

Anwendungsbegleitende Datenerhebung zu Risdiplom: Prüfung des Studienprotokolls (Version 1.0) und des statistischen Analyseplans (Version 1.0)

2. Addendum zum Projekt A21-131

A horizontal bar composed of 18 rectangular segments of varying shades of blue and grey. The word 'ADDENDUM' is written in white capital letters on a dark blue segment that spans across the middle of the bar.

ADDENDUM

Projekt: A23-87

Version: 1.1

Stand: 27.10.2023

IQWiG-Berichte – Nr. 1647

DOI: 10.60584/A23-87_V1.1

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Anwendungsbegleitende Datenerhebung zu Risdiplom: Prüfung des Studienprotokolls (Version 1.0) und des statistischen Analyseplans (Version 1.0) – 2. Addendum zum Projekt A21-131

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

16.08.2023

Interne Projektnummer

A23-87

DOI-URL

https://dx.doi.org/10.60584/A23-87_V1.1

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

Zitiervorschlag

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Anwendungsbegleitende Datenerhebung zu Risdiplom: Prüfung des Studienprotokolls (Version 1.0) und des statistischen Analyseplans (Version 1.0); 2. Addendum zum Projekt A21-131 [online]. 2023 [Zugriff: TT.MM.JJJJ]. URL: https://dx.doi.org/10.60584/A23-87_V1.1.

Schlagwörter

Risdiplom, Muskelatrophie – Spinale, Nutzenbewertung, Register, Forschungsdesign, Peer Review – Forschung

Keywords

Risdiplom, Muscular Atrophy – Spinal, Benefit Assessment, Registries, Research Design, Peer Review – Research

An dem Addendum beteiligte Mitarbeiterinnen und Mitarbeiter des IQWiG

- Anna-Katharina Barnert
- Lars Beckmann
- Tatjana Hermanns
- Katharina Hirsch
- Simone Johner
- Volker Vervölgyi

Inhaltsverzeichnis

	Seite
Tabellenverzeichnis	vi
Abkürzungsverzeichnis.....	vii
1 Hintergrund.....	1
1.1 Änderung im Vergleich zur Version 1.0.....	1
2 Prüfung der Unterlagen zur Planung der AbD von Risdiplom.....	2
2.1 Allgemeine Anmerkungen zu den vom pU vorgelegten Unterlagen	2
2.2 Spezifische Anmerkungen zum SP und SAP	2
2.2.1 Fragestellung gemäß PICO	2
2.2.1.1 Population.....	3
2.2.1.2 Intervention und Kontrolle	4
2.2.1.3 Endpunkte (Outcomes).....	4
2.2.2 Datenquelle	14
2.2.3 Indexdatum, Zuordnung zu den Behandlungsgruppen und Beobachtungsdauer	15
2.2.4 Studienpopulation der Registerstudie	16
2.2.4.1 Fallzahlplanung	16
2.2.5 Identifikation und Erhebung von Confoundern, Adjustierung für Confounder... 17	17
2.2.5.1 Identifikation und Erhebung wichtiger Confounder.....	18
2.2.5.2 Berücksichtigung wichtiger Confounder bei der Analyse.....	18
2.2.6 Anmerkungen zu den geplanten Auswertungen für die Endpunkte der AbD	20
2.2.7 Untersuchung auf potenzielle Effektmodifikatoren und Subgruppenanalysen ..	21
2.2.8 Qualität und Vollständigkeit der Datenerhebung.....	21
2.2.8.1 Maßnahmen zur Sicherstellung der Qualität und Vollständigkeit der Datenerhebung.....	21
2.2.8.2 Umgang mit fehlenden Daten	22
2.2.9 Weitere Anpassungen der AbD	23
2.2.10 Interpretation der Ergebnisse der AbD	24
2.2.11 Zeitplan für die AbD.....	24
2.3 Zusammenfassung und Fazit	25

3	Literatur	26
Anhang A	Prüfung der Informationsbeschaffung zur Identifizierung von Confoundern ..	30
Anhang B	Gegenüberstellung der Operationalisierungen zwischen der AbD zu Risdiplom und Onasemnogen-Abeparvovec	32

Tabellenverzeichnis

	Seite
Tabelle 1: Geplante Instrumente und Operationalisierungen zur Erhebung der motorischen Funktion.....	6
Tabelle 2: Quellen der Informationsbeschaffung zur Identifizierung von Confoundern.....	30
Tabelle 3: Unterschiede in den geplanten Endpunkten und Operationalisierungen zu Risdiplom im Vergleich zu Onasemnogen-Abeparvovec.....	32

Abkürzungsverzeichnis

Abkürzung	Bedeutung
6MWT	6 Minuten Gehstest
AbD	Anwendungsbegleitende Datenerhebung
CHOP-INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
CTCAE	Common Terminology Criteria for Adverse Events
FSW	Fine Stratification Weights
G-BA	Gemeinsamer Bundesausschuss
HFMSE	Hammersmith Functional Motor Scale Expanded
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IPTW	Inverse Probability of Treatment Weighting
ITT	Intention to treat
MMRM	Modell für Messwiederholungen (Mixed Model Repeated measures)
PICO	Patient, Intervention, Comparator, Outcome
PS	Propensity Score
pU	pharmazeutischer Unternehmer
RULM	Revised Upper Limb Module
SAP	statistischer Analyseplan
SGB	Sozialgesetzbuch
SMA	Spinale Muskelatrophie
SMD	Standardisierte Mittelwertsdifferenz
SMN	Survival Motor Neuron
SP	Studienprotokoll
SUE	schwerwiegendes unerwünschtes Ereignis
UE	unerwünschtes Ereignis

1 Hintergrund

Der Gemeinsame Bundesausschuss (G-BA) hat das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) am 16.08.2023 mit der Prüfung des Studienprotokolls (SP) und des statistischen Analyseplans (SAP) zur anwendungsbegleitenden Datenerhebung (AbD) zu Risdiplam beauftragt.

In seiner Sitzung am 21.07.2022 hat der G-BA beschlossen, eine AbD und Auswertungen nach § 35a Absatz 3b Satz 1 SGB V für den Wirkstoff Risdiplam in der Indikation „spinale Muskelatrophie“ zu fordern [1,2]. Dem Beschluss liegt unter anderem das vom IQWiG erstellte Konzept für eine AbD zu Risdiplam zugrunde (Rapid Report A21-131 vom 15.02.2022) [3].

Zur Prüfung, ob die Anforderungen des G-BA an die AbD und an Auswertungen durch die vom pharmazeutischen Unternehmer (pU) erstellten Unterlagen zum SP und SAP umgesetzt worden sind, hat der G-BA dem IQWiG diese Unterlagen übermittelt [4,5] und mit der Prüfung dieser Unterlagen beauftragt. Neben dem G-BA-Beschluss zu Risdiplam sollen die Inhalte der diesbezüglichen Beratung des pU zur Studienplanung der AbD berücksichtigt werden (Beratungsanforderung 2022-B-215) [6].

1.1 Änderung im Vergleich zur Version 1.0

Die vorliegende Version 1.1 vom 27.10.2023 ersetzt die Version 1.0 des 2. Addendums zum Projekt A21-131 vom 29.09.2023. Neben einer redaktionellen Anpassung wurde folgende Änderung in der Version 1.1 im Vergleich zu Version 1.0 vorgenommen:

- Im Abschnitt 2.2.1.3 wurde die fehlerhafte Beschreibung der Operationalisierung der dauerhaften Beatmung in der AbD zu Onasemnogen-Abeparvovec korrigiert.

Die Änderung hat keine Auswirkung auf das Fazit des Addendums.

2 Prüfung der Unterlagen zur Planung der AbD von Risdiplom

2.1 Allgemeine Anmerkungen zu den vom pU vorgelegten Unterlagen

Der pU hat, wie im G-BA-Beschluss zur AbD gefordert [1,2], ein SP und einen SAP zur Prüfung vorgelegt, beide Dokumente in der Version 1.0 [4,5].

Gemäß Beschluss des G-BA ist die AbD als nicht-randomisierter Vergleich von Risdiplom und parallel erhobener sowie zeitlich nicht parallel erhobener Daten zu Nusinersen und Onasemnogen-Abeparvovec innerhalb einer Datenquelle, sofern die zeitlich nicht parallel erhobenen Daten ebenfalls die im G-BA Beschluss genannten Anforderungen an die Datenqualität entsprechen, umzusetzen [1]. Im Falle einer vergleichenden Registerstudie soll das SMARtCARE-Register als primäre Datenquelle genutzt werden, sofern die im G-BA-Beschluss aufgeführten Qualitätskriterien erfüllt sind. Der pU wählt das SMARtCARE-Register als Datenquelle für die AbD.

In den nachfolgenden Abschnitten werden die vom pU vorgelegten Dokumente dahin gehend geprüft, ob durch sie die Durchführung der AbD unter Verwendung des SMARtCARE-Registers einschließlich der zugehörigen Auswertungen ausreichend genau beschrieben wird, ob die Planung der AbD einschließlich der Auswertung der Daten inhaltlich und methodisch sachgerecht ist, und ob die im G-BA-Beschluss genannten Anforderungen adäquat umgesetzt wurden.

2.2 Spezifische Anmerkungen zum SP und SAP

Zur Prüfung des SP / SAP lagen zum SMARtCARE-Register zwar das Registerprotokoll aber kein Kodierhandbuch oder eine Übersicht über den aktuellen Datensatz vor. Für die Endpunkte, einige Einschlusskriterien und die Confounder macht der pU im SP und SAP tabellarische Angaben zu den jeweiligen Dokumentationsfeldern im Register. Ohne das Kodierhandbuch geht aus diesen Angaben jedoch nicht immer hervor, was genau dokumentiert wird. Eine Einschätzung der jeweiligen angegebenen Dokumentationsfelder im SMARtCARE-Register findet sich in den entsprechenden Prüfabschnitten (z. B. zu den Endpunkten im Abschnitt 2.2.1.3). Für die adäquate Bewertung der Dokumentation der relevanten Daten soll der pU alle relevanten Datenfelder mit der entsprechenden Operationalisierung im Register im SP und SAP angeben und die aktuelle Version des Kodierhandbuchs zur ergänzenden Information vorlegen.

2.2.1 Fragestellung gemäß PICO

Die Fragestellung des G-BA für die AbD und die anschließende Auswertung der Daten sind im Beschluss des G-BA mithilfe des PICO-Schemas niedergelegt (PICO: Population, Intervention, Kontrolle, Outcome). Die folgenden Abschnitte beurteilen die Umsetzung des PICO im SP und SAP.

2.2.1.1 Population

Gemäß G-BA-Beschluss soll die AbD für folgende Populationen durchgeführt werden:

- präsymptomatische Patientinnen und Patienten mit einer 5q-assoziierten spinalen Muskelatrophie (SMA) und bis zu 3 Kopien des Survival Motor Neuron (SMN)2-Gens
- symptomatische Patientinnen und Patienten mit einer klinisch diagnostizierten Typ-1-SMA
- symptomatische Patientinnen und Patienten mit einer klinisch diagnostizierten Typ-2-SMA und bis zu 3 Kopien des SMN2-Gens
- symptomatische Patientinnen und Patienten mit einer klinisch diagnostizierten Typ-3-SMA und bis zu 3 Kopien des SMN2-Gens

Die Ein- und Ausschlusskriterien der geplanten AbD stimmen gemäß SP prinzipiell mit den Vorgaben des G-BA überein.

Laut G-BA Beschluss soll die AbD unter Berücksichtigung von zeitlich parallel erhobenen sowie zeitlich nicht parallel erhobenen Daten durchgeführt werden. Der pU berücksichtigt dies bei den Einschlusskriterien dahin gehend, dass Patientinnen und Patienten berücksichtigt werden, die ab dem Zeitpunkt der entsprechenden Zulassung der verschiedenen SMA-Therapien behandelt wurden. Dies sind für Nusinersen Patientinnen und Patienten mit einem Behandlungsstart ab dem 30.05.2017, für Onasemnogen-Abeparvovec ab dem 18.05.2020 und für Risdiplam ab dem 26.03.2021. Weitere Einschränkungen, insbesondere zu den erforderlichen Kriterien zur Berücksichtigung nicht parallel erhobener Daten (entsprechend der im G-BA Beschluss [1] genannten Anforderungen an die Datenqualität), macht der pU nicht. Dies muss ergänzt werden.

Darüber hinaus benennt der pU als Einschlusskriterium, dass die Visite, die als Baselinevisite definiert wird, zwischen 6 Wochen vor und 3 Wochen nach der ersten Applikation der SMA-Medikation liegen muss. Da dieses Vorgehen weder inhaltlich sachgerecht ist (siehe Abschnitt 2.2.3), noch sinnvoll als Einschlusskriterium erscheint, sollte dieses Einschlusskriterium gestrichen werden.

