



IQWiG-Berichte – Nr. 1265

**Anwendungsbegleitende
Datenerhebung zu
Onasemnogen-Abeparvovec:
Prüfung des Studienprotokolls
und des statistischen
Analyseplans**

2. Addendum zum Auftrag A20-61

Addendum

Auftrag: A21-149
Version: 1.0
Stand: 21.12.2021

Impressum

Herausgeber

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema

Anwendungsbegleitende Datenerhebung zu Onasemnogen-Abeparvovec: Prüfung des Studienprotokolls und des statistischen Analyseplans – 2. Addendum zum Auftrag A20-61

Auftraggeber

Gemeinsamer Bundesausschuss

Datum des Auftrags

24.11.2021

Interne Auftragsnummer

A21-149

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

An dem Addendum beteiligte Mitarbeiterinnen und Mitarbeiter des IQWiG

- Beate Wieseler
- Ralf Bender
- Thomas Kaiser
- Guido Skipka

Schlagwörter

Onasemnogen Abeparvovec, Muskelatrophie – spinale, Register, Nutzenbewertung, Forschungsdesign, Peer Review – Forschung

Keywords

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

Inhaltsverzeichnis

	Seite
Abkürzungsverzeichnis.....	iv
1 Hintergrund.....	1
2 Prüfung der Unterlagen zur Planung der AbD von Onasemnogen-Abeparvovec	2
2.1 Allgemeine Anmerkungen zu den vom pU vorgelegten Unterlagen	2
2.2 Anmerkungen zur Umsetzung des vom G-BA formulierten Anpassungsbedarfs	2
2.2.1 Fragestellung gemäß PICO: Patientenpopulation	2
2.2.2 Fragestellung gemäß PICO: Outcome (Morbidität).....	3
2.2.3 Fragestellung gemäß PICO: Outcome (Nebenwirkungen).....	4
2.2.4 Studiendesign: prospektive / retrospektive Datenerhebung	5
2.2.5 Studiendesign: Auswahl von Confoundern	6
2.2.6 Datenquelle 1: Einbindung des RESTORE-Registers.....	7
2.2.7 Datenquelle 2: SMARtCARE-Zentren außerhalb Deutschlands.....	9
2.2.8 Datenquelle 3: Einschluss von Studienzentren, die die Qualitätssicherungsrichtlinie des G-BA nicht erfüllen.....	9
2.2.9 Auswertung der Datenerhebung; Fallzahlplanung	10
2.2.10 Auswertung der Datenerhebung: Confounderadjustierung 1	12
2.2.11 Auswertung der Datenerhebung: Confounderadjustierung 2	12
2.2.12 Auswertung der Datenerhebung: Confounderadjustierung 3	14
2.2.13 Auswertung der Datenerhebung: Analyse der Endpunkte 1	14
2.2.14 Auswertung der Datenerhebung: Analyse der Endpunkte 2	15
2.2.15 Auswertung der Datenerhebung: Analyse der Endpunkte 3	15
2.2.16 Auswertung der Datenerhebung: Berücksichtigung verschobener Hypothesengrenzen	16
2.2.17 Auswertung der Datenerhebung: Subgruppenanalysen.....	17
2.2.18 Auswertung der Datenerhebung: Umgang mit fehlenden Confoundern.....	17
2.2.19 Auswertung der Datenerhebung: Umgang mit Behandlungswechseln 1	18
2.2.20 Auswertung der Datenerhebung: Umgang mit Behandlungswechseln 2	18
2.2.21 Auswertung der Datenerhebung: Umgang mit Behandlungswechseln 3	19
2.2.22 Auswertung der Datenerhebung: Geplante Analysen	20
3 Literatur	23

