts ist nachvollziehbar. Der CHOP INTEND ist speziell zur Beurteilung motorischer Fähigkeiten bei Säuglingen mit SMA Typ I entwickelt worden.‘</i></p> <p>Stellt man nun beide Instrumente gegenüber, zeigen sich gravierende Unterschiede.</p> <p>Der CHOP INTEND erfasst anhand von 16 Items aktive Bewegungen und ausgelöste Reflexbewegungen, wohingegen der HINE-2 anhand von 8 Items die motorische Funktion hinsichtlich umfassender motorischer Meilensteine wie beispielsweise dem Sitzen beurteilt.</p>	

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Es wird anhand des CHOP INTEND eher die Beweglichkeit als die motorische Funktion im Gesamten erfasst. Es erfolgt beispielsweise die Extension der Knie. Kann das Kind sichtbar die Knie in einen 45° Winkel beugen? Dies wird anhand des HINE beispielsweise gar nicht erfasst. Mittels des HINE-2 wird aber wiederum die Fähigkeit zu Sitzen erfasst, was im CHOP INTEND nicht enthalten ist. Der CHOP INTEND ist spezifischer mit mehr Items, um auch feinere Veränderungen zu erkennen und von daher eine wichtige Ergänzung zu der HINE-Skala.</p> <p><u>Vorgeschlagene Änderung:</u> HINE-2 und CHOP INTEND sind validierte Instrumente zur Erfassung motorischer Fähigkeiten bei Patienten mit SMA. Die Instrumente HINE-2 und CHOP INTEND sind als unabhängige Messinstrumente zu betrachten und sind beide notwendig, um eine allumfassende Beschreibung des Bewegungsspektrums von SMA-Patienten zu erhalten.</p>	
Novartis	4.3.4.4 (S.12)	<p><u>Anmerkung:</u> Das IQWiG führt aus, dass in der eingeschlossenen Studie ENDEAR zwei Endpunkte erhoben wurden, die das Erreichen motorischer Meilensteine erfassen. Weiter wird beschrieben, dass beide Endpunkte bislang nicht abschliessend validiert wurden und es darüber hinaus keine validierte MID für beide Instrumente gibt.</p> <p>Dennoch wird der HINE-2 in der Nutzenbewertung berücksichtigt und der CHOP-INTEND von der Bewertung ausgeschlossen. Dieses Vorgehen ist für die Novartis Pharma GmbH nicht nachvollziehbar. Der CHOP INTEND ist ein zuverlässiges, einfach umzusetzendes und gut verträgliches Instrument um die motorischen Fähigkeiten im Besonderen von SMA Typ 1 und ähnlich schwachen Säuglinge mit neuromuskulären Erkrankungen zu messen und ist ein wichtiger Parameter in klinischen Studien (2).</p>	<p>Keine Änderung.</p> <p>Begründung: siehe Würdigung <u>in vorangehenden Zeilen</u> (Stellungnahme Avexis)</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>Zusätzlich beschreibt der G-BA den Endpunkt in der Nutzenbewertung zu Nusinersen als etabliertes und krankheitsspezifisches Instrument und bezieht die Ergebnisse des Endpunktes unterstützend zum primären Endpunkt der ENDEAR Studie (HINE Responder) mit in die Nutzenbewertung bei Nusinersen mit ein (3).</p> <p><u>Vorgeschlagene Änderung:</u></p> <p>Der CHOP-INTEND, welcher speziell bei Patienten mit SMA Typ-1, seine Anwendung findet und die Erreichung motorischer Meilensteine erfasst, sollte mit in die Nutzenbewertung einfließen, da er eine sinnvolle Ergänzung zum HINE-2 darstellt.</p>	
4.3.4.4 Ergebnisse zum Erreichen motorischer Meilensteine (nachgelieferte Daten zu Einzelitems des HINE-2)			
Biogen	S. 12 (4.3.4.4)	<p><i>Vorbericht IQWiG 4.3.4.4. Ergebnisse zum Erreichen motorischer Meilensteine</i></p> <p><i>Des Weiteren wurden Ergebnisse zu Responderanalysen für die erfassten einzelnen Items des HINE-2 dargestellt. Diese lagen allerdings nur auf Basis des Efficacy Sets vor</i></p> <p><u>Anmerkung:</u></p> <p>Die Ergebnisse der grafischen Darstellung der erfassten einzelnen Items des HINE-2 auf Basis der ITT sind im Anhang 'Tabellenvorlage zum Vergleich präsymptomatischer vs. symptomatischer Therapiebeginn' beigefügt.</p> <p>Es zeigt sich durchgehend für alle Einzelitems des HINE-2, dass eine präsymptomatische Behandlung bei 2 SMN2-Kopien bei der Mehrheit der Kinder zum Erreichen der Meilensteine führt, wohingegen symptomatisch behandelte Kinder die Meilensteine meist nicht erreichen:</p>	<p>Kenntnisnahme.</p> <p>Es erfolgte eine Anpassung im Bericht dahingehend, dass die nachgelieferten Daten – sofern sie hierfür geeignet waren – im Bericht dargestellt und zur Nutzenableitung herangezogen wurden.</p>

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<ul style="list-style-type: none"> ▪ Dreht sich im Sitzen: 14 von 15 präsymptomatisch behandelten Kindern, 1 von 34 symptomatisch behandelten Kindern ▪ Krabbeln: 11 von 15 präsymptomatisch behandelten Kindern, 0 von 34 symptomatisch behandelten Kindern ▪ Freies Stehen: 9 von 15 präsymptomatisch behandelten Kindern, 0 von 34 symptomatisch behandelten Kindern ▪ Freies Gehen: 7 von 15 präsymptomatisch behandelten Kindern, 0 von 34 symptomatisch behandelten Kindern <p><u>Vorgeschlagene Änderung:</u> Ein Vorteil eines präsymptomatischen Therapiebeginns zeigt sich auch auf der Ebene der einzelnen Items des HINE-2.</p>	
5 Einordnung des Arbeitsergebnisses			
Biogen	S. 19 (5.)	<p><i>Vorbericht IQWiG 5. Einordnung des Arbeitsergebnisses</i> <i>Es bleiben auch viele weitere Fragen unbeantwortet. Bislang ist beispielsweise ungeklärt, wie lange die Behandlung fortgesetzt werden sollte. Die Verabreichung bei Patientinnen und Patienten mit Skoliose [23] oder Spondylodese [38] gestaltet sich zudem häufig schwierig und bedarf häufig einer Steuerung mithilfe bildgebender Verfahren mit Kontrastmittel, Anästhesie und meist auch einer stationären Überwachung [39,40]. Auch werden mögliche Auswirkungen der mit der intrathekalen Verabreichung einhergehenden wiederholten Anästhesien auf die Hirnentwicklung von Kleinkindern diskutiert [41-43]. Ferner steigt bei einer Steuerung der Lumbalpunktionen mittels radiologischer Verfahren das Krebsrisiko [38,44,45].</i></p> <p><u>Anmerkung:</u> Der Ausschuss für Humanarzneimittel (CHMP) der Europäischen Arzneimittel-Agentur (EMA) stuft das Nutzen-Risiko-Verhältnis auch nach Einreichung von regelmäßig aktualisierten Unbedenklichkeitsberichten von Nusinersen als</p>	Kenntnisnahme. Keine Änderung.

Stellungnehmender	Bezug (Abschnitt / Seite im Vorbericht)	Wortlaut der Stellungnahme und des Änderungsvorschlags	Würdigung der Stellungnahme und Konsequenz
		<p>unverändert ein. Die Daten der noch laufenden Studie SHINE zeigen, dass mittlerweile Patienten mit einer Behandlungsdauer im Mittel von 1084 Tagen (960 Tage bei Patienten mit infantiler SMA und 1162 Tage bei Patienten mit später einsetzender SMA) auch bei langfristiger Behandlung keine nachteiligen Effekte erfahren.</p> <p><u>Vorgeschlagene Änderung:</u> Das CHMP stuft das Nutzen-Risiko-Verhältnis auch nach den regelmäßig vorzulegenden Unterlagen zur Sicherheit von Nusinersen als unverändert ein. Weitere Ergebnisse zur Langzeitsicherheit und -wirksamkeit von Nusinersen werden nach weiterer Beobachtung in der Studie SHINE vorliegen.</p>	
<p>Pilotprojekt Neugeborenen-screening SMA / Deutsche Gesellschaft für Kinder- und Jugendmedizin^a</p>	<p>S.21 erster Abschnitt</p>	<p>Da der Zweck der genetischen Reihenuntersuchung auf SMA die Identifikation von erkrankten Kindern und nicht von einer Anlageträgerschaft ist, steht der Nachweis einer Anlageträgerschaft für SMA dieser genetischen Reihenuntersuchung nicht entgegen. Die Gesetzesbegründung im GenDG bezieht sich auf genetische Reihenuntersuchungen mit dem Zweck, Anlageträgerschaften zu identifizieren, die sich bei den Untersuchten selbst gesundheitlich nicht auswirken (Heterozygotenscreening, z.B. β-Thalassämie-Screening bei Erwachsenen). Dies ist nach §16 GenDG nicht zulässig</p> <p><u>Vorgeschlagene Änderung:</u> Wir schlagen vor, den ersten Absatz auf S.21 [„Mit manchen Testverfahren Beachtung finden“] zu streichen.</p>	<p>Keine Änderung. Begründung: Die Auffassung der Stellungnehmenden wird geteilt, dass allein die Möglichkeit des Nachweises einer Anlageträgerschaft für SMA einer genetischen Reihenuntersuchung nicht entgegensteht. Diese Absicht wird mit dem Passus im Bericht auch nicht verfolgt. Vielmehr hat er informativen Charakter, um darauf hinzuweisen, dass bei einer Einführung eines Neugeborenen-screenings auf SMA in Deutschland die Identifikation von heterozygoten Trägern und die Verarbeitung dieser Information theoretisch möglich wären, aber nicht erlaubt wären.</p>
<p>a: Die Stellungnahme des Pilotprojekts Neugeborenen-screening SMA und die Stellungnahme der Deutschen Gesellschaft für Kinder- und Jugendmedizin waren wortgleich.</p>			
<p>BSC: Best supportive Care; CHMP: Ausschuss für Humanarzneimittel; RCT: randomisierte kontrollierte Studie; SMA: spinale Muskelatrophie</p>			