Der pU definiert als Ausschlusskriterium eine Vorbehandlung der Patientinnen und Patienten mit Risdiplam, Nusinersen oder Onasemnogen-Abeparvovec vor Einschluss in das Register. Dies ist grundsätzlich nachvollziehbar, jedoch sollte sichergestellt sein, dass über dieses Ausschlusskriterium nicht auch solche Patientinnen und Patienten per se ausgeschlossen werden, die eine Vorbehandlung lediglich in Form einer Brückentherapie erhalten haben. Deshalb sollte im Sinne der Erreichung einer hinreichend hohen Fallzahl für solche Patientinnen und Patienten zunächst geprüft werden, ob entsprechende Baseline-Daten zum

Zeitpunkt des Therapiestarts in ausreichender Qualität vorliegen oder nachträglich erhoben werden können.

Der pU gibt an anderer Stelle im SP (Abschnitt 8.4.1) an, über welche Seite der Dokumentationsfelder des SMARTCARE-Registers die Dokumentation der Haupt-Einschlusskriterien erfolgt. Die Angaben sind nicht vollständig für alle Ein- und Ausschlusskriterien und enthalten keine Angaben zu den spezifischen Datenfeldern. Eine genaue Beurteilung zur Abgrenzung der Patientenpopulationen über die entsprechenden Datenfelder lässt sich auf Basis dieser Informationen nicht vornehmen. So gibt der pU bspw. an, das Merkmal Alter bei Symptombeginn über die Seite „Baseline – Clinical diagnosis“ zu erfassen. Wie spezifisch die Erfassung über dieses Datenfeld ist, kann aus der Angabe nicht abgeleitet werden. Wie oben ausgeführt müssen die Datenfelder, über die eine Abgrenzung der relevanten Patientenpopulation erfolgt, im SP und SAP konkretisiert werden.

2.2.1.2 Intervention und Kontrolle

Der pU schließt Patientinnen und Patienten, die mit Risdiplam oder Onasemnogen-Abeparvovec oder Nusinersen gemäß Zulassung behandelt wurden, in die Datenerhebung ein. Dieses Vorgehen ist adäquat.

2.2.1.3 Endpunkte (Outcomes)

Allgemeine Anmerkungen

Der pU plant eine Vielzahl von Endpunkten (insbesondere zur Abbildung der motorischen Funktion) zu erheben, die die vom G-BA geforderten Endpunkt-Kategorien umfänglich abdecken. Die eingesetzten Instrumente und Operationalisierungen weichen jedoch in vielen Punkten von den Erhebungen der zeitlich parallel und ebenfalls im SMARTCARE-Register laufenden AbD zu Onasemnogen-Abeparvovec [7] ab (siehe auch ergänzende Gegenüberstellung in Anhang B). Eine Begründung für dieses Vorgehen nennt der pU nicht. Um die Vergleichbarkeit der Ergebnisse auch zwischen den AbDs zu gewährleisten, ist eine einheitliche Vorgehensweise zu bevorzugen. Deshalb sollten möglichst für identische Endpunkte (beispielsweise beim kombinierten Endpunkt Tod / dauerhafte Beatmung) einheitliche und im Verfahren zu Onasemnogen-Abeparvovec als sachgerecht bewertete Operationalisierungen gewählt werden. Ein dem Verfahren zu Onasemnogen-Abeparvovec entsprechendes Vorgehen, das er auch entsprechend begründet, wählt der pU auch zur Auswahl der relevanten Confounder (siehe Abschnitt 2.2.5.1).

Problematisch für eine Bewertung des Zusatznutzens ist die Vielzahl der Endpunkte zur Beschreibung der motorischen Funktion. Diese Multiplizität sollte verringert werden, indem die relevanten Endpunkte selektiert und die Endpunkte insgesamt hierarchisiert werden. Zu bevorzugen ist dabei ein zur AbD zu Onasemnogen-Abeparvovec einheitliches Vorgehen. Die Entscheidungen müssen im SP prä-spezifiziert werden.

Für eine Reihe von Endpunkten plant der pU Auswertungen zu mehreren Zeitpunkten (z. B. 12, 24 und 36 Monate nach Behandlungsstart). In diesen Situationen sollte immer die Auswertung, die den längsten Beobachtungszeitraum berücksichtigt, die primäre Analyse darstellen.

Darüber hinaus ist weder im SP noch im SAP eine Übersicht zu den geplanten Erhebungen enthalten. Es ist somit nicht nachvollziehbar, wie häufig die einzelnen Erhebungen durchgeführt werden sollen und ob insbesondere die Erhebung aller motorischen Endpunkte bei jeder Visite erfolgt. Aus den Angaben in den Studienunterlagen geht lediglich hervor, dass der pU davon ausgeht, dass die Visiten alle 4 Monate stattfinden. Um, vor allem vor dem Hintergrund der großen Anzahl an zu erhebenden Endpunkten, eine Übersicht zu den geplanten Erhebungszeitpunkten im Beobachtungsverlauf zu haben, sollte ein Erhebungsplan ergänzt werden.

Der pU berücksichtigt die vom G-BA im Beschluss zur AbD geforderten Endpunkte wie folgt:

Mortalität

Der pU plant den Endpunkt Gesamtmortalität in der Operationalisierung Zeit bis zum Tod als Endpunkt in allen Patientenpopulationen zu erheben. Dieses Vorgehen ist sachgerecht.

Morbidität

Motorische Funktion (erhoben mit altersgeeigneten Instrumenten)

Gemäß Beschluss des G-BA soll die motorische Funktion über altersgeeignete Instrumente erfolgen. Der pU plant hierfür mehrere Instrumente einzusetzen (Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders [CHOP-INTEND], Hammersmith Functional Motor Scale Expanded [HFMSE], Revised Upper Limb Module [RULM], 6-Minuten-Gehtest und Bayley-III). Welche Instrumente eingesetzt werden sollen, unterscheidet der pU anhand der zu untersuchenden Populationen (präsymptomatisch, SMA Typ 1, Typ 2 oder Typ 3). Die folgende Tabelle fasst die geplanten Erhebungen zur motorischen Funktion in den entsprechenden Patientenpopulationen zusammen.

Tabelle 1: Geplante Instrumente und Operationalisierungen zur Erhebung der motorischen Funktion

Instrument	Population	Operationalisierung
CHOP-INTEND	SMA Typ 1	<ul style="list-style-type: none"> ▪ Änderung des Scores von Baseline zu einem Alter von 12, 24 und 36 Monaten ▪ Anteil der Patientinnen und Patienten mit einem Score ≥ 40 zu einem Alter von 12 Monaten
HFMSE	SMA Typ 2 und SMA Typ 3	<ul style="list-style-type: none"> ▪ Änderung des Scores von Baseline 12, 24 und 36 Monate nach Behandlungsbeginn
RULM	SMA Typ 2 und SMA Typ 3	<ul style="list-style-type: none"> ▪ Änderung des Scores von Baseline 12, 24 und 36 Monate nach Behandlungsbeginn
Bayley III	Präsymptomatisch und SMA Typ 1	<ul style="list-style-type: none"> ▪ Score in Subskalen zur Grobmotorik zu einem Alter von 12, 24 und 36 Monaten
6MWT	SMA Typ 3	<ul style="list-style-type: none"> ▪ Anteil der Patientinnen und Patienten mit einer Verbesserung/Verschlechterung über eine Gehstrecke von > 30 m 12, 24 und 36 Monate nach Behandlungsbeginn; ▪ Änderung von Baseline
6MWT: 6-Minuten-Gehtest; CHOP-INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; HFMSE: Hammersmith Functional Motor Scale Expanded; RULM: Revised Upper Limb Module		

Die Einsatzfähigkeit der Instrumente ist grundsätzlich abhängig von den motorischen Fähigkeiten ebenso wie dem Alter der Patientinnen und Patienten. Aufgrund der vom pU gewählten ausschließlich populationsabhängigen Einteilung ohne weitere Angaben zum Altersbereich der Patientinnen und Patienten, bei dem die Instrumente erhoben werden sollen, ist nicht nachvollziehbar, wie mit Patientinnen und Patienten umgegangen wird, die zwar den entsprechenden SMA-Typ aufweisen, jedoch noch nicht das Alter erreicht haben, ab dem der Einsatz der jeweiligen Instrumente empfohlen wird. Dies sollte in den Studienunterlagen entsprechend umgesetzt werden.

Für keines der genannten Instrumente legt der pU Unterlagen vor, mit denen deren Eignung beurteilt werden kann. Die einzelnen Instrumente werden im Folgenden beschrieben und ihre Eignung beurteilt.

CHOP-INTEND

Der CHOP-INTEND ist ein Instrument zur Erfassung der motorischen Funktionsfähigkeit. Er wurde für SMA Typ 1 Patientinnen und Patienten entwickelt und besteht aus 16 Domänen, die jeweils mit einem Wert von 0 (nicht funktional) bis 4 (voll funktional) bewertet werden. Es ergibt sich ein Gesamtwert von 0 bis 64 Punkten, wobei höhere Werte einer besseren motorischen Funktionsfähigkeit entsprechen. Der CHOP-INTEND wird auch in der AbD zu Onasemnogen-Abeparvovec zur Erhebung der motorischen Funktion in allen Patientenpopulationen eingesetzt und eignet sich für eine Altersspanne zwischen 1 und 38 Monaten. Laut Registerprotokoll des SMARtCARE-Registers (Version 5.0) wird der CHOP-

INTEND bei allen Kindern < 2 Jahren erhoben und darüber hinaus bei allen Kindern > 2 Jahre, die nicht sitzen können.

HFMSE

Mit dem HFMSE können motorische Fähigkeiten bei Patientinnen und Patienten mit SMA Typ 2 und 3 überprüft werden. Das Instrument umfasst 33 Testelemente und operationalisiert überwiegend grobmotorische Funktionen (Positionswechsel im Liegen, Krabbeln, Aufstehen über den Kniestand, Stehen, Gehen und Springen) [8-11]. Die 33 Testelemente werden jeweils auf einer Drei-Punkte-Skala (0: ist nicht in der Lage, 1: ist mit Hilfestellung in der Lage, 2: ist ohne Hilfe in der Lage) bewertet, sodass sich eine Maximalpunktzahl von 66 ergibt. Höhere Werte bedeuten eine bessere motorische Funktionalität. Der HFMSE eignet sich für Patientinnen und Patienten ab 2 Jahren [12]. Laut Registerprotokoll des SMARtCARE-Registers wird der HFMSE bei allen Kindern > 2 Jahre, die die Fähigkeit haben zu sitzen, erhoben.

RULM

Der RULM ist ein Instrument zur Untersuchung der motorischen Funktion der oberen Extremitäten bei Patientinnen und Patienten mit SMA Typ 2 und 3 [13,14]. Validität und Reliabilität wurden gezeigt. Das Instrument umfasst 19 Elemente zur Testung der proximalen und distalen motorischen Funktionen der Arme und Hände sowie ein Eingangs-Element zur Klassifikation der Funktionalität. 18 der 19 Testelemente werden mit einer Drei-Punkte-Skala (0: ist nicht in der Lage, 1: ist in der Lage [modifizierte Aufgabe], 2: ist ohne Schwierigkeiten in der Lage), 1 Testelement wird mit einer Zwei-Punkte-Skala (0: ist nicht in der Lage, 1: ist in der Lage) bewertet, sodass sich eine Maximalpunktzahl von 37 ergibt. Eine höhere Gesamtpunktzahl entspricht einem besseren Funktionsstatus. Der RULM bietet eine sinnvolle Ergänzung zum HFMSE, um speziell die Funktionalität der oberen Extremitäten zu überprüfen. Der RULM eignet sich für Patientinnen und Patienten \geq 30 Monate [12]. Laut Registerprotokoll des SMARtCARE-Registers wird der RULM bei allen Kindern > 2 Jahre, die in einem Rollstuhl sitzen können, erhoben.

Bayley III

Der Bayley III ist ein international etabliertes Instrument zur Beurteilung der Entwicklung vom Säuglingsalter bis zu einem Alter von 42 Monaten in 5 Bereichen: kognitives, sprachliches, motorisches, sozial-emotionales und adaptives Verhalten. Die Dauer der Durchführung des gesamten Verfahrens wird laut Manual für Kinder im Alter von < 12 Monaten mit 50 Minuten angegeben, für Kinder im Alter von \geq 13 Monaten mit 90 Minuten [15].

Die motorische Skala des Bayley III kann laut Manual separat eingesetzt werden, sie unterteilt sich in eine Subskala zur Grobmotorik (72 Aufgaben) und eine Subskala zur Feinmotorik (66 Aufgaben).

Laut Protokoll des SMArtCARE-Registers wird der Bayley III optional bei Kindern in einem Alter bis zu 42 Monaten eingesetzt.