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AbD	Anwendungsbegleitende Datenerhebung
CHOP-INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
EFS	Event-free-survival
G-BA	Gemeinsamer Bundesausschuss
HINE	Hammersmith Infant Neurological Examination
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
MedDRA	Medical Dictionary for Regulatory Affairs
PICO	Patient, Intervention, Comparator, Outcome
PS	Propensity Score
pU	pharmazeutischer Unternehmer
QS-RL	Qualitätssicherungs-Richtlinie
SAP	Statistischer Analyseplan
SMD	Standardisierte Mittelwertdifferenz
SGB	Sozialgesetzbuch
SMA	spinale Muskelatrophie
SMN	Survival of Motor Neuron
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 24.11.2021 mit der Prüfung des überarbeiteten Studienprotokolls und des statistischen Analyseplans (SAP) zur anwendungsbegleitenden Datenerhebung (AbD) zu Onasemnogen-Abeparvovec beauftragt.

In seiner Sitzung am 04.02.2021 hat der G-BA beschlossen, eine AbD und Auswertungen nach § 35a Absatz 3b Satz 1 SGB V für den Wirkstoff Onasemnogen-Abeparvovec in der Behandlung der spinalen Muskelatrophie zu fordern [1,2]. Dem Beschluss liegt unter anderem das vom IQWiG erstellte Konzept für eine AbD zu Onasemnogen-Abeparvovec zugrunde (Rapid Report A20-61 vom 01.10.2020 [3]).

Zur Prüfung, ob die Anforderungen des G-BA an die AbD und an Auswertungen durch die vom pharmazeutischen Unternehmer (pU) erstellten Unterlagen zum Studienprotokoll und SAP umgesetzt worden sind, hat der G-BA dem IQWiG die erste Version dieser Unterlagen übermittelt [4,5] und es mit der Prüfung dieser Unterlagen beauftragt. Neben dem G-BA-Beschluss zur Onasemnogen-Abeparvovec sollen die Inhalte der diesbezüglichen Beratungen des pU zur Studienplanung der AbD (2021-B-190 [6], 2021-B-122 [7]) berücksichtigt werden.

Auf Basis der Ergebnisse dieser Prüfung (Addendum A21-107 zum Rapid Report A20-61) hat der G-BA mit Schreiben vom 28.09.2021 den pU schriftlich darüber informiert, bei welchen Angaben im vorgelegten Studienprotokoll und SAP Anpassungsbedarf besteht. Der pU hat am 23.11.2021 überarbeitete Unterlagen beim G-BA eingereicht. Der G-BA hat dem IQWiG diese Unterlagen am 24.11.2021 übermittelt.

2 Prüfung der Unterlagen zur Planung der AbD von Onasemnogen-Abeparvovec

2.1 Allgemeine Anmerkungen zu den vom pU vorgelegten Unterlagen

Der pU hat das Studienprotokoll und den SAP (jeweils in der Version vom 18.11.2021) auf Basis der Anforderungen des G-BA, die ihm mit Schreiben vom 28.09.2021 übermittelt wurden, überarbeitet. Die Prüfung des Protokolls und SAP bezieht sich ausschließlich auf die Punkte aus der Liste des G-BA und die vom pU in seiner Übersicht der Anpassungen in Studienprotokoll und SAP beschriebenen Änderungen. Darüber hinaus wird ausschließlich die vom pU als G-BA Ansatz benannte Studienplanung kommentiert (siehe Abschnitt 2.2.1).

Die Prüfung der Änderungen des Studienprotokolls und des SAP hat zu einigen Punkten weiteren Anpassungsbedarf ergeben. Diese verbleibenden Mängel können in einer weiteren Revision der Studienunterlagen behoben werden. Die Bewertung der Umsetzung und der abschließenden Methodik der Studie können nach Einschätzung des IQWiG im Rahmen der Nutzenbewertung unter Berücksichtigung der Ergebnisse der AbD erfolgen.