A5 Literatur

1. Verhaart IEC, Robertson A, Wilson IJ, Aartsma-Rus A, Cameron S, Jones CC et al. Prevalence, incidence and carrier frequency of 5q-linked spinal muscular atrophy: a literature review. *Orphanet J Rare Dis* 2017; 12(1): 124.
2. Orphanet. Spinale Muskelatrophie, proximale, Typ 1 [online]. 07.2009 [Zugriff: 25.01.2019]. URL: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=DE&Expert=83330.
3. Orphanet. Spinale Muskelatrophie, proximale [online]. 07.2009 [Zugriff: 12.02.2019]. URL: https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=DE&Expert=70.
4. Prior TW. Perspectives and diagnostic considerations in spinal muscular atrophy. *Genet Med* 2010; 12(3): 145-152.
5. Lefebvre S, Burglen L, Reboullet S, Clermont O, Burlet P, Viollet L et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 1995; 80(1): 155-165.
6. Wang CH, Finkel RS, Bertini ES, Schroth M, Simonds A, Wong B et al. Consensus statement for standard of care in spinal muscular atrophy. *J Child Neurol* 2007; 22(8): 1027-1049.
7. Taylor JL, Lee FK, Yazdanpanah GK, Staropoli JF, Liu M, Carulli JP et al. Newborn blood spot screening test using multiplexed real-time PCR to simultaneously screen for spinal muscular atrophy and severe combined immunodeficiency. *Clin Chem* 2015; 61(2): 412-419.
8. Dobrowolski SF, Pham HT, Downes FP, Prior TW, Naylor EW, Swoboda KJ. Newborn screening for spinal muscular atrophy by calibrated short-amplicon melt profiling. *Clin Chem* 2012; 58(6): 1033-1039.
9. Michelson D, Ciafaloni E, Ashwal S, Lewis E, Narayanaswami P, Oskoui M et al. Evidence in focus: nusinersen use in spinal muscular atrophy; report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology* 2018; 91(20): 923-933.
10. Dubowitz V. Chaos in the classification of SMA: a possible resolution. *Neuromuscul Disord* 1995; 5(1): 3-5.
11. Arnold WD, Kassar D, Kissel JT. Spinal muscular atrophy: diagnosis and management in a new therapeutic era. *Muscle Nerve* 2015; 51(2): 157-167.
12. Munsat TL, Davies KE. International SMA Consortium Meeting (26-28 June 1992, Bonn, Germany): meeting report. *Neuromuscul Disord* 1992; 2(5-6): 423-428.
13. Russman BS. Spinal muscular atrophy: clinical classification and disease heterogeneity. *J Child Neurol* 2007; 22(8): 946-951.

14. Chien YH, Chiang SC, Weng WC, Lee NC, Lin CJ, Hsieh WS et al. Presymptomatic diagnosis of spinal muscular atrophy through newborn screening. *J Pediatr* 2017; 190: 124-129.e1.
15. Lin CW, Kalb SJ, Yeh WS. Delay in diagnosis of spinal muscular atrophy: a systematic literature review. *Pediatr Neurol* 2015; 53(4): 293-300.
16. Deutsches Institut für Medizinische Dokumentation und Information. G12.-: spinale Muskelatrophie und verwandte Syndrome [online]. In: ICD-10-GM Version 2019. 21.09.2018 [Zugriff: 17.05.2019]. URL: <https://www.dimdi.de/static/de/klassifikationen/icd/icd-10-gm/kode-suche/htmlgm2019/block-g10-g14.htm#G12>.
17. Deutsches Institut für Medizinische Dokumentation und Information. Datenbankeintrag zu: Spinraza 12mg Injektionslösung; Biogen; Ampullen [online]. In: ABDA-Fertigarzneimittel. [Zugriff: 27.11.2018]. URL: <https://www.dimdi.de/dynamic/de/arzneimittel/arzneimittel-recherchieren/abda-datenbank/datenbankinformation-abda-fertigarzneimittel/>.
18. European Medicines Agency. Spinraza: European public assessment report; product information [Deutsch] [online]. 27.12.2018 [Zugriff: 25.01.2019]. URL: https://www.ema.europa.eu/documents/product-information/spinraza-epar-product-information_de.pdf.
19. Hoffmann-La Roche. A study to evaluate long term safety, tolerability, and effectiveness of olesoxime in patients with spinal muscular atrophy (SMA): study details [online]. In: ClinicalTrials.gov. 30.10.2018 [Zugriff: 06.12.2018]. URL: <https://clinicaltrials.gov/ct2/show/NCT02628743>.
20. AveXis. Single-dose gene replacement therapy clinical trial for patients with spinal muscular atrophy type 1 (STRIVE-EU): study details [online]. In: ClinicalTrials.gov. 01.10.2018 [Zugriff: 06.12.2018]. URL: <https://clinicaltrials.gov/ct2/show/study/NCT03461289>.
21. Food and Drug Administration. FDA approves innovative gene therapy to treat pediatric patients with spinal muscular atrophy, a rare disease and leading genetic cause of infant mortality [online]. In: ClinicalTrials.gov. 24.05.2019 [Zugriff: 18.06.2019]. URL: <https://www.fda.gov/news-events/press-announcements/fda-approves-innovative-gene-therapy-treat-pediatric-patients-spinal-muscular-atrophy-rare-disease>.
22. Messina S. New directions for SMA therapy. *J Clin Med* 2018; 7(9): 251.
23. Vill K, Blaschek A, Schara U, Kölbel H, Hohenfellner K, Harms E et al. Spinale Muskelatrophie: Zeit für das Neugeborenencreening? *Nervenarzt* 2017; 88(12): 1358-1366.
24. Gemeinsamer Bundesausschuss. Richtlinien des Bundesausschusses der Ärzte und Krankenkassen über die Früherkennung von Krankheiten bei Kindern bis zur Vollendung des 6. Lebensjahres („Kinder-Richtlinien“) [online]. 16.12.2010 [Zugriff: 05.09.2013]. URL: http://www.g-ba.de/downloads/62-492-506/RL_Kinder_2010-12-16.pdf.

25. Ionis Pharmaceuticals. 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: study ISIS 396443-CS3B; clinical study report [unveröffentlicht]. 2017.
26. Biogen. Nusinersen (Spinraza): Dossier zur Nutzenbewertung gemäß § 35a SGB V; Modul 4 A; 5q-assozierte spinale Muskelatrophie (SMA) [online]. 30.06.2017 [Zugriff: 04.07.2019]. URL: https://www.g-ba.de/downloads/92-975-1991/2017-06-30_Modul4A_Nusinersen.pdf.
27. Biogen. Anhang zur Stellungnahme: Vergleich präsymptomatischer vs. symptomatischer Therapiebeginn. [Demnächst verfügbar unter: <https://www.iqwig.de/de/projekte-ergebnisse/projekte/nichtmedikamentose-verfahren/s-projekte/s18-02-neugeborenencreening-auf-spinale-muskelatrophie-sma.10782.html#documents> im Dokument "Dokumentation der Anhörung zum Vorbericht"].
28. 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: study 232SM201; clinical study report [unveröffentlicht]. 2019.
29. De Vivo DC, Bertini E, Swoboda KJ, Hwu WL, Crawford TO, Finkel RS et al. Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: interim efficacy and safety results from the phase 2 NURTURE study. *Neuromuscul Disord* 2019; 29(11): 842-856.
30. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 5.0. Köln: IQWiG; 2017. URL: <https://www.iqwig.de/download/Allgemeine-Methoden-Version-5-0.pdf>.
31. Czibere L, Burggraf S, Fleige T, Glück B, Keitel LM, Landt O et al. High-throughput genetic newborn screening for spinal muscular atrophy by rapid nucleic acid extraction from dried blood spots and 384-well qPCR. *Eur J Hum Genet* 30.07.2019 [Epub ahead of print].
32. Vill K, Kölbl H, Schwartz O, Blaschek A, Olgemöller B, Harms E et al. One year of newborn screening for SMA: results of a German pilot project. *Journal of neuromuscular diseases* 2019; 6(4): 503-515.
33. Kariyawasam DST, Russell JS, Wiley V, Alexander IE, Farrar MA. The implementation of newborn screening for spinal muscular atrophy: the Australian experience. *Genet Med* 14.10.2019 [Epub ahead of print].
34. Kraszewski JN, Kay DM, Stevens CF, Koval C, Haser B, Ortiz V et al. Pilot study of population-based newborn screening for spinal muscular atrophy in New York state. *Genet Med* 2018; 20(6): 608-613.