6 Minuten Gehstest (6MWT)

Der 6MWT erfasst die Gehstrecke, die innerhalb von 6 Minuten zurückgelegt werden kann. Der 6MWT eignet sich für Patientinnen und Patienten ab 4 Jahren [12]. Laut Protokoll des SMArtCARE-Registers wird der 6 Minuten Gehstest bei allen Kindern > 2 Jahre erhoben.

Weitere Anmerkungen zur Operationalisierung der Endpunkte

Der pU wählt abhängig von der auszuwertenden Population unterschiedliche Operationalisierungen der Endpunkte zur Erhebung der motorischen Funktion. Diese werden im Folgenden in Abhängigkeit von der Population beschrieben und bewertet.

Präsymptomatische Patientinnen und Patienten mit einer 5q-assoziierten SMA und bis zu 3 Kopien des SMN2-Gens

Die motorische Funktion soll bei präsymptomatischen Patientinnen und Patienten ausschließlich über die Subskala zu den grobmotorischen Fähigkeiten des Bayley III erhoben werden. Für die Bayley Skala werden im Manual verschiedenen Auswertungsarten beschrieben (*scaled scores, composite scores, percentile ranks, and growth scores*). Aus den Angaben im SP / SAP geht nicht hervor, welchen Score der pU plant auszuwerten. Dies muss ergänzt werden.

Der pU plant die Veränderung des Scores im Vergleich zu Baseline im Alter von 12, 24 und 36 Monaten auszuwerten. Es ist anzunehmen, dass für diese Population das Alter der Patientinnen und Patienten annähernd der Zeit nach Behandlungsbeginn entspricht, wenn die Patientinnen und Patienten primär über das Neugeborenencreening identifiziert werden und eine Behandlung zeitnah nach Geburt begonnen wird. Dennoch sollte, insbesondere vor dem Hintergrund einer Vereinheitlichung mit der Datenerhebung zu Onasemnogen-Abeparvec, für alle Operationalisierungen der Auswertungszeitpunkt abhängig von der Zeit seit Behandlungsbeginn gewählt werden.

Der pU wählt für die Erhebung der motorischen Funktion bei den präsymptomatischen Patientinnen und Patienten ausschließlich die Subskala zur Grobmotorik des Bayley III. Im Sinne der Vergleichbarkeit mit der AbD zu Onasemnogen-Abeparvec ist es sinnvoll alternativ oder zumindest zusätzlich den CHOP-INTEND einzusetzen, der sich ebenfalls für die entsprechende Altersgruppe eignet.

Symptomatische Patientinnen und Patienten mit einer klinisch diagnostizierten Typ-1-SMA

Die motorische Funktion soll bei Patientinnen und Patienten mit SMA Typ 1 über den CHOP-INTEND erhoben werden. Der pU plant zum einen Auswertungen zum Anteil der Patientinnen und Patienten mit einem Score von ≥ 40 zum Alter von 12 Monaten und zum anderen

Auswertungen zu einer Veränderung im Gesamtscore im Vergleich zu Baseline im Alter von 12, 24 und 36 Monaten. Der pU liefert keine Angaben dazu, ob ein Score von 40 eine geeignete Responseschwelle für den CHOP-INTEND darstellt. Dies muss entsprechend nachgereicht werden. Darüber hinaus scheint eine Auswertung lediglich im Alter von 12 Monaten bei einer Beobachtungsdauer von insgesamt 36 Monaten nicht sinnvoll. Die gewählten Operationalisierungen sollten grundsätzlich den gesamten Beobachtungszeitraum abdecken.

Neben dem CHOP-INTEND wird analog zu den präsymptomatischen Patientinnen und Patienten auch der Bayley III zur Erhebung der grobmotorischen Funktion eingesetzt.

Symptomatische Patientinnen und Patienten mit einer klinisch diagnostizierten Typ-2-SMA und bis zu 3 Kopien des SMN2-Gens

Bei Patientinnen und Patienten mit einer Typ-2-SMA sollen der HFMSE und der RULM erhoben werden. Der pU plant für diese Endpunkte Auswertungen zur Änderung 12, 24 und 36 Monate nach Behandlungsbeginn im Vergleich zu Baseline.

Wie beschrieben, wird im Beschluss des G-BA auf den altersgerechten Einsatz der Instrumente zur Erhebung der motorischen Funktion verwiesen. Dies ist insbesondere bei Patientinnen und Patienten mit SMA Typ 2 von Bedeutung. Die Erkrankung manifestiert sich hier in der Regel in einem Alter zwischen 6 und 15 Lebensmonaten [16]. Sowohl der HFMSE als auch der RULM sind für dieses Alter, entsprechend der in den Tests zu erfüllenden Aufgaben (z. B. eine Tasse zum Mund führen), noch nicht empfohlen. Da der pU die Baselineerhebung zum Behandlungsstart plant, erscheint eine Operationalisierung über die Änderungen im Vergleich zu Baseline in dieser Population nicht sachgerecht. Für die adäquate Erhebung zu Baseline sind darüber hinaus auch die Empfehlungen in Abschnitt 2.2.3 zum Indexdatum zu beachten. Der pU sollte im SP spezifizieren, wie er sicherstellt altersgeeignete Instrumente zur Erhebung der motorischen Funktion einzusetzen. Dies betrifft insbesondere Operationalisierungen, in denen der pU Veränderungen gegenüber Baseline erheben möchte.

Symptomatische Patientinnen und Patienten mit einer klinisch diagnostizierten Typ-3-SMA und bis zu 3 Kopien des SMN2-Gens

Die Erhebung der motorischen Funktion erfolgt für die Patientinnen und Patienten mit einer SMA Typ 3 über den RULM und HFMSE analog zur Population der Patientinnen und Patienten mit SMA Typ 2. Die Erkrankung manifestiert sich ab einem Alter von 18 Monaten [12]. Entsprechend zur Population mit einem SMA Typ 2 sollte der altersgeeignete Einsatz der Instrumente daher ebenfalls geprüft werden.

Neben dem HFMSE und dem RULM plant der pU die 6-Minuten-Gehstrecke bei Patientinnen und Patienten mit SMA Typ 3 zu erheben. Die vom pU gewählten Operationalisierungen umfassen die 1) Verbesserung bzw. die Verschlechterung der zurückgelegten Gehstrecke um > 30 m 12, 24 und 36 Monate nach Therapiestart, sowie 2) die absolute Änderung und 3) die

relative Änderung der Gehstrecke gegenüber Baseline 12, 24 und 36 Monate nach Therapiestart. Der pU liefert keine Begründung dazu, warum er eine Veränderung von > 30 m als Responseschwelle wählt. Sofern nicht nachvollziehbar darzulegen ist, dass es sich bei dieser Responseschwelle um eine klinisch relevante Änderung in der betrachteten Patientenpopulation handelt, sind stetigen Analysen zu bevorzugen. Sofern verschiedene stetige Analysen geplant werden, sind diese zu hierarchisieren, bzw. ist eine primäre Operationalisierung zu wählen. Übergreifend macht der pU für die Operationalisierungen dieses Endpunkts die Einschränkung auf Patientinnen und Patienten, die zu Baseline gehfähig waren. Je nach Anteil der nicht gehfähigen Patienten kann dies zur Folge haben, dass dieser Endpunkt nicht mehr für die ganze Population aussagekräftig ist. Dies ist bei der Interpretation der Ergebnisse zu berücksichtigen. Darüber hinaus ist auch für dieses Instrument der altersgeeignete Einsatz zu prüfen. Der 6-Minuten-Gehtest ist erst ab einem Alter von 4 Jahren geeignet [12]. Der SMA Typ 3 kann sich aber bereits ab einem Alter von 18 Monaten manifestieren. Dies kann dazu führen, dass für einen relevanten Teil der Patientinnen und Patienten keine adäquate Baseline-Erhebung erfolgen kann. Insofern ist fraglich, ob die oben genannten Operationalisierungen, die den Baselinewert berücksichtigen, adäquat sind. Darüber hinaus ist auch fraglich, ob Erhebungen zu Monat 12 und ggf. 24 nach Behandlungsbeginn für diesen Patientinnen und Patienten sinnvoll interpretierbar sind. Eine Auswertung der Gehstrecke zu Monat 36 nach Behandlungsbeginn ohne Berücksichtigung der Baselinewerte ist in dieser Situation ggf. zu bevorzugen.

Erreichen motorischer Meilensteine („Motor development milestones“ der WHO)

Neben der oben beschriebenen Erfassung der motorischen Funktion über verschiedene Instrumente, plant der pU eine Auswahl von 2 der motorischen Meilensteine der WHO (Sitzen ohne Unterstützung und Gehen ohne Unterstützung) zu erfassen. In der Auswahl der zu erhebenden Meilensteine fehlt der Meilenstein Stehen ohne Unterstützung. Eine Begründung zur Nichtberücksichtigung dieses Meilensteins liefert der pU nicht. Der Meilenstein Stehen ohne Unterstützung sollte ergänzt werden.

Ebenfalls nicht erhoben wird ein Endpunkt zum Erhalt der motorischen Funktion. In der AbD zu Onasemnogen-Abeparvovec wird dieser Endpunkt als Zeit vom Erreichen eines motorischen Meilensteins bis zum vollständigen Verlust dieser Fähigkeit operationalisiert [7]. Dieser Endpunkt sollte ebenfalls ergänzt werden und die Auswertungsart spezifiziert werden.

Der pU plant die Auswertung der Meilensteine in Abhängigkeit von der betrachteten Population. So plant der pU für die präsymptomatischen Patientinnen und Patienten und in der Population mit SMA Typ 1 beide Meilensteine zu erheben, in der Population mit SMA Typ 2 nur den Meilenstein Gehen ohne Unterstützung. In der Population mit SMA Typ 3 plant er keine Auswertung der motorischen Meilensteine. Dies ist sachgerecht.

Nicht sachgerecht ist die vom pU gewählte Operationalisierung des Anteils der Patientinnen und Patienten, die den Meilenstein zu einem altersgerechten Zeitpunkt erreichen. Er plant dafür nur das Ergebnis der ersten Visite innerhalb eines Zeitfensters, bis zu dem gesunde Kinder diesen Meilenstein in der Regel erreichen (Sitzen: 9,5 bis 13,5 Monate; Gehen: 18 bis 22 Monate), heranzuziehen. Patientinnen und Patienten, die zu dieser Visite den Meilenstein nicht erreicht (oder bereits wieder verloren) haben, werden als Non-Responder gewertet. Mit der Wahl dieser Operationalisierung wird somit nur ein kleiner Ausschnitt der gesamten Beobachtungsdauer berücksichtigt. Vor diesem Hintergrund werden anstelle der Operationalisierung des pU Ereigniszeitanalysen oder Analysen zum Anteil der Patientinnen und Patienten, die zu einem zu definierenden Zeitpunkt einen motorischen Meilenstein erreicht haben (z. B. 12, 24, 36 Monate nach Behandlungsstart), empfohlen. Eine entsprechende Angabe des Auswertungszeitpunkts ist auch für den vom pU genannten Endpunkt Anteil der Patientinnen und Patienten, die den Meilenstein Gehen ohne Unterstützung erreicht haben, zu ergänzen.

Respiratorische Funktion (Notwendigkeit der [dauerhaften] Beatmung)

In der AbD sollen Endpunkte zur respiratorischen Funktion zum einen im kombinierten Endpunkt Tod / dauerhafte Beatmung und zum anderen als Einzelkomponente erhoben werden, jeweils als Ereigniszeitanalysen. Eine dauerhafte Beatmung ist dabei operationalisiert als eine Beatmung über täglich mindestens 16 Stunden an mindestens 21 aufeinanderfolgenden Tagen. In der AbD zu Onasemnogen-Abeparvec wird die dauerhafte Beatmung als 2 aufeinanderfolgende Dokumentationen einer Beatmung für ≥ 16 Stunden/Tag operationalisiert [7]. Warum der pU erst ab 21 aufeinanderfolgenden Tagen von einer dauerhaften Beatmung ausgeht, ist nicht nachvollziehbar. Es wird empfohlen die Operationalisierung der dauerhaften Beatmung an die in der AbD von Onasemnogen-Abeparvec anzugleichen.

Neben der dauerhaften Beatmung plant der pU auch die Zeit bis zur Erforderlichkeit jeglicher respiratorischer Unterstützung zu erheben. Die respiratorische Unterstützung ist dabei als Nutzung eines Beatmungsgeräts operationalisiert. Die vom pU zur Operationalisierung genannten Dokumentationsfelder des SMARtCARE-Registers enthalten darüber hinaus die Angabe, ob es sich um eine invasive oder nicht invasive Beatmung handelt. Es ist unklar, wie diese Information in der Operationalisierung berücksichtigt werden soll oder ob dies nicht vielmehr in die Operationalisierung eines separaten Endpunkts eingehen müsste („Art der Beatmung“ wie in der AbD zu Onasemnogen-Abeparvec [7]).

bulbäre Funktion (Schluck- und Sprachfähigkeit, Notwendigkeit nicht oraler Ernährungsunterstützung)

Die bulbäre Funktion, erhoben über die Verschlechterung der Schluckfunktion und die Notwendigkeit einer nicht oralen Ernährung und sind jeweils über den Anteil der Patientinnen

und Patienten zum Alter von 12, 24 und 36 Monaten (präsymptomatische Patientinnen und mit SMA Typ 1) bzw. 12, 24 und 36 Monate nach Therapiestart (SMA Typ 2 und 3) operationalisiert. Das ist grundsätzlich sachgerecht, weicht aber beispielsweise von der Operationalisierung in der AbD zu Onasemnogen-Abeparvovec ab, in der jeweils Ereigniszeitanalysen für die Endpunkte geplant sind. Dies sollte wie oben beschrieben vereinheitlicht werden.