35. Gemeinsamer Bundesausschuss. Nusinersen: Nutzenbewertung 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 [online]. 02.10.2017 [Zugriff: 12.07.2019]. URL: https://www.g-ba.de/downloads/92-975-1992/2017-07-01_Nutzenbewertung-G-BA_Nusinersen_D-294.pdf.
36. Finkel RS, Mercuri E, Darras BT, Connolly AM, Kuntz NL, Kirschner J et al. Supplementary appendix to "Finkel RS, Mercuri E, Darras BT et al. Nusinersen versus sham control in infantile-onset spinal muscular atrophy. N Engl J Med 2017;377: 1723-32" [online]. [Zugriff: 20.11.2018]. URL: https://www.nejm.org/doi/suppl/10.1056/NEJMoa1702752/suppl_file/nejmoa1702752_appendix.pdf.
37. Prakash V. Spinraza: a rare disease success story. Gene Ther 2017; 24(9): 497.
38. Biogen. A study for participants with spinal muscular atrophy (SMA) who previously participated in nusinersen (ISIS 396443) investigational studies (SHINE): study details [online]. In: ClinicalTrials.gov. 30.01.2019 [Zugriff: 13.05.2019]. URL: <https://clinicaltrials.gov/ct2/show/NCT02594124>.
39. Mercuri E, Darras BT, Chiriboga CA, Day JW, Campbell C, Connolly AM et al. Nusinersen versus sham control in later-onset spinal muscular atrophy. N Engl J Med 2018; 378(7): 625-635.
40. Biogen. A study to assess the efficacy and safety of nusinersen (ISIS 396443) in participants with later-onset spinal muscular atrophy (SMA) (CHERISH): study details [online]. In: ClinicalTrials.gov. 21.02.2018 [Zugriff: 14.06.2019]. URL: <https://clinicaltrials.gov/ct2/show/NCT02292537>.
41. Ionis Pharmaceuticals. 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 [online]. In: EU Clinical Trials Register. [Zugriff: 14.06.2019]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-001947-18.
42. Ionis Pharmaceuticals. 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]: study ISIS 396443-CS4; clinical study report synopsis [online]. In: Biogen Clinical Trial Results. 06.06.2017 [Zugriff: 14.06.2019]. URL: <http://clinicalresearch.biogen.com/Content/Studies/CS4%20Biogen.com%20Packet.pdf>.
43. Biogen. A study to assess the safety and tolerability of nusinersen (ISIS 396443) in participants with spinal muscular atrophy (SMA) (EMBRACE): study details [online]. In: ClinicalTrials.gov. 16.01.2019 [Zugriff: 14.06.2019]. URL: <https://clinicaltrials.gov/ct2/show/NCT02462759>.

44. Biogen Idec Research. 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 [online]. In: EU Clinical Trials Register. [Zugriff: 14.06.2019]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2014-003657-33.
45. Stolte B, Totzeck A, Kizina K, Bolz S, Pietruck L, Monninghoff C et al. Feasibility and safety of intrathecal treatment with nusinersen in adult patients with spinal muscular atrophy. *Ther Adv Neurol Disord* 2018; 11: 1756286418803246.
46. Lorenz HM, Kühnle I, Edler J, Hobbiebrunken E, Wilichowski E, Tsaknakis K et al. Intrathekale Nusinersen-Therapie bei Kindern mit Spinaler Muskelatrophie und Wirbelsäulendeformitäten. *Klin Padiatr* 2018; 230(4): 231-233.
47. Gemeinsamer Bundesausschuss. Wirkstoff Nusinersen: mündliche Anhörung; 5. Kapitel § 19 Abs. 2 Verfahrensordnung des Gemeinsamen Bundesausschusses; stenografisches Wortprotokoll [online]. [Zugriff: 31.07.2019]. URL: https://www.g-ba.de/downloads/91-1031-298/2017_11_06_Wortprotokoll_Nusinersen_D-294.pdf.
48. Haché M, Swoboda KJ, Sethna N, Farrow-Gillespie A, Khandji A, Xia S et al. Intrathecal injections in children with spinal muscular atrophy: nusinersen clinical trial experience. *J Child Neurol* 2016; 31(7): 899-906.
49. Psaty BM, Platt R, Altman RB. Neurotoxicity of generic anesthesia agents in infants and children: an orphan research question in search of a sponsor. *JAMA* 2015; 313(15): 1515-1516.
50. Food and Drug Administration. FDA approves label changes for use of general anesthetic and sedation drugs in young children [online]. 27.04.2017 [Zugriff: 14.06.2019]. URL: <https://www.fda.gov/media/104705/download>.
51. Pearce MS, Salotti JA, Little MP, McHugh K, Lee C, Kim KP et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet* 2012; 380(9840): 499-505.
52. Wurster CD, Winter B, Wollinsky K, Ludolph AC, Uzelac Z, Witzel S et al. Intrathecal administration of nusinersen in adolescent and adult SMA type 2 and 3 patients. *J Neurol* 2019; 266(1): 183-194.
53. Bundesregierung. Gesetzentwurf der Bundesregierung: Entwurf eines Gesetzes über genetische Untersuchungen bei Menschen (Gendiagnostikgesetz – GenDG); Drucksache 16/10532 [online]. 13.10.2008 [Zugriff: 14.03.2019]. URL: <http://dip21.bundestag.de/dip21/btd/16/105/1610532.pdf>.
54. King NMP, Bishop CE. New treatments for serious conditions: ethical implications. *Gene Ther* 2017; 24(9): 534-538.

55. Glascock J, Sampson J, Haidet-Phillips A, Connolly A, Darras B, Day J et al. Treatment algorithm for infants diagnosed with spinal muscular atrophy through newborn screening. *J Neuromuscul Dis* 2018; 5(2): 145-158.
56. 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 spoSIS 396443-CS4: clinical trial results [online]. In: EU Clinical Trials Register. 03.05.2019 [Zugriff: 27.06.2019]. URL: <https://www.clinicaltrialsregister.eu/ctr-search/trial/2014-003657-33/results>.
57. Lowes LP, Alfano LN, Arnold WD, Shell R, Prior TW, McColly M et al. Impact of age and motor function in a phase 1/2a study of infants with SMA type 1 receiving single-dose gene replacement therapy. *Pediatr Neurol* 13.05.2019 [Epub ahead of print].
58. Nennstiel U, Hohenfellner K, Müller-Felber W, Röschinger W. Neugeborenen-Screening auf Cystinose und Spinale Muskelatrophie. Bayer Aztebl 2018; (3): 107.
59. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Neugeborenencreening auf 5q-assoziierte spinale Muskelatrophie: Dokumentation der Anhörung zum Vorbericht; Auftrag S18-02. Köln: IQWiG; 2020.
60. ICH Expert Working Group. ICH harmonised tripartite guideline: structure and content of clinical study reports; E3; current step 4 version [online]. 30.11.1995 [Zugriff: 05.09.2018]. URL: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf.
61. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ et al. CONSORT 2010: explanation and elaboration; updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c869.
62. Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004; 94(3): 361-366.
63. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.
64. Pepe MS, Alonzo TA. Comparing disease screening tests when true disease status is ascertained only for screen positives. *Biostatistics* 2001; 2(3): 249-260.
65. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med* 2003; 138(1): W1-12.

66. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155(8): 529-536.
67. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. *Methods for meta-analysis in medical research*. Chichester: Wiley; 2000.
68. Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. *Cochrane Database Syst Rev* 2015; (Suppl 1): 25-27.
69. Kuss O. Statistical methods for meta-analyses including information from studies without any events: add nothing to nothing and succeed nevertheless. *Stat Med* 2015; 34(7): 1097-1116.
70. Leemis LM, Trivedi KS. A comparison of approximate interval estimators for the Bernoulli parameter. *Am Stat* 1996; 50(1): 63-68.
71. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005; 58(10): 982-990.
72. Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006; 59(12): 1331-1332.
73. Menke J. Bivariate random-effects meta-analysis of sensitivity and specificity with SAS PROC GLIMMIX. *Methods Inf Med* 2010; 49(1): 54-64.
74. Hotelling H. The generalization of student's ratio. *Ann Math Stat* 1931; 2(3): 360-378.
75. Trikalinos TA, Trow P, Schmid CH. Simulation-based comparison of methods for meta-analysis of proportions and rates [online]. 11.2013 [Zugriff: 22.01.2019]. URL: https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0076998/pdf/PubMedHealth_PMH0076998.pdf.
76. Biogen. A study to assess the efficacy and safety of nusinersen (ISIS 396443) in infants with spinal muscular atrophy (ENDEAR): study details [online]. In: *ClinicalTrials.gov*. 28.07.2017 [Zugriff: 13.05.2019]. URL: <https://clinicaltrials.gov/ct2/show/NCT02193074>.
77. Isis Pharmaceuticals. 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 [online]. In: *EU Clinical Trials Register*. [Zugriff: 13.05.2019]. URL: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-004422-29.

78. Isis Pharmaceuticals. Eine randomisierte, doppelblinde, scheinbehandlungskontrollierte Phase-3-Studie zur Beurteilung der klinischen Wirksamkeit und Sicherheit von intrathekal verabreichtem ISIS 396443 bei Patienten mit infantiler spinaler Muskelatrophie [online]. In: Deutsches Register Klinischer Studien. [Zugriff: 13.05.2019]. URL: <http://www.drks.de/DRKS00009141>.
79. Ionis Pharmaceuticals. 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: study ISIS 396443-CS3B; clinical study report synopsis [online]. In: Biogen Clinical Trial Results. 10.02.2017 [Zugriff: 13.05.2019]. URL: <http://clinicalresearch.biogen.com/Content/Studies/CS3b%20Biogen.com%20Packet.pdf>.
80. Biogen. A study to assess the efficacy and safety of nusinersen (ISIS 396443) in infants with spinal muscular atrophy (ENDEAR): study results [online]. In: ClinicalTrials.gov. 28.07.2017 [Zugriff: 13.05.2019]. URL: <https://clinicaltrials.gov/ct2/show/results/NCT02193074>.
81. Isis Pharmaceuticals. 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: clinical trial results [online]. In: EU Clinical Trials Register. 15.07.2017 [Zugriff: 13.05.2019]. URL: <https://www.clinicaltrialsregister.eu/ctr-search/trial/2013-004422-29/results>.
82. National Taiwan University Hospital. Newborn screening for spinal muscular atrophy: study details [online]. In: ClinicalTrials.gov. 02.11.2016 [Zugriff: 13.05.2019]. URL: <https://clinicaltrials.gov/ct2/show/NCT02123186>.
83. 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: study ISIS 396443-CS3B; clinical study report erratum 3 [unveröffentlicht]. 2017.
84. Ionis Pharmaceuticals. 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: study ISIS 396443-CS3B; clinical study report erratum 4 [unveröffentlicht]. 2017.
85. Servais L, Farrar M, Finkel RS, Kirschner J, Muntoni F, Sun P et al. Nusinersen demonstrates greater efficacy in infants with shorter disease duration: final results from the ENDEAR study in infants with spinal muscular atrophy (SMA) [Poster]. 22nd International Annual Congress of the World Muscle Society; 03-07102017; Saint Malo, Frankreich.
86. 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: study ISIS 396443-CS3B; Zusatzanalysen [unveröffentlicht]. 2019.

87. Finkel RS, Mercuri E, Darras BT, Connolly AM, Kuntz NL, Kirschner J et al. Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *N Engl J Med* 2017; 377(18): 1723-1732.
88. Ionis Pharmaceuticals. A study to assess the efficacy, safety, tolerability, and pharmacokinetics of multiple doses of nusinersen (ISIS 396443) delivered intrathecally to patients with infantile-onset spinal muscular atrophy: study ISIS396443-CS3A; clinical study report [unveröffentlicht]. 2017.
89. Kaohsiung Medical University Chung-Ho Memorial Hospital. Neonatal spinal muscular atrophy (SMA) screening (SMA): study details [online]. In: *ClinicalTrials.gov*. 08.08.2018 [Zugriff: 13.05.2019]. URL: <https://clinicaltrials.gov/ct2/show/NCT03217578>.
90. Centre Hospitalier Régional de la Citadelle. Sun may arise on SMA: newborn screening of SMA in Belgium; study details [online]. In: *ClinicalTrials.gov*. 20.11.2018 [Zugriff: 13.05.2019]. URL: <https://clinicaltrials.gov/ct2/show/NCT03554343>.
91. Boemer F, Caberg JH, Dideberg V, Dardenne D, Bours V, Hiligsmann M et al. Newborn screening for SMA in Southern Belgium. *Neuromuscul Disord* 2019; 29(5): 343-349.
92. RTI International. Early check: expanded screening in newborns; study details [online]. In: *ClinicalTrials.gov*. 13.12.2018 [Zugriff: 13.05.2019]. URL: <https://clinicaltrials.gov/ct2/show/NCT03655223>.
93. Martín Andrés A, Silva Mato A. Choosing the optimal unconditioned test for comparing two independent proportions. *Computat Stat Data Anal* 1994; 17(5): 555-574.
94. Bartels B, Montes J, Van der Pol WL, De Groot JF. Physical exercise training for type 3 spinal muscular atrophy. *Cochrane Database Syst Rev* 2019; (3): CD012120.
95. Meylemans A, De Bleecker J. Current evidence for treatment with nusinersen for spinal muscular atrophy: a systematic review. *Acta Neurol Belg* 06.08.2019 [Epub ahead of print].
96. Calucho M, Bernal S, Alias L, March F, Vencesla A, Rodriguez-Alvarez FJ et al. Correlation between SMA type and SMN2 copy number revisited: an analysis of 625 unrelated Spanish patients and a compilation of 2834 reported cases. *Neuromuscul Disord* 2018; 28(3): 208-215.

A6 Studienlisten

A6.1 Liste der gesichteten systematischen Übersichten

A6.1.1 Vergleichende Interventionsstudien zum Therapiebeginn

1. Bartels B, Montes J, Van der Pol WL, De Groot JF. Physical exercise training for type 3 spinal muscular atrophy. *Cochrane Database Syst Rev* 2019; (3): CD012120.
2. Meylemans A, De Bleecker J. Current evidence for treatment with nusinersen for spinal muscular atrophy: a systematic review. *Acta Neurol Belg* 06.08.2019 [Epub ahead of print].
3. Michelson D, Ciafaloni E, Ashwal S, Lewis E, Narayanaswami P, Oskoui M et al. Evidence in focus: nusinersen use in spinal muscular atrophy; report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology* 2018; 91(20): 923-933.

A6.2 Liste der ausgeschlossenen Publikationen mit Ausschlussgründen

A6.2.1 Vergleichende Interventionsstudien der Screeningkette sowie Studien zur diagnostischen Güte

Nicht INS1/D1

1. Ar Rochmah M, Harahap NIF, Niba ETE, Nakanishi K, Awano H, Morioka I et al. Genetic screening of spinal muscular atrophy using a real-time modified COP-PCR technique with dried blood-spot DNA. *Brain Dev* 2017; 39(9): 774-782.
2. Arkblad E, Tulinius M, Kroksmark AK, Henricsson M, Darin N. A population-based study of genotypic and phenotypic variability in children with spinal muscular atrophy. *Acta Paediatr* 2009; 98(5): 865-872.
3. Bishop KM, Montes J, Finkel RS. Motor milestone assessment of infants with spinal muscular atrophy using the hammersmith infant neurological Exam-Part 2: experience from a nusinersen clinical study. *Muscle Nerve* 2018; 57(1): 142-146.
4. Cao YY, Zhang WH, Qu YJ, Bai JL, Jin YW, Wang H et al. Diagnosis of spinal muscular atrophy: a simple method for quantifying the relative amount of survival motor neuron gene 1/2 using sanger DNA sequencing. *Chin Med J (Engl)* 2018; 131(24): 2921-2929.
5. Chen WJ, Wu ZY, Wang N, Lin MT, Mu-Rong SX. Rapid diagnosis of spinal muscular atrophy using denaturing high-performance liquid chromatography. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2005; 22(3): 291-293.
6. Corcia P, Khoris J, Couratier P, Mayeux-Portas V, Bieth E, De Toffol B et al. SMN1 gene study in three families in which ALS and spinal muscular atrophy co-exist. *Neurology* 2002; 59(9): 1464-1466.
7. Evans MC, Cherry JJ, Androphy EJ. Differential regulation of the SMN2 gene by individual HDAC proteins. *Biochem Biophys Res Commun* 2011; 414(1): 25-30.

8. Feldkötter M, Schwarzer V, Wirth R, Wienker TF, Wirth B. Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am J Hum Genet* 2002; 70(2): 358-368.
9. Gitlin JM, Fischbeck K, Crawford TO, Cwik V, Fleischman A, Gonye K et al. Carrier testing for spinal muscular atrophy. *Genet Med* 2010; 12(10): 621-622.
10. Harahap NI, Harahap IS, Kaszynski RH, Nurputra DK, Hartomo TB, Pham HT et al. Spinal muscular atrophy patient detection and carrier screening using dried blood spots on filter paper. *Genet Test Mol Biomarkers* 2012; 16(2): 123-129.
11. Ibrahim S, Moatter T, Saleem AF. Spinal muscular atrophy: clinical spectrum and genetic mutations in Pakistani children. *Neurol India* 2012; 60(3): 294-298.
12. Kim CA, Passos-Bueno MR, Marie SK, Cerqueira A, Conti U, Marques-Dias MJ et al. Clinical and molecular analysis of spinal muscular atrophy in Brazilian patients. *Genet Mol Biol* 1999; 22(4): 487-492.
13. NIHR Horizon Scanning Research & Intelligence Centre. Nusinersen for paediatric-onset spinal muscular atrophy: first line. Birmingham: NIHR HSRIC; 2016.
14. Rashnonejad A, Onay H, Atik T, Atan Sahin O, Gokben S, Tekgul H et al. Molecular genetic analysis of survival motor neuron gene in 460 Turkish cases with suspicious spinal muscular atrophy disease. *Iranian Journal of Child Neurology* 2016; 10(4): 30-35.
15. Sivanesan S, Howell MD, Didonato CJ, Singh RN. Antisense oligonucleotide mediated therapy of spinal muscular atrophy. *Transl Neurosci* 2013; 4(1): 1-7.
16. Sumner CJ, Kolb SJ, Harmison GG, Jeffries NO, Schadt K, Finkel RS et al. SMN mRNA and protein levels in peripheral blood: biomarkers for SMA clinical trials. *Neurology* 2006; 66(7): 1067-1073.
17. Zappata S, Tiziano F, Neri G, Brahe C. Deletions in the SMN gene in infantile and adult spinal muscular atrophy patients from the same family. *Hum Genet* 1996; 97(3): 315-318.
18. Zarkov M, Stojadinovic A, Sekulic S, Barjaktarovic I, Peric S, Kekovic G et al. Association between the SMN2 gene copy number and clinical characteristics of patients with spinal muscular atrophy with homozygous deletion of exon 7 of the SMN1 gene. *Vojnosanit Pregl* 2015; 72(10): 859-863.

Nicht INS2

1. Canadian Agency for Drugs and Technologies in Health. Clinical review report: nusinersen (Spinraza); CADTH common drug review [online]. 01.2018 [Zugriff: 28.05.2019]. URL: https://www.ncbi.nlm.nih.gov/books/NBK533989/pdf/Bookshelf_NBK533989.pdf.
2. MacKenzie A. Sense in antisense therapy for spinal muscular atrophy. *N Engl J Med* 2012; 366(8): 761-763.

3. Mailman MD, Heinz JW, Papp AC, Snyder PJ, Sedra MS, Wirth B et al. Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. *Genet Med* 2002; 4(1): 20-26.
4. Michelson D, Ciafaloni E, Ashwal S, Lewis E, Narayanaswami P, Oskoui M et al. Evidence in focus: nusinersen use in spinal muscular atrophy; report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology* 2018; 91(20): 923-933.
5. Rose NC, Wick M. Current recommendations: screening for Mendelian disorders. *Semin Perinatol* 2016; 40(1): 23-28.