Die Sprachfähigkeit soll über die Subskalen expressive und rezepptive Sprache des Bayley III zu einem Alter von 12, 24 und 36 Monaten bei präsymptomatischen Patientinnen und Patienten und Patientinnen und Patienten mit SMA Typ 1 erhoben werden. Die Entwicklung der Sprachfähigkeit erscheint auch bei Patientinnen und Patienten mit SMA Typ 2 und Typ 3 sinnvoll und sollte ergänzt werden. Die rezepptive Sprache wird im Bayley über 49 Aufgaben und die expressive Sprache über 48 Aufgaben erfasst. Die Erfassung des Bayley III ist im SMARtCARE bislang keine verpflichtende Untersuchung.

Der pU plant als Operationalisierung den Anteil der Patientinnen und Patienten zu erheben, die altersgerechte Scores in den Subskalen expressive Sprache und rezepptive Sprache im Alter von 12, 24 und 36 Monaten zeigen. Altersgerechte Scores sollen dabei definiert werden als Altersäquivalente, die im Bayley III Manual [15] beschrieben werden und das durchschnittliche Alter angeben zu dem ein bestimmter Rohwert typisch ist. Der geringste Score, der für das entsprechende Alter ± 2 Monate angegeben ist wird als untere Grenze angenommen. Bereits im Manual wird als Limitation der Altersäquivalente genannt, dass diese keine gleichmäßig verteilten Intervalle innerhalb einer Skala abbilden und daher kleine Änderungen der Rohwerte zu großen Änderungen der Äquivalente führen können. Die Operationalisierung sollte entsprechend angepasst werden, z. B. als Veränderung im Vergleich zu Baseline.

Weitere Komplikationen der Erkrankung (z.B. Schmerz, orthopädische Komplikationen)

Das Auftreten von regelmäßigen Schmerzen und Fatigue soll jeweils als Anteil zu einem Alter von 12, 24 und 36 Monaten bei präsymptomatischen Patientinnen und Patienten und Patientinnen und Patienten mit SMA Typ 1 bzw. 12, 24 und 36 Monate nach Behandlungsbeginn bei Patientinnen und Patienten mit SMA Typ 2 und 3 erhoben werden. Auf Basis der vorliegenden Angaben scheint die Erhebung lediglich durch eine Abfrage, ob Schmerzen oder Fatigue vorliegen zu erfolgen, ohne dass ein bestimmtes Instrument zur Erhebung dieser Endpunkte verwendet wird. Die alleinige Abfrage, ob Schmerzen oder Fatigue vorliegen, ohne Verwendung eines geeigneten, für alle Patientinnen und Patienten verwendeten Instruments, ist keine valide Erhebung dieser beiden Symptome. Der pU sollte prüfen, ob altersgerechte, valide Instrumente zur Erhebung dieser Symptome vorliegen und diese entsprechend anwenden.

Orthopädische Komplikationen werden über einen kombinierten Endpunkt aus schwerer Skoliose (Cobb-Winkel ≥ 40 Grad) und einer orthopädischen Operation erhoben. Dargestellt wird die Zeit bis zum Auftreten des kombinierten Endpunkts, sowie die Zeit bis zum Auftreten der Einzelkomponenten des kombinierten Endpunkts (Zeit bis zur ersten Dokumentation einer schweren Skoliose bzw. zur ersten Dokumentation einer orthopädischen Operation). Der pU begründet nicht, warum er die Skoliose lediglich auf schwere Ereignisse einschränkt. In der AbD zu Onasemnogen-Abepravovec wird für den kombinierten Endpunkt jegliches Auftreten einer Skoliose dokumentiert [7]. Eine Vereinheitlichung der Operationalisierungen ist sinnvoll, mindestens sollte die Erfassung jeglichen Auftretens einer Skoliose ergänzt werden.

Der pU ergänzt die vom G-BA festgelegten Endpunkte um einen Endpunkt zu geplanten Hospitalisierungen. Der pU plant die Anzahl der geplanten Hospitalisierungen, ohne Berücksichtigung von Hospitalisierung zur Applikation der SMA-Medikation, zu erheben. Es bleibt unklar, in welchem Zusammenhang diese geplanten Hospitalisierungen mit der durch die SMA verursachte Morbidität stehen und wie sie von den bereits erhobenen Endpunkten (z. B. zu orthopädischen Operationen) abgegrenzt werden können. Insofern ist der Endpunkt in der geplanten AbD verzichtbar.

Unerwünschte Ereignisse

Gemäß Forderungsbeschluss des G-BA sollen für die AbD folgende UEs erfasst werden:

- schwerwiegende unerwünschte Ereignisse (SUE)
- unerwünschte Ereignisse, die zu einer Hospitalisierung führen
- schwerwiegende spezifische unerwünschte Ereignisse:
 - Netzhaut-Toxizität
 - Wirkung auf Epithelgewebe
 - Thrombozytopenie
 - renale Toxizität
 - Hydrozephalus
 - Hepatotoxizität
 - kardiale Ereignisse
 - Entzündung der Spinalganglionzellen

Im SP ist die Erhebung der vom G-BA geforderten UE Endpunkte vorgesehen, bezüglich der Operationalisierung und der geplanten Auswertungen ergeben sich allerdings Kritikpunkte:

Der pU plant für alle UE-Endpunkte – ausgenommen dem Endpunkt ungeplante Hospitalisierungen – Ereigniszeitanalysen. Bei vergleichbarer Beobachtungsdauer sollten Auswertungen zum Anteil der Patientinnen und Patienten mit Ereignis bei den UE-Endpunkten vorgezogen werden.

Es fehlt eine Beschreibung der Operationalisierung der SUEs. Analog zur AbD zu Onasemnogen-Abeparvovec stellt die Annäherung über UEs, die zur Hospitalisierung oder einer Verlängerung der Hospitalisierung führen bzw. UEs, die zum Tod führen eine sachgerechte Operationalisierung dar. Dies ist im SP/SAP entsprechend klarzustellen und die entsprechenden Dokumentationsfelder im SMArtCARE-Register zu ergänzen.

Für die spezifischen SUEs weicht die Benennung zu denen der AbD zu Onasemnogen-Abeparvovec ab. Da die für Onasemnogen-Abeparvovec definierten spezifischen UEs neben der allgemeinen Erfassung von UEs aktiv im SMArtCARE-Register abgefragt werden, ist für die gleichen spezifischen UEs in beiden AbDs auf eine einheitliche Benennung zu achten (z. B. Hepatotoxizität bei Onasemnogen-Abeparvovec und Hepatopathie bei Risdiplam).

Darüber hinaus plant der pU die Anzahl ungeplanter Hospitalisierungen zu erheben. Es ist zunächst unklar, wie dieser Endpunkt von den bereits erhobenen UEs, die zu einer Hospitalisierung führen abzugrenzen ist. Anhand der vom pU im SAP gewählten Dokumentationsfelder ergibt sich, dass die Operationalisierung deckungsgleich ist mit der Operationalisierung zu UEs, die zu einer Hospitalisierung führen und es sich entsprechend um den gleichen Endpunkt handelt, der damit verzichtbar ist. Wie oben beschrieben sind für den verbleibenden Endpunkt UEs, die zu einer Hospitalisierung führen anstelle der vom pU geplanten Operationalisierungen (Anzahl Ereignisse bzw. Ereigniszeitanalyse) bei vergleichbarer Beobachtungsdauer Auswertungen zum Anteil der Patientinnen und Patienten mit Ereignis zu bevorzugen.

2.2.2 Datenquelle

Der pU wählt das SMArtCARE-Register als Datenquelle für die AbD. Das Register ist für die AbD geeignet, da es grundsätzlich die notwendigen Qualitätskriterien erfüllt [3], und wurde vom G-BA als primär relevantes Register benannt [1]. Das SMArtCARE-Register stellt darüber hinaus auch die primäre Datenquelle für die anwendungsbegleitende Datenerhebung zu Onasemnogen-Abeparvovec dar und wurde bereits in vielen Teilen an die Anforderungen der AbD angepasst [7]. Wie bereits beschrieben, liegt das aktuelle Kodierhandbuch des SMArtCARE-Registers nicht vor, wodurch die Überprüfung der Operationalisierungen im Register nicht für alle Daten (Endpunkte, Confounder, etc.) möglich ist.

Der G-BA weist darüber hinaus auf die Einbindung weiterer Register hin, sofern diese die notwendigen Anforderungen erfüllen. Der pU beschreibt im Protokoll, dass er ggf. weitere Register einbinden möchte, macht darüber hinaus aber keine weiteren Angaben. Für die AbD

ist es erforderlich die Einbindung weiterer Register vorab zu benennen und die Eignung der jeweiligen Register entsprechend der vom G-BA im Beschluss genannten Kriterien zu prüfen. Vor dem Start der AbD müssen alle als Datenquelle dienenden Register feststehen und an die Anforderungen der AbD angepasst sein.

Weder im SP noch im SAP gibt es Angaben dazu, wie häufig die Meldung an das SMARtCARE-Register erfolgt. Die Dokumentation der erhobenen Daten sollte grundsätzlich möglichst für alle Patienten einheitlich direkt im Anschluss an die jeweilige Visite erfolgen, um Meldeverzögerungen zu vermeiden.

2.2.3 Indexdatum, Zuordnung zu den Behandlungsgruppen und Beobachtungsdauer

Indexdatum

Der pU beschreibt als Indexdatum, d. h. als Datum für den Beobachtungsstart, die Visite festzulegen, die zwischen 6 Wochen vor und 3 Wochen nach der ersten Applikation der SMA-Medikation liegt. Dieses Vorgehen ist nicht sachgerecht. Es ist nicht adäquat, für einen Teil der Patientinnen und Patienten das Indexdatum vor und für einen anderen Teil der Patientinnen und Patienten nach dem Start der Behandlung zu erfassen. Wie in der Niederschrift zur Beratungsgespräch [6] ausgeführt, sollte als Indexdatum entsprechend des Vorgehens einer Target Trial Emulation der Tag der Therapieentscheidung herangezogen werden. Sollte dies nicht möglich sein, sollte eine bestmögliche Näherung daran erfolgen. Dadurch, dass die Applikation der SMA-Medikation ggf. nicht unmittelbar nach der Therapieentscheidung erfolgt und insbesondere bei einer Therapie mit Onasemnogen-Abeparvovec Zeit zwischen der Therapieentscheidung und der Applikation vergehen kann, kann es erforderlich sein, eine Brückentherapie einzusetzen. Diese Brückentherapie ist als Teil des Therapiekonzepts zu betrachten und muss daher durch den Beobachtungszeitraum abgedeckt werden.

Zuordnung zu den Behandlungsgruppen

Der pU beschreibt nicht explizit wie die Zuordnung der Patientinnen und Patienten zu den Behandlungsgruppen erfolgt. Auf Basis der vom pU gewählten Estimands ist davon auszugehen, dass die Zuordnung zu den Behandlungsgruppen auf Basis der Therapie erfolgt, die die Patientinnen und Patienten zum Zeitpunkt des Einschlusses in die AbD erhalten haben. Dies ist entsprechend der Ausführungen im Beratungsgespräch [6] sachgerecht.

Unklar bleibt dabei jedoch, wie mit Patientinnen und Patienten, die eine Brückentherapie (z. B. bis zum Einsatz der Gentherapie) erhalten haben, umgegangen wird. Dies ist zu ergänzen. Wie im Beratungsgespräch [6] ausgeführt, könnte eine solche Brückentherapie bei einigen Patientinnen und Patienten eine Rolle spielen. Da eine Brückentherapie nicht auf Weiterbehandlung ausgelegt ist, kann sie den nachfolgenden Therapien zugeordnet werden.

Beobachtungsdauer

Der pU macht weder im SP noch im SAP Angaben dazu, wie lange die Patientinnen und Patienten nachbeobachtet werden sollen. Laut G-BA Beschluss soll die Nachbeobachtung 36 Monate betragen. Das Studienende ist derzeit für den 1. Januar 2026 geplant. Somit liegt zwischen dem geplanten Start der AbD im 4. Quartal 2023 und dem Studienende nur eine Zeit von maximal 27 Monaten. Somit wäre es für neu in das Register aufgenommene Patientinnen und Patienten nicht möglich diese 36 Monate nachzubeobachten. Davon werden insbesondere die Patientinnen und Patienten betroffen sein, die entsprechend der Anwendungsgebietserweiterung von Risdiplam ab einem Alter von 0 Monaten behandelt werden. Zur Vermeidung verkürzter Beobachtungszeiten müssen entsprechende Anpassungen vorgenommen werden.

Im Zusammenhang mit der Auswertung der UE-Endpunkt beschreibt der pU, dass er für den Endpunkt ungeplante Hospitalisierungen nur Ereignisse vor dem Behandlungswechsel in die Analyse einschließen will und die Nachbeobachtung „für diese Patientinnen und Patienten“ mit dem Behandlungswechsel endet. Es bleibt unklar, auf welche Patientinnen und Patienten er diese Aussage bezieht. Grundsätzlich sind für die AbD alle Patientinnen und Patienten für mind. 36 Monate nachzubeobachten, unabhängig von einem Behandlungswechsel.

2.2.4 Studienpopulation der Registerstudie

2.2.4.1 Fallzahlplanung

Im SAP und SP gibt der pU für jede der 4 Populationen Angaben zur Fallzahlplanung an. Bezüglich der präsymptomatischen Patientinnen und Patienten (Fragestellung 1) sowie der symptomatischen Patientinnen und Patienten mit Typ-3-SMA (Fragestellung 4) gibt der pU an zunächst auf eine Fallzahlkalkulation zu verzichten, da aktuell nicht genügend Evidenz vorliege, um diese im ausreichenden Maß zu informieren. Eine entsprechende Fallzahlkalkulation soll unter Berücksichtigung der Ergebnisse der ersten Interimsanalyse vorgenommen werden. Dieses Vorgehen ist grundsätzlich nachvollziehbar.

Als Endpunkt für die Fallzahlschätzung wählt der pU für die Fragestellung 1 ungeplante Hospitalisierungen operationalisiert als Anzahl ungeplanter Hospitalisierungen, mit dem entsprechenden Effektmaß Rate Ratio. Es wird davon ausgegangen, dass die ungeplanten Hospitalisierungen deckungsgleich zum UE-Endpunkt UEs die zu einer Hospitalisierung führen anzusehen sind. Die Durchführung der Fallzahlschätzung für einen UE-Endpunkt ist grundsätzlich sachgerecht [6], allerdings ist die Operationalisierung des Endpunkts wie in Abschnitt 2.2.1.3 beschrieben anzupassen.

Für Fragestellung 4 und Fragestellung 3 (Patientinnen und Patienten mit einer Typ-2-SMA) gibt der pU an die Fallzahlplanung auf Grundlage des RULM, operationalisiert als Änderung zu Monat 12 im Gesamtscore zu Baseline mit dem entsprechenden Effektmaß Cohen's d (als

SMD) vorzunehmen. Wie in Abschnitt 2.2.1.3 beschrieben, ist die vom pU verwendete Operationalisierung nicht sachgerecht, die Fallzahlplanung für die Fragestellungen 3 und 4 sind entsprechend anzupassen.

Für die symptomatischen Patientinnen und Patienten mit einer Typ-1-SMA (Fragestellung 2) gibt der pU an die Fallzahlplanung auf Grundlage des kombinierten Endpunkts Tod oder andauernde Beatmung, operationalisiert über die Zeit bis zum Ereignis, vorzunehmen. Die hierbei angenommenen Ereignisanteile für den Vergleichsarm mit 22 % sowie die zu erwartenden Effektstärken, auf denen die Fallzahlplanung beruht, sind nicht nachvollziehbar. Die Annahmen für die vorgelegte Fallzahlplanung sollten entsprechend genauer begründet werden.

Auffällig ist, dass sowohl für Fragestellung 2 als auch für Fragestellung 3 die Annahmen für die Effektstärke jeweils nahe an der vom pU formulierten verschobenen Nullhypothese liegen. Dieses Vorgehen führt zwangsläufig zu sehr großen Fallzahlen. Die angenommenen Effektstärken sind nicht ausreichend begründet. Sofern der pU die Annahmen begründen kann, sollen die für die Fallzahlplanung herangezogenen Endpunkte überdacht werden, da ansonsten bezogen auf die für die AbD zur Verfügung stehenden Patientinnen und Patienten unrealistisch hohe Fallzahlen resultieren.

Wie bereits beschrieben, geht aus der Niederschrift des G-BA hervor [6], dass die Fallzahlplanung auch auf Basis von Endpunkten zu Nebenwirkungen beruhen kann, um eine Überlegenheit bei vergleichbarem Nutzen nachzuweisen. Es ist vom pU deshalb zu prüfen, ob ggf. andere Endpunkte für die Fallzahlplanung herangezogen werden können, um eine adäquate Fallzahl zu erreichen.

Weiterhin ist anzumerken, dass der pU sich nicht dazu äußert, auf welchen Verteilungsverhältnissen bezüglich der Behandlungsarme die jeweiligen Fallzahlplanungen beruhen. Dies ist ebenfalls anzugeben.

2.2.5 Identifikation und Erhebung von Confoundern, Adjustierung für Confounder

Einen zentralen Aspekt bei vergleichenden Studien ohne Randomisierung stellt die adäquate Adjustierung für Confounder dar, um interpretierbare Schätzungen des interessierenden Effekts zu erhalten. Für eine adäquate Kontrolle für Confounder ist es daher unabhängig von der verwendeten Methodik zur Adjustierung notwendig [17-19],

- vorab alle wichtigen Confounder (inklusive wichtiger Wechselwirkungen) zu identifizieren und diese wichtigen Confounder in der Registerstudie vollständig zu erheben sowie
- die wichtigen Confounder in angemessener Form bei der Auswertung im Modell zu berücksichtigen.

Das diesbezüglich geplante Vorgehen des pU wird im Folgenden beschrieben und bewertet.

2.2.5.1 Identifikation und Erhebung wichtiger Confounder

Zur Identifikation der für die vorliegende Fragestellung relevanten Confounder verweist der pU auf 3 Quellen: 1) auf die Ergebnisse einer systematischen Literaturrecherche im Studienprotokoll zu Onasemnogen-Abeparvovec (Version 2.02 vom 18.11.2021) [20] sowie 2) auf das Ergebnis einer orientierenden Suche nach prognostischen Faktoren und Effektmodifikatoren im Rahmen des MAIC aus dem Nutzendossier zur Bewertung von Risdiplom [21,22]. Darüber hinaus berücksichtigt der pU 3) die Hinweise aus der Beratung mit dem G-BA und IQWiG (2022-B-215) [6,23,24] bei der Anpassung der Confounder (siehe auch Anhang A). Darüber hinaus beschreibt der pU, dass die Confounder-Definition der erneuten Nutzenbewertung von Onasemnogen-Abeparvovec übernommen werden soll, um die beiden AbDs konsistenter zu machen.

Das Vorgehen des pU zur Identifizierung der Confounder ist grundsätzlich sachgerecht. Im Gegensatz zur AbD von Onasemnogen-Abeparvovec sind jedoch auch Patientinnen und Patienten mit SMA Typ 3 Teil der Fragestellung. Der pU beschreibt selbst im Studienprotokoll, dass überprüft werden muss, ob weitere Confounder für diese Patientenpopulation von Relevanz sind. Er verweist in der Folge auf die für die Nutzenbewertung zu Risdiplom vorgelegten Confounder der MAICs für Patientinnen und Patienten mit SMA Typ 3. Dies ist nicht nachvollziehbar, da im Dossier zur genannten Nutzenbewertung für Typ 3 SMA keine Confounderauswahl vorgelegen hat [21]. Es kann somit nicht ausgeschlossen werden, dass für Patientinnen und Patienten mit Typ 3 SMA noch weitere relevante Confounder zu berücksichtigen sind. Dies muss vom pU systematisch überprüft werden.

2.2.5.2 Berücksichtigung wichtiger Confounder bei der Analyse

Bezüglich der vom pU ausgewählten Confounder ist nicht ersichtlich, auf welche Weise diese in das Modell zur Propensity Score (PS) Berechnung eingehen (z. B. stetig oder dichotom). Weitere Unklarheit besteht bezüglich der vom pU vorgenommenen Einteilung in „less important“ und „very important“. Welchen Einfluss diese Einteilung auf die Modellierung des PS hat, ist anhand der Ausführungen des pU nicht ersichtlich. Dies ist vom pU zu beschreiben. Grundsätzlich sollten alle relevanten Confounder im Modell berücksichtigt werden.

Auswahl des Verfahrens für die Confounderadjustierung

Der pU beschreibt, dass er zur Confounderadjustierung gewichtungsbasierte PS Verfahren verwenden will. Hierbei benennt er das Inverse Probability of Treatment Weighting (IPTW) sowie das Verfahren der Fine Stratification Weights (FSW). Für die FSW wählt der pU eine feste Stratabreite von 0,1, was einer Anzahl von 10 Strata entspricht. Das beschriebene Vorgehen ist sachgerecht.

Zum Umgang mit extremen Gewichten bei Nutzung des IPTW-Verfahrens plant der pU Propensity Scores die $< 0,05$ bzw. $> 0,95$ liegend mittels des entsprechenden Schwellenwerts zu ersetzen. Ein solches Vorgehen ist durch geeignete Literatur zu begründen. Sofern keine geeignete Literatur vorliegt, sollte das Verfahren angepasst werden, z. B. durch das Trunkieren von extremen Gewichten.

Als Entscheidungskriterium dafür, welches der 2 beschriebenen Verfahren als primäre Analyse genommen wird, gibt der pU ein 2-stufiges Verfahren an. Zunächst soll das Verfahren mit der besseren Balanciertheit herangezogen werden. Die Balanciertheit ist hinreichend, wenn die standardisierte Mittelwertdifferenz (SMD) für alle Confounder $< 0,2$ ist (siehe unten). Weisen beide Verfahren eine hinreichende Balanciertheit auf, soll das Verfahren mit der besten allgemeinen Balanciertheit („overall measure of balance“) verwendet werden. Als mögliches Maß für diese gibt der pU die C-Statistik in der gewichteten Population an. Dieses Vorgehen ist sachgerecht.

Für die Untersuchung der Überlappung soll die flächenmäßige Überlappung der PS-Score-Dichten berechnet werden. Wenn diese mindestens 50 % beträgt, wird die Überlappung als ausreichend angesehen. Dieses Vorgehen entspricht dem Vorgehen in der AbD zu Onasemnogen-Abeparvovec [7]. Da aktuell kein etabliertes Kriterium zur Einordnung der Überlappung existiert, gibt der pU des Weiteren an, sein Vorgehen ggf. zu ändern, sobald adäquate Kriterien existieren. Das vom pU beschriebene Vorgehen zur Untersuchung der Überlappung ist sachgerecht.

Zur Bewertung der Balanciertheit wird die SMD für die einzelnen Confounder herangezogen. Dabei wird von einer ausreichenden Balanciertheit ausgegangen, wenn bei allen Confoundern ein Unterschied von höchstens 0,2 SMD festgestellt wird. Die in der Literatur vorgeschlagenen akzeptablen Grenzen für eine ausreichende Balanciertheit liegen im Bereich zwischen 0,1 und 0,25 [25]. Die vom pU vorgeschlagene Grenze von 0,2 ist zwar sachgerecht, stellt allerdings ein schwaches Kriterium für eine ausreichende Balanciertheit dar. Dies ist bei der Interpretation der Ergebnisse auf Basis einer verschobenen Hypothesengrenze entsprechend zu berücksichtigen [26].

Sofern bezüglich des PS Scores keine ausreichende Überlappung und Balanciertheit erreicht werden kann, plant der pU naive Analysen durchzuführen. Hierbei geht er nicht darauf ein, welche Auswirkungen dies auf die Interpretation der Ergebnisse hat.

Beschreibung der Zielpopulation nach PS-Gewichtung

Ein zentraler Aspekt bei nicht randomisierten Vergleichen mit Confounderadjustierung ist die detaillierte Beschreibung der Zielpopulation, die sich aus der Anwendung eines PS-Verfahrens ergibt. Die durch das PS-Verfahren gebildete Patientenpopulation (= Analysepopulation) muss repräsentativ für die jeweilige Fragestellung sein. Sollten relevante Unterschiede zwischen der

Analysepopulation und der ursprünglichen Gesamtpopulation erkennbar sein, ist anzugeben auf welche Teilpopulation sich die Ergebnisse der durchgeführten Analysen beziehen. Eine detaillierte Beschreibung des pU bezüglich der Prüfung der Analysepopulation und Zielpopulation fehlen in den vorgelegten Dokumenten. Dies ist vom pU zu überarbeiten.

2.2.6 Anmerkungen zu den geplanten Auswertungen für die Endpunkte der AbD

Im SP und SAP sind zu den verschiedenen Operationalisierungen der Endpunkte jeweils endpunktspezifische Ausführungen zu den geplanten Auswertungen enthalten. Diese werden nachfolgend adressiert.