Nicht D2

1. Anhuf D, Eggermann T, Rudnik-Schoneborn S, Zerres K. Determination of SMN1 and SMN2 copy number using TaqMan technology. *Hum Mutat* 2003; 22(1): 74-78.
2. Arkblad EL, Darin N, Berg K, Kimber E, Brandberg G, Lindberg C et al. Multiplex ligation-dependent probe amplification improves diagnostics in spinal muscular atrophy. *Neuromuscul Disord* 2006; 16(12): 830-838.
3. Claborn MK, Stevens DL, Walker CK, Gildon BL. Nusinersen: a treatment for spinal muscular atrophy. *Ann Pharmacother* 2019; 53(1): 61-69.
4. Dobrowolski SF, Pham HT, Downes FP, Prior TW, Naylor EW, Swoboda KJ. Newborn screening for spinal muscular atrophy by calibrated short-amplicon melt profiling. *Clin Chem* 2012; 58(6): 1033-1039.
5. Eggermann T, Eggermann K, Elbracht M, Zerres K, Rudnik-Schöneborn S. A new splice site mutation in the SMN1 gene causes discrepant results in SMN1 deletion screening approaches. *Neuromuscul Disord* 2008; 18(2): 146-149.
6. Haider MZ, Moosa A, Dalal H, Habib Y, Reynold L. Gene deletion patterns in spinal muscular atrophy patients with different clinical phenotypes. *J Biomed Sci* 2001; 8(2): 191-196.
7. Hamzi K, Bellayou H, Itri M, Nadifi S. PCR-RFLP, sequencing, and quantification in molecular diagnosis of spinal muscular atrophy: limits and advantages. *J Mol Neurosci* 2013; 50(2): 270-274.
8. Jedrzejowska M, Borkowska J, Zimowski J, Kostera-Pruszczyk A, Milewski M, Jurek M et al. Unaffected patients with a homozygous absence of the SMN1 gene. *Eur J Hum Genet* 2008; 16(8): 930-934.
9. Kesari A, Idris MM, Chandak GR, Mittal B. Genotype-phenotype correlation of SMN locus genes in spinal muscular atrophy patients from India. *Exp Mol Med* 2005; 37(3): 147-154.
10. Koul R, Al Futaisi A, Chacko A, Rao V, Simsek M, Muralitharan S et al. Clinical and genetic study of spinal muscular atrophies in Oman. *J Child Neurol* 2007; 22(10): 1227-1230.

11. Kubo Y, Nishio H, Saito K. A new method for SMN1 and hybrid SMN gene analysis in spinal muscular atrophy using long-range PCR followed by sequencing. *J Hum Genet* 2015; 60(5): 233-239.
12. Liang YH, Chen XL, Yu ZS, Chen CY, Bi S, Mao LG et al. Deletion analysis of SMN1 and NAIP genes in Southern Chinese children with spinal muscular atrophy. *J Zhejiang Univ Sci B* 2009; 10(1): 29-34.
13. Marini M, Sasongko TH, Watihayati MS, Atif AB, Hayati F, Gunadi et al. Allele-specific PCR for a cost-effective & time-efficient diagnostic screening of spinal muscular atrophy. *Indian J Med Res* 2012; 135(1): 31-35.
14. Niba ETE, Ar Rochmah M, Harahap NIF, Awano H, Morioka I, Iijima K et al. SMA diagnosis: detection of SMN1 deletion with real-time mCOP-PCR system using fresh blood DNA. *Kobe J Med Sci* 2017; 63(3): E80-E83.
15. Ogino S, Leonard DG, Rennert H, Wilson RB. Spinal muscular atrophy genetic testing experience at an academic medical center. *J Mol Diagn* 2002; 4(1): 53-58.
16. Prior TW, Snyder PJ, Rink BD, Pearl DK, Pyatt RE, Mihal DC et al. Newborn and carrier screening for spinal muscular atrophy. *Am J Med Genet A* 2010; 152A(7): 1608-1616.
17. Seo J, Singh NN, Ottesen EW, Lee BM, Singh RN. A novel human-specific splice isoform alters the critical C-terminus of survival motor neuron protein. *Sci Rep* 2016; 6: 30778.
18. Shafeghati Y, Teymourian S, Babamohammadi G, Afrouzan F, Almadani N, Karimi-Nejad R et al. Molecular diagnosis Iranian patients with spinal muscular atrophy. *Arch Iran Med* 2004; 7(1): 47-52.
19. Sheng-Yuan Z, Xiong F, Chen YJ, Yan TZ, Zeng J, Li L et al. Molecular characterization of SMN copy number derived from carrier screening and from core families with SMA in a Chinese population. *Eur J Hum Genet* 2010; 18(9): 978-984.
20. Shin S, Park SS, Hwang YS, Lee KW, Chung SG, Lee YJ et al. Deletion of SMN and NAIP genes in Korean patients with spinal muscular atrophy. *J Korean Med Sci* 2000; 15(1): 93-98.
21. Taylor JL, Lee FK, Yazdanpanah GK, Staropoli JF, Liu M, Carulli JP et al. Newborn blood spot screening test using multiplexed real-time PCR to simultaneously screen for spinal muscular atrophy and severe combined immunodeficiency. *Clin Chem* 2015; 61(2): 412-419.
22. Vidal-Folch N, Gavrilov D, Raymond K, Rinaldo P, Tortorelli S, Matern D et al. Multiplex droplet digital PCR method applicable to newborn screening, carrier status, and assessment of spinal muscular atrophy. *Clin Chem* 2018; 64(12): 1753-1761.
23. Wadman RI, Stam M, Jansen MD, Van der Weegen Y, Wijngaarde CA, Harschnitz O et al. A comparative study of SMN protein and mRNA in blood and fibroblasts in patients with spinal muscular atrophy and healthy controls. *PLoS One* 2016; 11(11): e0167087.

24. Zuluaga-Sanchez S, Teynor M, Knight C, Thompson R, Lundqvist T, Ekelund M et al. Cost effectiveness of nusinersen in the treatment of patients with infantile-onset and later-onset spinal muscular atrophy in Sweden. *Pharmacoeconomics* 2019; 37(6): 04. 845–865.

Nicht INS3

1. Zaldivar T, Montejo Y, Acevedo AM, Guerra R, Vargas J, Garofalo N et al. Evidence of reduced frequency of spinal muscular atrophy type I in the Cuban population. *Neurology* 2005; 65(4): 636-638.

Nicht D4

1. Niba ETE, Rochmah MA, Harahap NIF, Awano H, Morioka I, Iijima K et al. Spinal muscular atrophy: advanced version of screening system with real-time mCOP-PCR and PCR-RFLP for SMN1 deletion. *Kobe J Med Sci* 2019; 65(2): E49-E53.

Nicht INS5

1. Grafe KA. Spinale Muskelatrophie: Test soll ins Neugeborenen-Screening. *Pharm Ztg* 2018; 163(34): 16.

2. Howell RR. Fifty years of newborn screening. *Mol Genet Metab* 2014; 113(1-2): 4-5.

3. Ke Q, Zhao ZY, Mendell JR, Baker M, Wiley V, Kwon JM et al. Progress in treatment and newborn screening for Duchenne muscular dystrophy and spinal muscular atrophy. *World J Pediatr* 2019; 15(3): 219-225.

4. Rudnik-Schöneborn S, Zerres K. Spinale Muskelatrophien. *Med Genet* 2009; 21(3): 349-357.

5. Rudnik-Schöneborn S, Zerres K, Hahnen E, Meng G, Voit T, Hanefeld F et al. Apparent autosomal recessive inheritance in families with proximal spinal muscular atrophy affecting individuals in two generations. *Am J Hum Genet* 1996; 59(5): 1163-1165.

6. Swoboda KJ. Seize the day: newborn screening for SMA. *Am J Med Genet A* 2010; 152A(7): 1605-1607.

7. Wadman M. Newborn screening urged for fatal neurological disorder. *Science* 2018; 360(6396): 1385.

Nicht D5

1. Czech C, Tang W, Bugawan T, Mano C, Horn C, Iglesias VA et al. Biomarker for spinal muscular atrophy: expression of SMN in peripheral blood of SMA patients and healthy controls. *PLoS One* 2015; 10(10): e0139950.

2. De Sanctis R, Pane M, Coratti G, Palermo C, Leone D, Pera MC et al. Clinical phenotypes and trajectories of disease progression in type 1 spinal muscular atrophy. *Neuromuscul Disord* 2018; 28(1): 24-28.

3. He J, Zhang QJ, Lin QF, Chen YF, Lin XZ, Lin MT et al. Molecular analysis of SMN1, SMN2, NAIP, GTF2H2, and H4F5 genes in 157 Chinese patients with spinal muscular atrophy. *Gene* 2013; 518(2): 325-329.
4. Kabra M, Arora S, Maria A, Aggarwal R. Preserved umbilical cord facilitates antenatal diagnosis of spinal muscular atrophy. *Indian Pediatr* 2003; 40(5): 415-418.
5. Kaneko K, Arakawa R, Urano M, Aoki R, Saito K. Relationships between long-term observations of motor milestones and genotype analysis results in childhood-onset Japanese spinal muscular atrophy patients. *Brain Dev* 2017; 39(9): 763-773.
6. Liu Z, Zhang P, He X, Liu S, Tang S, Zhang R et al. New multiplex real-time PCR approach to detect gene mutations for spinal muscular atrophy. *BMC Neurol* 2016; 16(1): 141.
7. Pyatt RE, Mihal DC, Prior TW. Assessment of liquid microbead arrays for the screening of newborns for spinal muscular atrophy. *Clin Chem* 2007; 53(11): 1879-1885.
8. Pyatt RE, Prior TW. A feasibility study for the newborn screening of spinal muscular atrophy. *Genet Med* 2006; 8(7): 428-437.
9. Watihayati MS, Zabidi-Hussin AMH, Tang TH, Matsuo M, Nishio H, Zilfalil BA. Corrigendum: deletion analyses of SMN1 and NAIP genes in Malaysian spinal muscular atrophy patients (*Pediatrics International* 2007; 49 (1): 11-14). *Pediatr Int* 2007; 49(2): 303.

Nicht INS6/D6

1. Vinas PCI, Martin HI, Zaldivar VT, Garofalo GN, Zayas GM, Guerra BR et al. Molecular evaluation in infantile spinal muscular atrophy. *Rev Chil Pediatr* 2013; 84(5): 499-504.