Der pU gibt an, für Responderanalysen drei Effektmaße angeben zu wollen, attributables Risiko, Odds Ratio und relatives Risiko, jeweils mit 95%-Konfidenzintervall. Dabei gibt der pU nicht an, welches Modell er zur Berechnung verwenden will, ebenfalls fehlen Angaben zur geplanten Teststatistik. Diese Angaben sind zu ergänzen.

Zur Betrachtung von Ereigniszeitanalysen plant der pU die Verwendung des Cox Proportional Hazards Modell zur Schätzung des Effektmaßes Hazard Ratio. Als Teststatistik ist der Log-Rank-Test angegeben. Zur grafischen Darstellung sollen Kaplan-Meier-Kurven verwendet werden. Dieses Vorgehen ist sachgerecht.

Für stetige Auswertungen gibt der pU an, ein gemischtes Modell für Messwiederholungen (Mixed Model Repeated measures [MMRM]) mit den Kovariablen Behandlung, Visite, Interaktion zwischen Behandlung und Visite, adjustiert für den Wert zu Studienbeginn, verwenden zu wollen. Als Effektmaß zieht der pU Cohen's d nebst 95%-Konfidenzintervall heran. Eine genaue Definition des Effektmaßes Cohen's d im Zusammenhang mit der MMRM-Analyse ist nicht beschrieben. Weiterhin ist eine grafische Darstellung der Ergebnisse geplant. Angaben zur Teststatistik macht der pU nicht. Das Vorgehen ist prinzipiell sachgerecht, jedoch sollten die fehlenden Angaben zum Effektmaß und zur Teststatistik vom pU adressiert werden.

Für die stetigen Auswertungen zum Endpunkt 6MWT gibt der pU an, ebenfalls Cohen's d als Effektmaß verwenden zu wollen. Dabei ist zu beachten das die Ergebnisse dieses Endpunkts auf einer natürlichen Skala beruhen. Daher ist die Verwendung von Cohen's d nicht sachgerecht, vielmehr sollte die Interpretation der Relevanz der Ergebnisse auf Basis der Skala des Instruments (also der Wegstrecke) erfolgen. Dies ist vom pU zu ändern.

Der pU beschreibt, dass die UE-Endpunkte unter Verwendung einer Hypothetical Strategy ausgewertet werden sollen. Dazu werden in beiden Behandlungsarmen die Patientinnen und Patienten zum Zeitpunkt des Behandlungswechsels zensiert. Dies ist nicht sachgerecht (siehe auch Abschnitt 2.2.9 zu Estimands).

Sensitivitätsanalysen

Der pU gibt an, als Sensitivitätsanalysen Analysen zu wählen, bei denen Beobachtungen nach jeglichem Behandlungswechsel nicht in die Analyse eingehen (siehe auch Abschnitt 2.2.3). Dieses Vorgehen ist sachgerecht.

Darüber hinaus sollten, wie in der Beratung des pU [6] angemerkt, aufgrund einer möglichen Heterogenität zwischen den beiden Therapieoptionen im Vergleichsarm, Heterogenitätsuntersuchungen in der Datenauswertung durchgeführt werden. Solche Untersuchungen hat der pU bislang nicht geplant, und sollte dies entsprechend als Sensitivitätsanalysen in den Studienunterlagen ergänzen.

2.2.7 Untersuchung auf potenzielle Effektmodifikatoren und Subgruppenanalysen

Im Beschluss wird gefordert, dass der pU Angaben zu Subgruppenanalysen anhand der Kopienanzahl des SMN2-Gens für präsymptomatische Patientinnen und Patienten vorlegen soll, um zu überprüfen, ob eine gemeinsame Auswertung sachgerecht ist [1]. Der pU plant Subgruppenanalysen für das Merkmal SMN2-Kopienzahl in allen Patientenpopulationen durchzuführen. Dieses Vorgehen ist sachgerecht.

Darüber hinaus plant der pU für verschiedene Merkmale Subgruppenanalysen durchzuführen. Die Wahl der potenziellen Effektmodifikatoren entspricht dabei weitestgehend denen, die auch in der AbD zu Onasemnogen-Abeparvovec untersucht werden sollen. So plant er anstelle des Alters zu Behandlungsbeginn das Alter zu Studieneinschluss (präsymptomatisch, SMA Typ 2 und SMA Typ 3) bzw. das Alter bei Diagnose (SMA Typ 3) als Subgruppenmerkmale auszuwerten. Dies sollte vereinheitlicht und auf alle zu untersuchenden Populationen ausgeweitet werden.

Grundsätzlich sollten inhaltlich sinnvolle Trennwerte gewählt werden, die sich ggf. zwischen den einzelnen zu untersuchenden Populationen unterscheiden. Die Einteilung der Subgruppen auf Basis des Medians ist ohne inhaltliche Rationale nicht sachgerecht.

Der pU gibt nicht an, mit welchen Methoden er die Subgruppenanalysen durchführen möchte. Weiterhin fehlen Angaben dazu, wie der pU mit den Ergebnissen bezüglich deren Interpretierbarkeit umgehen will. Diese Angaben müssen ergänzt werden.

2.2.8 Qualität und Vollständigkeit der Datenerhebung

2.2.8.1 Maßnahmen zur Sicherstellung der Qualität und Vollständigkeit der Datenerhebung

Zur Sicherstellung der Qualität und Vollständigkeit der Datenerhebung werden gemäß SP SAS-Plausibilitätsprüfungen durchgeführt. Dies sollte im SP konkretisiert werden.

Gemäß Beschluss [1] beschreibt der pU, dass eine Source Data Verification für 100 % der Patientinnen und Patienten je Erhebungszentrum für den primären Endpunkt und für mindestens 10 % zufällig ausgewählter Patienten je Erhebungszentrum für alle weiteren Endpunkte über den Zeitraum seit Beginn der Datensammlung durchgeführt werden soll. Dies ist sachgerecht.

Insgesamt macht der der pU keine Angaben zu den Konsequenzen, die aus diesen Prüfungen gezogen werden. Dies sollte ergänzt werden.

Spezifische Maßnahmen zur Sicherstellung der Vollständigkeit der Datenerhebungen benennt der pU jedoch nicht. Für relevante Daten, die im Zuge der AbD erhoben werden sollen, müssen alle relevanten Datenfelder spätestens mit Start der AbD verpflichtend zu dokumentieren sein.

2.2.8.2 Umgang mit fehlenden Daten

Gemäß des G-BA-Beschlusses zur AbD sind Angaben zum erwarteten Umfang und den Gründen für fehlende Daten, sowie Maßnahmen zur Vermeidung fehlender Daten und Auswertungsstrategien zum Umgang mit fehlenden Daten darzulegen [1].

Der pU plant fehlende Datumsangaben wie folgt zu ersetzen: Fehlt bei Nutzenendpunkten nur der Tag, wird der letzte Tag des jeweiligen Monats angenommen. Handelt es sich um einen Nebenwirkungsendpunkt, wird der 1. Tag des Monats ersetzt. War der Patient in diesem Monat ereignisfrei, wird der Tag in diesem Monat ersetzt, zu dem die Patientin oder der Patient zuletzt ereignisfrei war. Sofern der Monat fehlt, wird das Datum verwendet, an dem der Patient, die Patientin zuletzt als ereignisfrei galt. Während die geplante Ersetzung des fehlenden Tages sachgerecht ist, ebenso wie die geplante Ersetzung des Monats für Patientinnen und Patienten ohne Ereignis, führt die geplante Ersetzung des Monats für Patientinnen und Patienten mit Ereignis potenziell zu Verzerrungen und ist nicht sachgerecht. Es sind vielmehr entsprechende Anstrengungen erforderlich, um den Anteil fehlender Werte bei Datumsangaben zu minimieren.

Der pU gibt an für die geplante Studie sowohl Complete Case Analysen als auch multiple Imputation (MI) durchführen zu wollen. Den Ausführungen von Jakobsen 2017 [27] folgend, wird bei bis zu 5 % fehlenden Werten keine Ersetzung mittels MI vorgenommen. Sollten mehr als 40 % fehlende Werte vorliegen, sieht der pU eine Ersetzung mittels MI als nicht mehr sinnvoll an.

Der pU gibt an, dass wenn mehr als 40 % der Daten zu einer Variablen fehlen, diese fehlenden Werte nicht imputiert werden und diese Variable nicht als Confounder verwendet wird. Die daraus resultierenden Konsequenzen für die Auswertungen und Interpretation der Daten diskutiert der pU jedoch nicht. Dies ist vom pU zu ändern.

Bezüglich des Verfahrens zur multiplen Imputation gibt der pU an, die Confounder in sein Modell einzubinden, die zur Bildung des PS verwendet wurden. Der pU gibt an die multiple Imputation mittels Chained Equations (MICE) zu verwenden. Weiterhin gibt der pU an, für welchen Variablentyp er welche Modellierung wählt. Dieses Vorgehen bezüglich des Umgangs mit fehlenden Werten ist nachvollziehbar und sachgerecht.

2.2.9 Weitere Anpassungen der AbD

Fehlende Angaben

Gemäß G-BA-Beschluss zur AbD sind im SP und SAP Angaben folgenden Punkten zu machen:

- zum Umgang mit unplausiblen Daten und Ausreißern,
- Angaben zur Überprüfung, inwieweit die parallel erhobenen, sowie zeitlich nicht parallel erhobenen Daten zu Nusinersen und Onasemnogen-Abeparvovec für eine gepoolte Analyse geeignet sind,
- Angaben zur Überprüfung, inwieweit gegebenenfalls Daten zum Vergleich von Risdiplom gegenüber Nusinersen und Onasemnogen-Abeparvovec aus unterschiedlichen Datenquellen für eine gepoolte Analyse geeignet sind

Zu diesen Punkten finden sich keine Angaben in den vom pU vorgelegten Dokumenten. Dies muss ergänzt werden.

Estimand

Der pU bestimmt in Abhängigkeit vom primären Endpunkt zwei unterschiedliche Estimands. Zum einen den Estimand, welcher auf einer Treatment Policy Strategy beruht. Zum anderen den Estimand unter Verwendung einer Hypothetical Strategy, um bei präsymptomatischen Patientinnen und Patienten einen Behandlungswechsel zu adressieren. Dazu werden in beiden Behandlungsarmen die Patientinnen und Patienten zum Zeitpunkt des Behandlungswechsels zensiert. Der pU begründet dies damit, dass der von ihm gewählte primäre Endpunkt, ungeplante Hospitalisierungen, zu den Nebenwirkungen zählt. Im Rahmen einer AbD sind jedoch Estimands nach dem Hypothetical-Strategy-Ansatz auch für Endpunkte aus der Kategorie Nebenwirkungen nicht angemessen.

Der primäre Estimand einer AbD entspricht der Treatment Policy Strategy und beinhaltet die Auswertung nach dem Intention-to-Treat(ITT)-Prinzip für alle patientenrelevanten Endpunkte. Dies muss vom pU konsequent in SP und SAP berücksichtigt werden (siehe auch Abschnitt 2.2.3).

Abbruchkriterien

Gemäß G-BA-Beschluss zur AbD sind im SP und SAP Angaben zu Abbruchkriterien wegen Vergeblichkeit darzulegen [1]. Im SAP und SP finden sich Angaben zum Abbruchkriterium,

nämlich wenn die erwartete Power für die Schätzung des primären Endpunkts (und relevante sekundäre Endpunkte) unter 60 % liegt. Dieses Kriterium ist gemäß der Ausführungen in der Niederschrift im Beratungsgespräch sachgerecht [6]. Jedoch ist in Anbetracht der Kritikpunkte zur Fallzahlplanung (siehe Abschnitt unter 2.2.4.1) das Abbruchkriterium zu hinterfragen. Darüber hinaus bleibt unklar, wie der pU für die Populationen verfahren möchte, für die er zum aktuellen Zeitpunkt noch keine Fallzahlplanung vorgenommen hat. Weiterhin ist darauf zu achten und entsprechend im SP zu ergänzen, dass eine etwaige Entscheidung für einen Abbruch der AbD sowie für eine Änderung der Fallzahlschätzung in Abstimmung mit dem G-BA erfolgen sollte.

2.2.10 Interpretation der Ergebnisse der AbD

Aus einer nicht randomisierten Studie kann aufgrund potenziell unbekannter Confounder aus den in der Studie beobachteten Effekten erst ab einer bestimmten Effektstärke eine Aussage zum Nutzen oder Schaden einer Intervention abgeleitet werden. Eine (positive oder negative) Aussage zum Nutzen oder Schaden ergibt sich dann, wenn das Konfidenzintervall für den beobachteten Effekt ober- bzw. unterhalb einer zu definierenden Schwelle liegt (Test auf verschobene Nullhypothese). Die konkrete Schwelle ergibt sich durch die Qualität der Daten im Einzelfall, u. a. durch das Wissen über wichtige Confounder [26].

2.2.11 Zeitplan für die AbD

Im Beschluss des G-BA wird ausgeführt, dass der pU Daten zu 2 Zwischenanalysen vorlegen soll: die erste Zwischenanalyse soll 18 Monate nach Beginn der AbD (bezogen auf das Datum des Feststellungsbeschlusses) durchgeführt werden und die zweite 30 Monate nach Beginn der AbD. Das vom pU geplante Vorgehen weicht dahin gehend ab, dass er plant die Analysen in Relation zum Studienstart durchzuführen. Darüber hinaus sollten die Datenschnitte für die Zwischenanalysen jeweils 4 Monate vor dem jeweiligen Berichtszeitpunkt erfolgen und nicht – wie aktuell vom pU geplant – 6 Monate.

Angaben zum Verlauf der Datenerhebung sollen 6, 18 und 30 Monate nach Beginn der AbD vorgelegt werden. Dies plant der pU für den Zeitpunkt 6 Monate anhand eines Datenschnitts zu Studienbeginn, in den nur die retrospektiv eingeschlossenen Patientinnen und Patienten eingehen. Dieses Vorgehen erscheint nicht sachgerecht. Da für die Verlaufserhebung lediglich deskriptive Analysen erforderlich sind, sollte eine kürzere Latenzzeit ausreichend sein, die ermöglicht auch bereits zur ersten Verlaufserhebung Daten zu prospektiv eingeschlossenen Patientinnen und Patienten abzubilden.

2.3 Zusammenfassung und Fazit

Die Prüfung des SP (Version 1.0) und SAP (Version 1.0) für die AbD zu Risdiplom hat ergeben, dass die Anforderungen des G-BA zur Durchführung der AbD einschließlich der zugehörigen Auswertung der Daten zum Teil, aber nicht vollständig umgesetzt wurden.

Eine Anpassung ist insbesondere in folgenden Bereichen erforderlich:

- Ausführungen zu den relevanten Datenfeldern im SMARtCARE-Register, möglichst mit ergänzender Vorlage durch das Kodierhandbuch (siehe Abschnitt 2.2)
- Vereinheitlichung der Endpunkt-Operationalisierungen zur AbD zu Onasemogen-Abeparvovec (siehe Abschnitt 2.2.1.3)
- Erstellung eines Erhebungsplans, in dem alle geplanten Erhebungen zu den jeweiligen Zeitpunkten dargestellt werden (siehe Abschnitt 2.2.1.3)
- Anpassung der Bestimmung des Indexdatums und Anpassung der Beobachtungsdauer (siehe Abschnitt 2.2.3)
- Anpassung an geforderte Änderungen der Fallzahlplanungen für die Fragestellungen (siehe Abschnitt 2.2.4.1)
- Adjustierung für Confounder: Systematische Prüfung, ob weitere Confounder für die Patientinnen und Patienten mit SMA Typ 3 berücksichtigt werden müssen (siehe Abschnitt 2.2.5)
- Beschreibung von Heterogenitätsanalysen hinsichtlich zeitlich parallel und nicht-parallel erhobener Daten (siehe Abschnitt 2.2.6)
- Konkretisierung der Auswertung stetiger Daten mittels MMRM (Definition von Cohen's d und Teststatistik) (siehe Abschnitt 2.2.6)
- Spezifizierung der Angaben zu geplanten Analysen (z. B. zur Durchführung der Subgruppenanalysen) (siehe Abschnitt 2.2.7)
- Konsequente Umsetzung der Treatment Policy Strategy mit ITT-Analyse als Estimand (siehe Abschnitt 2.2.9)
- Ergänzung der fehlenden Angaben gemäß G-BA Beschluss (siehe Abschnitt 2.2.9)

3 Literatur

1. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V; Risdiplam (spinale Muskelatrophie); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen [online]. 2022 [Zugriff: 12.09.2023]. URL: https://www.g-ba.de/downloads/39-261-5536/2022-07-21_AM-RL-XII_Risdiplam_AbD-004_Forderung_BAnz.pdf.
2. Gemeinsamer Bundesausschuss. Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V; Risdiplam (spinale Muskelatrophie); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen [online]. 2022 [Zugriff: 12.09.2023]. URL: https://www.g-ba.de/downloads/40-268-8667/2022-07-21_AM-RL-XII_Risdiplam_AbD-004_Forderung_TrG.pdf.
3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzept für eine anwendungsbegleitende Datenerhebung – Risdiplam; Rapid Report [online]. 2022 [Zugriff: 11.07.2023]. URL: https://www.iqwig.de/download/a21-131_anwendungsbegleitende-datenerhebung-risdiplam_rapid-report_v1-0.pdf.
4. Roche. Evaluation of a Real World Data Collection for the Reassessment of the Additional Benefit of Evrysdi (Risdiplam); NIS Protocol (Secondary Data Use); Version 1.0 [unveröffentlicht]. 2023.
5. F. Hoffmann-La Roch. Evaluation of a Real World Data Collection for the Reassessment of the Additional Benefit of Evrysdi (Risdiplam); Statistical Analysis Plan Version 1.0 [unveröffentlicht]. 2023.
6. Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV; Beratungsanforderung 2022-B-215; Wirkstoff: Risdiplam [unveröffentlicht]. 2022.
7. Novartis Gene Therapies EU. Routine data collection and evaluations of onasemnogene abeparvovec in Germany; Study Protocol [online]. 2022 [Zugriff: 12.09.2023]. URL: https://www.g-ba.de/downloads/40-268-8964/2022-10-20_AM-RL-AbD_Onasemnogen-Abeparvovec_2020-AbD-001_Feststellung_Studienprotokoll.pdf.
8. Finkel R, Bertini E, Muntoni F et al. 209th ENMC International Workshop: Outcome Measures and Clinical Trial Readiness in Spinal Muscular Atrophy 7-9 November 2014, Heemskerk, The Netherlands. *Neuromuscul Disord* 2015; 25(7): 593-602. <https://dx.doi.org/10.1016/j.nmd.2015.04.009>.

9. Glanzman AM, O'Hagen JM, McDermott MP et al. Validation of the Expanded Hammersmith Functional Motor Scale in spinal muscular atrophy type II and III. *J Child Neurol* 2011; 26(12): 1499-1507. <https://dx.doi.org/10.1177/0883073811420294>.
10. Mazzone E, De Sanctis R, Fanelli L et al. Hammersmith Functional Motor Scale and Motor Function Measure-20 in non ambulant SMA patients. *Neuromuscul Disord* 2014; 24(4): 347-352. <https://dx.doi.org/10.1016/j.nmd.2014.01.003>.
11. O'Hagen JM, Glanzman AM, McDermott MP et al. An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. *Neuromuscul Disord* 2007; 17(9-10): 693-697. <https://dx.doi.org/10.1016/j.nmd.2007.05.009>.
12. Pierzchlewicz K, Kepa I, Podogrodzki J et al. Spinal Muscular Atrophy: The Use of Functional Motor Scales in the Era of Disease-Modifying Treatment. *Child Neurol Open* 2021; 8: 1-9. <https://dx.doi.org/10.1177/2329048X211008725>.
13. Mazzone E, Bianco F, Martinelli D et al. Assessing upper limb function in nonambulant SMA patients: development of a new module. *Neuromuscul Disord* 2011; 21(6): 406-412. <https://dx.doi.org/10.1016/j.nmd.2011.02.014>.
14. Mazzone ES, Mayhew A, Montes J et al. Revised upper limb module for spinal muscular atrophy: Development of a new module. *Muscle Nerve* 2017; 55(6): 869-874. <https://dx.doi.org/10.1002/mus.25430>.
15. Bayley N. Bayley Scales of Infant and Toddler Development; Administration Manual. San Antonio: PsychCorp, Harcourt Assessment; 2006.
16. Kölbel H, Müller-Felber W. S1 Leitlinie: Spinale Muskelatrophie (SMA), Diagnostik und Therapie [online]. 2020 [Zugriff: 20.09.2023]. URL: https://register.awmf.org/assets/guidelines/022-030I_S1_Spinale-Muskelatropie-SMA-Diagnostik-Therapie_2021-07_1.pdf.
17. Berger ML, Mamdani M, Atkins D et al. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part I. *Value Health* 2009; 12(8): 1044-1052. <https://dx.doi.org/10.1111/j.1524-4733.2009.00600.x>.
18. Cox E, Martin BC, Van Staa T et al. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report—Part II. *Value Health* 2009; 12(8): 1053-1061. <https://dx.doi.org/10.1111/j.1524-4733.2009.00601.x>.

19. Pufulete M, Mahadevan K, Johnson TW et al. Confounders and co-interventions identified in non-randomized studies of interventions. *J Clin Epidemiol* 2022; 148: 115-123. <https://dx.doi.org/10.1016/j.jclinepi.2022.03.018>.
20. Novartis Gene Therapies EU. Routine data collection and evaluations of onasemnogene abeparvovec in Germany; Study Protocol [online]. 2021 [Zugriff: 20.09.2023]. URL: https://www.g-ba.de/downloads/17-98-5254/2022-01-20_AM-RL_Onasemnogen-Abeparvovec_abD_Studienprotokoll.pdf.
21. Roche Pharma. Risdiplam (Evrysdi); Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4A, 5q-assoziierte spinale Muskelatrophie (SMA) bei Patienten ab einem Alter von zwei Monaten, mit einer klinisch diagnostizierten SMA Typ I, Typ II oder Typ III oder mit einer bis vier Kopien des SMN2-Gens [online]. 2021 [Zugriff: 20.09.2023]. URL: <https://www.g-ba.de/bewertungsverfahren/nutzenbewertung/680/>.
22. Baranello G, Darras BT, Day JW et al. Risdiplam in Type 1 Spinal Muscular Atrophy. *N Engl J Med* 2021; 384(10): 915-923. <https://dx.doi.org/10.1056/NEJMoa2009965>.
23. Baranello G, Gorni K, Daigl M et al. Prognostic Factors and Treatment-Effect Modifiers in Spinal Muscular Atrophy. *Clin Pharmacol Ther* 2021; 110(6): 1435-1454. <https://dx.doi.org/10.1002/cpt.2247>.
24. Roche. Anlage I zum 5. Kapitel – Anforderungsformular für eine Beratung; Wirkstoff Risdiplam [unveröffentlicht]. 2022. Risdiplam (Evrysdi®).
25. Stuart EA, Lee BK, Leacy FP. Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *J Clin Epidemiol* 2013; 66(8 (Suppl)): S84-S90.e1. <https://dx.doi.org/10.1016/j.jclinepi.2013.01.013>.
26. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzepte zur Generierung versorgungsnaher Daten und deren Auswertung zum Zwecke der Nutzenbewertung von Arzneimitteln nach § 35a SGB V; Rapid Report [online]. 2020 [Zugriff: 26.07.2022]. URL: https://www.iqwig.de/download/a19-43_versorgungsnaher-daten-zum-zwecke-der-nutzenbewertung_rapid-report_v1-1.pdf.
27. Jakobsen JC, Gluud C, Wetterslev J et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. *BMC Med Res Methodol* 2017; 17(1): 162. <https://dx.doi.org/10.1186/s12874-017-0442-1>.

28. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.

Anwendungsbegleitende Datenerhebung zu Onasemnogen-Abeparvovec: Prüfung des Studienprotokolls und des statistischen Analyseplans; Addendum zum Auftrag A20-61 [online]. 2022 [Zugriff: 11.07.2023]. URL: https://www.iqwig.de/download/a21-107_anwendungsbegleitende-datenerhebung-onasemnogen-abeparvovec-pruefung-protokoll-und-sap_addendum-zum-auftrag-a20-61_v1-0.pdf.

Anhang A Prüfung der Informationsbeschaffung zur Identifizierung von Confoundern

Informationsbeschaffung

Zur Identifizierung von relevanten Confoundern verweist der pU auf folgende Quellen:

Tabelle 2: Quellen der Informationsbeschaffung zur Identifizierung von Confoundern

Dokument [Zitat]	Datum / ggf. letztes Suchdatum	Inhalt / Anforderung Informationsbeschaffung
Studienprotokoll Onasemnogen-Abeparvovec [20]	Version 2.02 vom 18.11.2021 / letzte Suche am 23.03.2021	Identifizierung von Confoundern durch systematische Recherche und Sichtung von Leitlinien und systematischen Übersichten (bibliografische Recherche in MEDLINE und der Cochrane Database of Systematic Reviews sowie eine Handsuche auf ausgewählten Webseiten nach Leitlinien)
Dossier zur Nutzenbewertung Risdiplam [21] Publikation Baranello 2021 NEJM [22]	22.04.2021 / Publikationen bis April 2019 berücksichtigt	Orientierende Literaturrecherche in MEDLINE und Embase nach prognostischen Faktoren und Effektmodifikatoren für SMA Typ I sowie eine Beratung durch interne und externe medizinische Expertinnen und Experten als Basis für die Auswahl der Faktoren im Rahmen des MAIC
Beratungsantrag von Roche [24] und Niederschrift zum Beratungsgespräch / Beratungsanforderung 2022-B-215 von G-BA [6] zu Risdiplam Publikation Baranello 2021 Clinical Pharmacology & Therapeutics [23]	08.11.2022 (Niederschrift)	Grundsätzliche positive Bewertung des vom pU vorgeschlagenen Vorgehens durch den G-BA / In der AbD zu Risdiplam sollen auch Daten für Patientinnen und Patienten mit Typ 3 SMA erhoben werden.