Nicht INS7/D7

1. Aartsma-Rus A. FDA approval of nusinersen for spinal muscular atrophy makes 2016 the year of splice modulating oligonucleotides. *Nucleic Acid Ther* 2017; 27(2): 67-69.
2. Boemer F, Caberg JH, Dideberg V, Dardenne D, Bours V, Hiligsmann M et al. Newborn screening for SMA in Southern Belgium. *Neuromuscul Disord* 2019; 29(5): 343-349.
3. Bürglen L, Spiegel R, Ignatius J, Cobben JM, Landrieu P, Lefebvre S et al. SMN gene deletion in variant of infantile spinal muscular atrophy. *Lancet* 1995; 346(8970): 316-317.
4. Clayton EW, Brothers KB. State-offered ethnically targeted reproductive genetic testing. *Genet Med* 2016; 18(2): 126-127.
5. Prasad V. Cost-effectiveness of nusinersen for spinal muscular atrophy: reply. *JAMA Pediatrics* 2018; 172(7): 701-702.
6. Prior TW, Snyder PJ, Rink BD, Pearl DK, Pyatt RE, Mihal DC et al. Newborn and carrier screening for spinal muscular atrophy. *Obstet Gynecol Surv* 2010; 65(11): 697-699.

A6.2.2 Vergleichende Interventionsstudien zum Therapiebeginn

Nicht INT1

1. Fehlings DL, Kirsch S, McComas A, Chipman M, Campbell K. Evaluation of therapeutic electrical stimulation to improve muscle strength and function in children with types II/III spinal muscular atrophy. *Dev Med Child Neurol* 2002; 44(11): 741-744.
2. Montes J, Garber CE, Kramer SS, Montgomery MJ, Dunaway S, Kamil-Rosenberg S et al. Single-blind, randomized, controlled clinical trial of exercise in ambulatory spinal muscular atrophy: why are the results negative? *J Neuromuscul Dis* 2015; 2(4): 463-470.
3. Shrader JA, Kats I, Kokkinis A, Zampieri C, Levy E, Joe GO et al. A randomized controlled trial of exercise in spinal and bulbar muscular atrophy. *Ann Clin Transl Neurol* 2015; 2(7): 739-747.

Nicht INT2

1. Bishop KM, Montes J, Finkel RS. Motor milestone assessment of infants with spinal muscular atrophy using the hammersmith infant neurological Exam-Part 2: experience from a nusinersen clinical study. *Muscle Nerve* 2018; 57(1): 142-146.
2. Darras BT, Chiriboga CA, Iannaccone ST, Swoboda KJ, Montes J, Mignon L et al. Nusinersen in later-onset spinal muscular atrophy: long-term results from the phase 1/2 studies. *Neurology* 2019; 92(21): e2492-e2506.
3. Porro F, Rinchetti P, Magri F, Riboldi G, Nizzardo M, Simone C et al. The wide spectrum of clinical phenotypes of spinal muscular atrophy with respiratory distress type 1: a systematic review. *J Neurol Sci* 2014; 346(1-2): 35-42.
4. Rustler V, Daggelmann J, Streckmann F, Bloch W, Baumann FT. Whole-body vibration in children with disabilities demonstrates therapeutic potentials for pediatric cancer populations: a systematic review. *Support Care Cancer* 2019; 27(2): 395-406.
5. Schoser B, Fong E, Geberhiwot T, Hughes D, Kissel JT, Madathil SC et al. Maximum inspiratory pressure as a clinically meaningful trial endpoint for neuromuscular diseases: a comprehensive review of the literature. *Orphanet J Rare Dis* 2017; 12(1): 52.

Nicht INT3

1. Mercuri E, Darras BT, Chiriboga CA, Day JW, Campbell C, Connolly AM et al. Nusinersen versus sham control in later-onset spinal muscular atrophy. *N Engl J Med* 2018; 378(7): 625-635.

Nicht INT5

1. Aartsma-Rus A. FDA approval of nusinersen for spinal muscular atrophy makes 2016 the year of splice modulating oligonucleotides. *Nucleic Acid Ther* 2017; 27(2): 67-69.
2. Aartsma-Rus A. Genetic therapies for spinal muscular atrophy type 1. *Lancet Neurol* 2018; 17(2): 111-112.

3. Ameri AA. Spinale Muskelatrophie (SMA): Antisense-Oligonukleotid Nusinersen bei SMA zugelassen. *Psychopharmakotherapie* 2017; 24(5): 239-240.
4. Aragon-Gawinska K, Seferian AM, Daron A, Gargaun E, Vuillerot C, Cances C et al. Nusinersen in patients older than 7 months with spinal muscular atrophy type 1: a cohort study. *Neurology* 2018; 91(14): e1312-e1318.
5. Burns JK, Kothary R, Parks RJ. Opening the window: the case for carrier and perinatal screening for spinal muscular atrophy. *Neuromuscul Disord* 2016; 26(9): 551-559.
6. Canadian Agency for Drugs and Technologies in Health. Pharmacoeconomic review report: nusinersen (Spinraza); CADTH common drug review [online]. 01.2018 [Zugriff: 27.05.2019]. URL: https://www.ncbi.nlm.nih.gov/books/NBK534118/pdf/Bookshelf_NBK534118.pdf.
7. Canadian Agency for Drugs and Technologies in Health. Pharmacoeconomic review report (resubmission): nusinersen (Spinraza); CADTH common drug review [online]. 04.2019 [Zugriff: 07.11.2019]. URL: <https://www.ncbi.nlm.nih.gov/books/n/srpe0576/pdf/>.
8. Chiriboga CA, Swoboda KJ, Darras BT, Iannaccone ST, Montes J, De Vivo DC et al. Results from a phase 1 study of nusinersen (ISIS-SMN(Rx)) in children with spinal muscular atrophy. *Neurology* 2016; 86(10): 890-897.
9. Claborn MK, Stevens DL, Walker CK, Gildon BL. Nusinersen: a treatment for spinal muscular atrophy. *Ann Pharmacother* 2019; 53(1): 61-69.
10. Corey DR. Synthetic nucleic acids and treatment of neurological diseases. *JAMA Neurology* 2016; 73(10): 1238-1242.
11. Corey DR. Nusinersen, an antisense oligonucleotide drug for spinal muscular atrophy. *Nat Neurosci* 2017; 20(4): 497-499.
12. Cunha MC, Oliveira AS, Labronici RH, Gabbai AA. Spinal muscular atrophy type II (intermediary) and III (Kugelberg-Welander): evolution of 50 patients with physiotherapy and hydrotherapy in a swimming pool. *Arq Neuropsiquiatr* 1996; 54(3): 402-406.
13. Darras BT, Farrar MA, Mercuri E, Finkel RS, Foster R, Hughes SG et al. An integrated safety analysis of infants and children with symptomatic spinal muscular atrophy (SMA) treated with nusinersen in seven clinical trials. *CNS Drugs* 2019; 33(9): 919-932.
14. Diener HC. Nusinersen im Vergleich zu Placebo bei Kindern mit Krankheitsbeginn nach dem 6. Lebensmonat. *Psychopharmakotherapie* 2018; 25(4): 215-216.
15. Farrar MA, Teoh HL, Carey KA, Cairns A, Forbes R, Herbert K et al. Nusinersen for SMA: expanded access programme. *J Neurol Neurosurg Psychiatry* 2018; 89(9): 937-942.
16. Finkel RS, Chiriboga CA, Vajsar J, Day JW, Montes J, De Vivo DC et al. Treatment of infantile-onset spinal muscular atrophy with nusinersen: a phase 2, open-label, dose-escalation study. *Lancet* 2016; 388(10063): 3017-3026.

17. Finsterer J, Soraru G. Onset manifestations of spinal and bulbar muscular atrophy (Kennedy's disease). *J Mol Neurosci* 2016; 58(3): 321-329.
18. Garg S. Management of scoliosis in patients with Duchenne muscular dystrophy and spinal muscular atrophy: a literature review. *J Pediatr Rehabil Med* 2016; 9(1): 23-29.
19. Geraci AP, Black K, Jin M, Rimler S, Evans A. Transforaminal lumbar puncture for intrathecal nusinersen administration. *Muscle Nerve* 2018; 58(1): E4-E5.
20. Gerrity MS, Prasad V, Obley AJ. Concerns about the approval of nusinersen sodium by the US food and drug administration. *JAMA Intern Med* 2018; 178(6): 743-744.
21. Hoy SM. Nusinersen: a review in 5q spinal muscular atrophy. *CNS Drugs* 2018; 32(7): 689-696.
22. Ioos C. There are other ways to manage spinal muscular atrophy type 1. *Chest* 2005; 127(4): 1463-1464.
23. La Foresta S, Faraone C, Sframeli M, Vita GL, Russo M, Profazio C et al. Intrathecal administration of nusinersen in type 1 SMA: successful psychological program in a single Italian center. *Neurol Sci* 2018; 39(11): 1961-1964.
24. Landfeldt E, Edström J, Sejersen T, Tulinius M, Lochmüller H, Kirschner J. Quality of life of patients with spinal muscular atrophy: a systematic review. *Eur J Paediatr Neurol* 2019; 23(3): 347-356.
25. Lewelt A, Krosschell KJ, Stoddard GJ, Weng C, Xue M, Marcus RL et al. Resistance strength training exercise in children with spinal muscular atrophy. *Muscle Nerve* 2015; 52(4): 559-567.
26. Montes J, Dunaway Young S, Mazzone ES, Pasternak A, Glanzman AM, Finkel RS et al. Nusinersen improves walking distance and reduces fatigue in later-onset spinal muscular atrophy. *Muscle Nerve* 2019; 60(4): 409-414.
27. Montes J, Garber CE, Kramer SS, Montgomery MJ, Dunaway S, Kamil-Rosenberg S et al. A randomized, controlled clinical trial of exercise in patients with spinal muscular atrophy: methods and baseline characteristics. *J Neuromuscul Dis* 2014; 1(2): 151-161.
28. Orrell RW. Motor neuron disease: systematic reviews of treatment for ALS and SMA. *Br Med Bull* 2010; 93(1): 145-159.
29. Pane M, Coratti G, Sansone VA, Messina S, Bruno C, Catteruccia M et al. Nusinersen in type 1 spinal muscular atrophy: twelve-month real-world data. *Ann Neurol* 2019; 86(3): 443-451.
30. Pane M, Palermo C, Messina S, Sansone VA, Bruno C, Catteruccia M et al. Nusinersen in type 1 SMA infants, children and young adults: preliminary results on motor function. *Neuromuscul Disord* 2018; 28(7): 582-585.
31. Pearlman S, Pourmand R. Evaluation of a patient presenting with progressive weakness and atrophy of the upper extremities. *J Clin Neuromuscul Dis* 2003; 5(1): 51-59.