Kommentar zur Informationsbeschaffung

Die Informationsbeschaffung zur Identifikation von Confoundern wurde bereits im Verfahren zur AbD von Onasemnogen-Abeparvovec als geeignet beurteilt [28]. Die Übernahme der Ergebnisse aus den o.g. Quellen für eine AbD zu Risdiplam wurde bei der Niederschrift zum Beratungsgespräch [6] als grundsätzlich geeignet angesehen.

Darüber hinaus ergeben sich aber für das aktuelle Verfahren folgenden Hinweise:

- Die systematischen Recherchen, auf die die Confounderidentifikation beruht, wurden am 23.03.2021 durchgeführt. Aktuelle Publikationen zu Confoundern, die nach diesem Zeitpunkt veröffentlicht wurden, und ihre Ergebnisse sind somit nicht berücksichtigt.

- Bei der Prüfung spezifischer Fragestellungen der Beratungsanforderung zu Risdiplom erfolgte der Hinweis des G-BAs, dass zu prüfen sei, inwieweit die Analyse von Daten zu Typ 3 SMA weitere Confounder berücksichtigen sollte.

Konsequenz

Die grundsätzliche Vorgehensweise bei der Informationsbeschaffung des pU nach Confoundern ist nachvollziehbar und adäquat. Der pU sollte jedoch systematisch prüfen, ob für die Analyse von Daten zu Typ 3 SMA weitere Confounder berücksichtigt werden müssen.

Anhang B Gegenüberstellung der Operationalisierungen zwischen der AbD zu Risdiplom und Onasemnogen-Abeparvovec

Tabelle 3: Unterschiede in den geplanten Endpunkten und Operationalisierungen zu Risdiplom im Vergleich zu Onasemnogen-Abeparvovec (mehreseitige Tabelle)

Endpunkt	Unterschiede ^a in der Operationalisierung (aus Studienprotokoll / Statistischer Analyseplan)		
	Risdiplom Operationalisierung	Population	Onasemnogen-Abeparvovec Operationalisierung
Gesamtmortalität oder dauerhafte Beatmung^b	<ul style="list-style-type: none"> ▪ Zeit bis zum Tod oder permanenten Beatmung definiert als Beatmung >16 Stunden/Tag für > 21 aufeinanderfolgende Tage 	<ul style="list-style-type: none"> ▪ alle 	<ul style="list-style-type: none"> ▪ Zeit bis zum Tod oder permanenten Beatmung definiert als 2 aufeinanderfolgende Dokumentationen einer Beatmung für ≥ 16 Stunden/Tag
Beatmung			
Notwendigkeit der Beatmung	<ul style="list-style-type: none"> ▪ Zeit bis zur ersten Dokumentation einer Beatmung 	<ul style="list-style-type: none"> ▪ alle 	<ul style="list-style-type: none"> ▪ Zeit bis zur ersten von 2 aufeinanderfolgenden Dokumentationen einer Beatmung
Art der Beatmung	-		<ul style="list-style-type: none"> ▪ Zeit bis zur ersten von 2 aufeinanderfolgenden Dokumentationen einer nicht invasiven oder invasiven Beatmung
Verbesserung der Zeit der Beatmung	-		<ul style="list-style-type: none"> ▪ Zeit bis zur ersten von 2 aufeinanderfolgenden Verbesserungen der Zeit der Beatmung im Vergleich zu Baseline
Erreichen motorischer Meilensteine			
Sitzen ohne Unterstützung	<ul style="list-style-type: none"> ▪ Anteil der Patienten, die den WHO Meilenstein altersentsprechend erreichen 	<ul style="list-style-type: none"> ▪ Präsymptomatisch und SMA Typ 1 	<ul style="list-style-type: none"> ▪ Zeit bis zum Erreichen des Meilensteins
Stehen ohne Unterstützung	-		<ul style="list-style-type: none"> ▪ Zeit bis zum Erreichen des Meilensteins
Laufen ohne Unterstützung	<ul style="list-style-type: none"> ▪ Anteil der Patienten, die den Meilenstein altersentsprechend erreichen ▪ Anteil der Patienten, die den Meilenstein erreichen 	<ul style="list-style-type: none"> ▪ Präsymptomatisch, SMA Typ 1 und Typ 2 	<ul style="list-style-type: none"> ▪ Zeit bis zum Erreichen des Meilensteins

Tabelle 3: Unterschiede in den geplanten Endpunkten und Operationalisierungen zu Risdiplom im Vergleich zu Onasemnogen-Abeparvovec (mehreseitige Tabelle)

Endpunkt	Unterschiede ^a in der Operationalisierung (aus Studienprotokoll / Statistischer Analyseplan)		
	Risdiplom Operationalisierung	Population	Onasemnogen-Abeparvovec Operationalisierung
Nachhaltigkeit des Erreichens der motorischen Meilensteine	-		<ul style="list-style-type: none"> ▪ Zeit vom Erreichen der Meilensteine (Sitzen, Stehen, Laufen) bis zum dauerhaften Verlust der Meilenstein-Fähigkeit
CHOP-INTEND	<ul style="list-style-type: none"> ▪ Anteil der Patienten mit CHOP-INTEND Score ≥ 40 im Alter von 12 Monaten ▪ Veränderung im Vergleich zu Baseline im CHOP-INTEND Score im Alter von 12, 24 und 36 Monaten 	<ul style="list-style-type: none"> ▪ SMA Typ 1 ▪ SMA Typ 1 	<ul style="list-style-type: none"> ▪ Veränderung im Vergleich zu Baseline im CHOP-INTEND Score zu 6 und 12 Monaten nach Behandlungsbeginn
HINE	-		<ul style="list-style-type: none"> ▪ Veränderung im Vergleich zu Baseline im HINE Score zu 12 und 24 Monaten nach Behandlungsbeginn
HFMSE	<ul style="list-style-type: none"> ▪ Veränderung im Vergleich zu Baseline im HFMSE Gesamtscore zu 12, 24 und 36 Monaten nach Behandlungsbeginn 	<ul style="list-style-type: none"> ▪ SMA Typ 2 und Typ 3 	-
RULM	<ul style="list-style-type: none"> ▪ Veränderung im Vergleich zu Baseline im RULM Gesamtscore zu 12, 24 und 36 Monaten nach Behandlungsbeginn 	<ul style="list-style-type: none"> ▪ SMA Typ 2 und Typ 3 	-
Gehfähigkeit	<ul style="list-style-type: none"> zu 12, 24 und 36 Monaten nach Behandlungsbeginn: ▪ Veränderung im Vergleich zu Baseline in der Gehstrecke ▪ Anteil an Patienten mit einer Veränderung der Gehstrecke > 30 m (Verbesserung und Verschlechterung) ▪ relative Veränderung im Vergleich zu Baseline in der Gehstrecke 	<ul style="list-style-type: none"> ▪ SMA Typ 3 	-
Schluckfähigkeit	<ul style="list-style-type: none"> ▪ Anteil der Patienten mit Verschlechterung der Schluckfähigkeit zu 12, 24 und 36 Monaten nach Behandlungsbeginn bzw. im Alter von 12, 24 und 36 Monaten^c 	<ul style="list-style-type: none"> ▪ alle 	<ul style="list-style-type: none"> ▪ Zeit bis zur ersten Dokumentation von Schwierigkeiten beim Schlucken

Tabelle 3: Unterschiede in den geplanten Endpunkten und Operationalisierungen zu Risdiplom im Vergleich zu Onasemnogen-Abeparvovec (mehrsseitige Tabelle)

Endpunkt	Unterschiede ^a in der Operationalisierung (aus Studienprotokoll / Statistischer Analyseplan)		
	Risdiplom Operationalisierung	Population	Onasemnogen-Abeparvovec Operationalisierung
Kaufähigkeit	-		▪ Zeit bis zur ersten Dokumentation von Schwierigkeiten beim Kauen
Notwendigkeit nicht-oraler Ernährungsunterstützung	▪ Anteil der Patienten mit Notwendigkeit nicht-oraler Ernährungsunterstützung zu 12, 24 und 36 Monaten nach Behandlungsbeginn bzw. im Alter von 12, 24 und 36 Monaten ^c	▪ alle	Zeit bis zur ersten nicht-oralen Ernährungsunterstützung (zu 2 aufeinanderfolgenden Dokumentationen)
Bayley III-Skala			
Subskala „grobmotorische Fähigkeiten“	▪ Veränderung im Vergleich zu Baseline im Alter von 12, 24 und 36 Monaten	▪ Präsymptomatisch, SMA Typ1	-
Subskala expressive Sprache und rezeptive Sprache	▪ Anteil der Patienten mit altersgerechten Scores im Alter von 12, 24 und 36 Monaten	▪ Präsymptomatisch, SMA Typ1	-
Skoliose oder orthopädische Operation ^b	▪ Zeit bis zur ersten Dokumentation einer schweren Skoliose ($\geq 40^\circ$ Cobb Winkel) oder orthopädischen Operation	▪ alle	▪ Zeit bis zur ersten Dokumentation einer Skoliose oder orthopädischen Operation
Schmerz	▪ Anteil der Patienten mit regelmäßigen Schmerzen zu 12, 24 und 36 Monaten nach Behandlungsbeginn bzw. im Alter von 12, 24 und 36 Monaten ^c	▪ alle	-
Fatigue	▪ Anteil der Patienten mit Fatigue zu 12, 24 und 36 Monaten nach Behandlungsbeginn bzw. im Alter von 12, 24 und 36 Monaten ^c	▪ alle	-
Geplante Hospitalisierungen	▪ Anzahl an geplanten Hospitalisierungen (ausschließlich Hospitalisierungen zur SMA-Behandlung)	▪ alle	▪ kumulative Anzahl an geplanten Hospitalisierungen (pro Patientenjahr) (einschließlich Hospitalisierungen zur SMA-Behandlung)
Ungeplante Hospitalisierungen	▪ Anzahl der Ereignisse	▪ alle	-
SUEs	▪ Zeit bis zum Ereignis	▪ alle	▪ Anzahl an Personen mit Ereignis sowie kumulative Anzahl an SUEs (pro Patientenjahr)

Tabelle 3: Unterschiede in den geplanten Endpunkten und Operationalisierungen zu Risdiplom im Vergleich zu Onasemnogen-Abeparvovec (mehrseitige Tabelle)

Endpunkt	Unterschiede ^a in der Operationalisierung (aus Studienprotokoll / Statistischer Analyseplan)		
	Risdiplom Operationalisierung	Population	Onasemnogen-Abeparvovec Operationalisierung
UEs, die zur Hospitalisierung führen	▪ Zeit bis zum Ereignis	▪ alle	▪ Anzahl an Personen mit Ereignis sowie kumulative Anzahl an UEs, die zur Hospitalisierung führen (pro Patientenjahr)
Spezifische SUEs	Zeit bis zum Ereignis ▪ retinopathy, ▪ effect on epithelial tissue, ▪ thrombocytopenia, ▪ nephropathy, ▪ hydrocephalus, ▪ hepatopathy, ▪ cardiac events, ▪ sensory neuropathy	▪ alle	Anzahl an Personen mit Ereignis sowie kumulative Anzahl an Ereignissen (pro Patientenjahr) ▪ hydrocephalus ▪ hepatotoxicity ▪ thrombocytopenia ▪ cardiac events ▪ dorsal root ganglia cell inflammation ▪ renal toxicity ▪ respiratory tract infection ▪ epileptic seizure ▪ post lumbar puncture syndrome

a. Unterschiede sind fett markiert.
b. Neben dem kombinierten Endpunkt sind auch jeweils die Einzelkomponenten als separate Endpunkte gelistet.
c. Für präsymptomatische Patientinnen und Patienten sowie Patientinnen und Patienten mit SMA-Typ1 sind die altersabhängigen Analysen geplant, für Patientinnen und Patienten mit SMA-Typ 2 oder SMA-Typ 3 sind die Analysen abhängig vom Behandlungsstart geplant.

CHOP-INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; HFMSE: Hammersmith Functional Motor Scale Expanded; HINE: Hammersmith Infant Neurological Examination; RULM: Revised Upper Limb Module; SAP: statistischer Analyseplan; SUE: schwerwiegendes unerwünschtes Ereignis; UE: unerwünschtes Ereignis