32. Pechmann A, Langer T, Schorling D, Stein S, Vogt S, Schara U et al. Evaluation of children with SMA type 1 under treatment with nusinersen within the expanded access program in Germany. *J Neuromuscul Dis* 2018; 5(2): 135-143.
33. Preisler N, Andersen G, Thogersen F, Crone C, Jeppesen TD, Wibrand F et al. Effect of aerobic training in patients with spinal and bulbar muscular atrophy (Kennedy disease). *Neurology* 2009; 72(4): 317-323.
34. Ross LF, Kwon JM. Spinal muscular atrophy: past, present, and future. *Neoreviews* 2019; 20(8): e437-e451.
35. Sansone VA, Albamonte E, Salmin F, Casiraghi J, Pirola A, Bettinelli M et al. Intrathecal nusinersen treatment for SMA in a dedicated neuromuscular clinic: an example of multidisciplinary and integrated care. *Neurol Sci* 2019; 40(2): 327-332.
36. Sansone VA, Pane M, Messina S, Bruno C, D'Amico A, Albamonte E et al. A 5-center experience with intrathecal administration of nusinersen in SMA1 in Italy letter to the editor of european journal of pediatric neurology regarding the manuscript "single-center experience with intrathecal administration of nusinersen in children with spinal muscular atrophy type 1" written by Pechmann and colleagues". *Eur J Paediatr Neurol* 2018; 22(4): 729-731.
37. Walter MC, Wenninger S, Thiele S, Stauber J, Hiebeler M, Greckl E et al. Safety and treatment effects of nusinersen in longstanding adult 5q-SMA type 3: a prospective observational study. *J Neuromuscul Dis* 2019; 6(4): 453-465.
38. Yong J, Moffett M, Lucas S. Implementing a global expanded access program (EAP) for infantile-onset spinal muscular atrophy (type I): understanding the imperative, impact and challenges. *J Neuromuscul Dis* 2019; 6(2): 227-231.
39. Zerres K, Rudnik-Schöneborn S, Dubowitz V, Emery AEH, Forst R, Granata C et al. Guidelines for symptomatic therapy in spinal muscular atrophy (SMA). *Acta Cardiologica* 1995; 7(2): 61-66.

Nicht INT6

1. Vasconcellos Jorge MGA, Carrapatoso BC, Fernandes ABS. Physical therapy in spinal muscular atrophy type I: a literature review. *Revista Neurociencias* 2013; 21(3): 402-407.

Nicht INT7

1. Nusinersen (Spinraza) for spinal muscular atrophy. *Med Lett Drugs Ther* 2017; 59(1517): 50-52.
2. Nusinersen use in spinal muscular atrophy. *Pediatrics* 2019; 143(1): e20183335.
3. Prasad V. Cost-effectiveness of Nusinersen for spinal muscular atrophy: reply. *JAMA Pediatrics* 2018; 172(7): 701-702.
4. Richardson RC. Cost-effectiveness of Nusinersen for spinal muscular atrophy. *JAMA Pediatrics* 2018; 172(7): 701.

A6.3 Liste der ausgeschlossenen Dokumente aus den durch den G-BA übermittelten Dokumenten mit Ausschlussgründen

Nicht INS5/INT5/D5

1. Deutsche Gesellschaft für Neugeborenencreening. Nationaler Screeningreport Deutschland 2016 [online]. 07.2018 [Zugriff: 05.02.2019]. URL: http://www.screening-dgns.de/Pdf/Screeningreports/DGNS-Screeningreport-d_2016.pdf.
2. Evidence-based Review Group. Evidence-based review of newborn screening for spinal muscular atrophy (SMA): final report (v5.2) [online]. 13.03.2018 [Zugriff: 05.02.2019]. URL: <https://www.hrsa.gov/sites/default/files/hrsa/advisory-committees/heritable-disorders/reports-recommendations/sma-final-report.pdf>.
3. Institute for Clinical and Economic Review. Spinraza and Zolgensma for spinal muscular atrophy: effectiveness and value: draft evidence report [online]. 20.12.2018 [Zugriff: 05.02.2019]. URL: https://icer-review.org/wp-content/uploads/2018/07/ICER_SMA_Draft_Evidence_Report_122018-1.pdf.
4. Prior TW, Finanger E. Spinal muscular atrophy [online]. In: GeneReviews. 22.12.2016 [Zugriff: 07.02.2019]. URL: <https://www.ncbi.nlm.nih.gov/books/NBK1352/>.
5. Rudnik-Schöneborn S, Zerres K. Spinale Muskelatrophien. Med Genet 2017; 29(1): 21-34.

Nicht INS7/INT7/D7

1. Bertini E, Hwu WL, Reyna SP, Farwell W, Gheuens S, Sun P et al. Efficacy and safety of nusinersen in infants with presymptomatic spinal muscular atrophy (SMA): interim results from the NURTURE study. Eur J Paediatr Neurol 2017; 21(Suppl 1): e14.
2. Mercuri E, Finkel R, Kirschner J, Chiriboga C, Kuntz N, Sun P et al. Efficacy and safety of nusinersen in children with later-onset spinal muscular atrophy (SMA): end of study results from the phase 3 CHERISH study. Neuromuscul Disord 2017; 27(Suppl 2): S210.

A7 Suchstrategien

A7.1 Suchstrategien in bibliografischen Datenbanken

A7.1.1 Suchstrategien für vergleichende Interventionsstudien der Screeningkette und Studien zur diagnostischen Güte

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) ALL 1946 to October 21, 2019

#	Searches
1	exp muscular atrophy, spinal/
2	(spinal* adj1 muscular* adj1 atroph*).ti,ab.
3	or/1-2
4	exp infant/
5	(newborn* or neonat* or pediatric* or infant*).ti,ab.
6	or/4-5
7	neonatal screening/
8	and/3,7
9	*mass screening/
10	screen*.ti,ab.
11	or/9-10
12	and/3,6,11
13	exp smn complex proteins/
14	exp exons/
15	(smn1* or smn2* or (smn* adj1 ("1" or "2"))).ti,ab.
16	(survival* adj3 motor* adj1 neuron*).ti,ab.
17	(exon7* or (exon* adj1 "7")).ti,ab.
18	or/13-17
19	and/3,6,18
20	or/8,12,19
21	20 not (exp animals/ not humans.sh.)
22	21 not (comment or editorial).pt.
23	22 and (english or german).lg.

2. Embase

Suchoberfläche: Ovid

- Embase 1974 to 2019 October 21

#	Searches
1	exp spinal muscular atrophy/
2	(spinal* adj1 muscular* adj1 atroph*).ti,ab.
3	or/1-2
4	exp infant/
5	(newborn* or neonat* or pediatric* or infant*).ti,ab.
6	or/4-5
7	newborn screening/
8	and/3,7
9	screen*.ti,ab.
10	and/3,6,9
11	survival motor neuron protein/
12	survival motor neuron protein 1/
13	survival motor neuron protein 2/
14	smn1 gene/
15	smn2 gene/
16	protein smn1/
17	protein smn2/
18	exon/
19	(smn1* or smn2* or (smn* adj1 ("1" or "2"))).ti,ab.
20	(survival* adj3 motor* adj1 neuron*).ti,ab.
21	(exon7* or (exon* adj1 "7")).ti,ab.
22	or/11-21
23	and/3,6,22
24	or/8,10,23
25	24 not medline.cr.
26	25 not (exp animal/ not exp human/)
27	26 not (conference abstract or conference review or editorial).pt.
28	27 and (english or german).lg.

3. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Database of Systematic Reviews: Issue 10 of 12, October 2019
- Cochrane Central Register of Controlled Trials: Issue 10 of 12, October 2019

ID	Search
#1	[mh "muscular atrophy, spinal"]
#2	(spinal* near/1 muscular* near/1 atroph*):ti,ab
#3	#1 or #2
#4	[mh "infant"]
#5	(newborn* or neonat* or pediatric* or infant*):ti,ab
#6	#4 or #5
#7	[mh ^"neonatal screening"]
#8	#3 and #7
#9	[mh ^"mass screening"]
#10	screen*:ti,ab
#11	#9 or #10
#12	#3 and #6 and #11
#13	[mh "smn complex proteins"]
#14	[mh "exons"]
#15	(smn1* or smn2* or (smn* near/1 (1 or 2))):ti,ab
#16	(survival* near/3 motor* near/1 neuron*):ti,ab
#17	(exon7* or (exon* near/1 7)):ti,ab
#18	#13 or #14 or #15 or #16 or #17
#19	#3 and #6 and #18
#20	#8 or #12 or #19 in Cochrane Reviews, Cochrane Protocols
#21	#8 or #12 or #19 in Trials

4. Health Technology Assessment Database

Suchoberfläche: Centre for Reviews and Dissemination

Line	Search
1	MeSH DESCRIPTOR Muscular Atrophy, Spinal EXPLODE ALL TREES
2	(spinal* AND muscular* AND atroph*)
3	#1 OR #2
4	MeSH DESCRIPTOR Infant EXPLODE ALL TREES
5	(newborn* OR neonat* OR pediatric* OR infant*)
6	#4 OR #5
7	MeSH DESCRIPTOR Neonatal Screening
8	#3 AND #7
9	MeSH DESCRIPTOR Mass Screening
10	(screen*)

Line	Search
11	#9 OR #10
12	#3 AND #6 AND #11
13	MeSH DESCRIPTOR SMN Complex Proteins EXPLODE ALL TREES
14	MeSH DESCRIPTOR Exons EXPLODE ALL TREES
15	(smn1* OR smn2* OR (smn* NEXT (1 OR 2)))
16	(survival* AND motor* AND neuron*)
17	(exon7* OR (exon* NEXT 7))
18	#13 OR #14 OR #15 OR #16 OR #17
19	#3 AND #6 AND #18
20	#8 OR #12 OR #19
21	(#20) IN HTA

A7.1.2 Suchstrategien für vergleichende Interventionsstudien zum Therapiebeginn

1. MEDLINE

Suchoberfläche: Ovid

- Ovid MEDLINE(R) 1946 to October Week 2 2019
- Ovid MEDLINE(R) Daily Update October 21, 2019

#	Searches
1	exp muscular atrophy, spinal/
2	(spinal* adj1 muscular* adj1 atroph*).ti,ab.
3	or/1-2
4	exp physical therapy modalities/
5	rehabilitation.fs.
6	(physiotherap* or exercise* or gymnastic* or massage* or mobilization* or stretching*).ti,ab.
7	(therap* adj3 (manual* or physical*)).ti,ab.
8	((muscle* or relaxation* or strength* or endurance* or fitness* or physical* or restistance*) adj1 training*).ti,ab.
9	nusinersen*.mp.
10	or/4-9
11	and/3,10
12	11 not (exp animals/ not humans.sh.)
13	12 not (comment or editorial).pt.
14	13 and (english or german).lg.

Suchoberfläche: Ovid

- Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations 1946 to October 21, 2019
- Ovid MEDLINE(R) Epub Ahead of Print October 21, 2019

#	Searches
1	(spinal* adj1 muscular* adj1 atroph*).ti,ab.
2	(physiotherap* or exercise* or gymnastic* or massage* or mobilization* or stretching*).ti,ab.
3	(therap* and (manual* or physical*)).ti,ab.
4	((muscle* or relaxation* or strength* or endurance* or fitness* or physical* or restistance*) and training*).ti,ab.
5	nusinersen*.mp.
6	or/2-5
7	and/1,6
8	7 not (comment or editorial).pt.
9	8 and (english or german).lg.

2. Embase

Suchoberfläche: Ovid

- Embase 1974 to 2019 October 21

#	Searches
1	spinal muscular atrophy/
2	exp hereditary spinal muscular atrophy/
3	(spinal* adj1 muscular* adj1 atroph*).ti,ab.
4	or/1-3
5	physiotherapy/
6	exp kinesiotherapy/
7	massage/
8	(physiotherap* or exercise* or gymnastic* or massage* or mobilization* or stretching*).ti,ab.
9	(therap* adj3 (manual* or physical*)).ti,ab.
10	((muscle* or relaxation* or strength* or endurance* or fitness* or physical* or restistance*) adj1 training*).ti,ab.
11	nusinersen*.mp.
12	or/5-11
13	and/4,12

#	Searches
14	13 not medline.cr.
15	14 not (exp animal/ not exp human/)
16	15 not (conference abstract or conference review or editorial).pt.
17	16 and (english or german).lg.

3. The Cochrane Library

Suchoberfläche: Wiley

- Cochrane Database of Systematic Reviews: Issue 10 of 12, October 2019
- Cochrane Central Register of Controlled Trials: Issue 10 of 12, October 2019

ID	Search
#1	[mh "muscular atrophy, spinal"]
#2	(spinal* near/1 muscular* near/1 atroph*):ti,ab
#3	#1 or #2
#4	[mh "physical therapy modalities"]
#5	[mh /rh]
#6	(physiotherap* or exercise* or gymnastic* or massage* or mobilization* or stretching*):ti,ab
#7	(therap* near/3 (manual* or physical*)):ti,ab
#8	((muscle* or relaxation* or strength* or endurance* or fitness* or physical* or restistance*) near/1 training*):ti,ab
#9	nusinersen*
#10	#4 or #5 or #6 or #7 or #8 or #9
#11	#3 and #10 in Cochrane Reviews, Cochrane Protocols
#12	#3 and #10 in Trials

4. Health Technology Assessment Database

Suchoberfläche: Centre for Reviews and Dissemination

Line	Search
1	MeSH DESCRIPTOR Muscular Atrophy, Spinal EXPLODE ALL TREES
2	(spinal* AND muscular* AND atroph*)
3	#1 OR #2
4	MeSH DESCRIPTOR Physical Therapy Modalities EXPLODE ALL TREES
5	(physiotherap* OR exercise* OR gymnastic* OR massage* OR mobilization* OR stretching*)

Line	Search
6	(therap* AND (manual* OR physical*))
7	((muscle* OR relaxation* OR strength* OR endurance* OR fitness* OR physical* OR restistance*) AND training*)
8	(nusinersen*)
9	#4 OR #5 OR #6 OR #7 OR #8
10	#3 AND #9
11	(#10) IN HTA

A7.2 Suche in Studienregistern

1. ClinicalTrials.gov

Anbieter: U.S. National Institutes of Health

- URL: <http://www.clinicaltrials.gov>
- Eingabeoberfläche: Advanced Search

Suchstrategie
spinal muscular atrophy

2. EU Clinical Trials Register

Anbieter: European Medicines Agency

- URL: <https://www.clinicaltrialsregister.eu/ctr-search/search>
- Eingabeoberfläche: Basic Search

Suchstrategie
nusinersen OR isis396443 OR isis-396443 OR (isis 396443)

3. International Clinical Trials Registry Platform Search Portal

Anbieter: World Health Organization

- URL: <http://apps.who.int/trialsearch/>
- Eingabeoberfläche: Standard Search

Suchstrategie
spinal muscular atrophy

4. Biogen Clinical Trial Results

Anbieter: Biogen

- URL: <http://clinicalresearch.biogen.com/Study/ByProduct/18>

Suchstrategie
Search by Product / SPINRAZA (nusinersen)

A8 Offenlegung potenzieller Interessenkonflikte der externen Sachverständigen

Im Folgenden sind die potenziellen Interessenkonflikte der externen Sachverständigen zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangaben der einzelnen Personen anhand des „Formblatts zur Offenlegung potenzieller Interessenkonflikte“ mit Stand 11/2016. Das aktuelle Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

Externe Sachverständige

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7
Mathes, Tim	nein	nein	ja	ja	nein	nein	nein
Prediger, Barbara	ja	nein	nein	ja	nein	nein	nein
Schoser, Benedikt	ja	ja	ja	ja	nein	nein	nein

Im „Formblatt zur Offenlegung potenzieller Interessenkonflikte“ (Version 11/2016) wurden folgende 7 Fragen gestellt:

Frage 1: Sind oder waren Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor bei einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere bei einem pharmazeutischen Unternehmen, Hersteller von Medizinprodukten oder einem industriellen Interessenverband angestellt, für diese selbstständig oder ehrenamtlich tätig bzw. sind oder waren Sie freiberuflich in eigener Praxis tätig? (Zu den oben genannten Einrichtungen zählen beispielsweise auch Kliniken, Einrichtungen der Selbstverwaltung, Fachgesellschaften, Auftragsinstitute)

Frage 2: Beraten Sie oder haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor ein Unternehmen, eine Institution oder einen Interessenverband im Gesundheitswesen, insbesondere ein pharmazeutisches Unternehmen, einen Hersteller von Medizinprodukten oder einen industriellen Interessenverband direkt oder indirekt beraten (z. B. als Gutachter, Sachverständiger, Mitglied eines Advisory Boards, Mitglied eines Data Safety Monitoring Boards (DSMB) oder Steering Committees)?

Frage 3: Haben Sie innerhalb des laufenden Jahres und der 3 Kalenderjahre davor direkt oder indirekt von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband Honorare erhalten (z. B. für Vorträge, Schulungstätigkeiten, Stellungnahmen oder Artikel)?

Frage 4: Haben Sie oder haben die von Ihnen unter Frage 1 genannten Einrichtungen innerhalb des laufenden Jahres und der 3 Kalenderjahre davor von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband finanzielle Unterstützung z. B. für Forschungsaktivitäten, die Durchführung klinischer Studien, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten? (Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.)

Frage 5: Haben Sie oder haben die von Ihnen unter Frage 1 genannten Einrichtungen innerhalb des laufenden Jahres und der 3 Kalenderjahre davor sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Unterstützung bei der Ausrichtung einer Veranstaltung, Übernahme von Reisekosten oder Teilnahmegebühren für Fortbildungen / Kongresse) erhalten von einem Unternehmen, einer Institution oder einem Interessenverband im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen, einem Hersteller von Medizinprodukten oder einem industriellen Interessenverband? (Sofern Sie in einer ausgedehnten Institution tätig sind, genügen Angaben zu Ihrer Arbeitseinheit, zum Beispiel Klinikabteilung, Forschungsgruppe etc.)

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile eines Unternehmens oder einer anderweitigen Institution im Gesundheitswesen, insbesondere von einem pharmazeutischen Unternehmen oder einem Hersteller von Medizinprodukten? Besitzen Sie Anteile eines „Branchenfonds“, der auf pharmazeutische Unternehmen oder Hersteller von Medizinprodukten ausgerichtet ist? Besitzen Sie Patente für ein pharmazeutisches Erzeugnis oder ein Medizinprodukt oder eine medizinische Methode oder Gebrauchsmuster für ein pharmazeutisches Erzeugnis oder ein Medizinprodukt?

Frage 7: Sind oder waren Sie jemals an der Erstellung einer Leitlinie oder Studie beteiligt, die eine mit diesem Projekt vergleichbare Thematik behandelt/e? Gibt es sonstige Umstände, die aus Sicht eines unvoreingenommenen Betrachters als Interessenkonflikt bewertet werden können (z. B. Aktivitäten in gesundheitsbezogenen Interessengruppierungen bzw. Selbsthilfegruppen, politische, akademische, wissenschaftliche oder persönliche Interessen)?