2.2 Anmerkungen zur Umsetzung des vom G-BA formulierten Anpassungsbedarfs

2.2.1 Fragestellung gemäß PICO: Patientenpopulation

Position des G-BA

Die Definition der Patientenpopulation und die Auswertung der Daten sollten entsprechend der Vorgaben des G-BA getrennt für präsymptomatische und symptomatische Patientinnen und Patienten erfolgen.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Definition von zwei Analyse-Ansätzen innerhalb der Studie, wobei im „G-BA-Ansatz“ eine Stratifizierung in vier Studienpopulationen erfolgt. Einer Stratifizierung prä-symptomatischer Patienten anhand der Kopienzahl des *SMN2*-Gens wurde im Rahmen des Beratungsgesprächs am 29. Juni 2021 zugestimmt (Frage 10). Konkordante Folgeanpassungen ergeben sich bei der Fallzahlplanung sowie den erwarteten Patientenzahlen.

Aufgrund der bereits im Beratungsgespräch vom 11. August 2021 vorgetragenen Herausforderungen einer Stratifizierung von Patienten anhand des Symptomstatus in der Routineversorgung sowie maßgeblichen Auswirkungen auf die notwendige Anzahl von Patienten erfolgt im NGT-Ansatz weiterhin eine Stratifizierung allein anhand der Kopienzahl des *SMN2*-Gens.

Für eine detaillierte Darstellung des Sachverhalts wird auf Abschnitt 8.1 des Studienprotokolls verwiesen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Der pU setzt die Anforderung des G-BA zur Definition der Studienpopulationen in der Fragestellung der AbD um. In einer Populationsdefinition, die er als G-BA Ansatz bezeichnet,

berücksichtigt er präsymptomatische und symptomatische Patientinnen und Patienten getrennt. Symptomatische Patientinnen und Patienten werden gemäß Anforderung des G-BA noch einmal in solche mit Typ-1-SMA und Typ-2-SMA unterteilt. Die prä-symptomatischen Patientinnen und Patienten stratifiziert der pU anhand der Kopienzahl des *SMN2*-Gens (≤ 2 Kopien / 3 Kopien). Diese Populationsdefinition entspricht den Anforderungen des G-BA und erscheint sachgerecht.

Ausgehend von dieser Populationsdefinition plant der pU die weitere Methodik der AbD. Dabei benennt er das an diese Population gebundene Vorgehen im Studienprotokoll und im SAP jeweils als G-BA Ansatz.

Parallel zum G-BA Ansatz beschreibt der pU im Studienprotokoll und im SAP ein Vorgehen, das er als Novartis Gene Therapies (NGT) Ansatz bezeichnet. Dieser Ansatz beruht auf einer im Vergleich zur ersten Protokollversion unveränderten Definition der Patientenpopulation, die lediglich zwischen Patientinnen und Patienten mit bis zu 2 bzw. mit 3 Kopien des *SMN2*-Gens unterscheidet, den Symptomstatus aber unberücksichtigt lässt. Diese Populationsdefinition und damit auch die darauf aufbauende Studienplanung entsprechen nicht den Anforderungen des G-BA und sind damit für eine der Datenerhebung folgende Nutzenbewertung nicht geeignet. Das als NGT Ansatz bezeichnete Vorgehen wird deshalb im Folgenden nicht weiter bewertet. Alle Ausführungen in den folgenden Kapiteln beziehen sich auf die als G-BA Ansatz bezeichnete Studienplanung.

2.2.2 Fragestellung gemäß PICO: Outcome (Morbidity)

Position des G-BA

Die durch die Vielzahl der Endpunkte zur Beschreibung der motorischen Funktion entstehende Multiplizität sollte verringert werden, indem die relevanten Endpunkte selektiert und die Endpunkte insgesamt hierarchisiert werden. Diese Entscheidungen müssen im Studienprotokoll präspezifiziert werden. Primär sollten Endpunkte, die den gesamten relevanten Beobachtungszeitraum abdecken, herangezogen werden.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

G-BA-Ansatz: Reduktion der Endpunkte zur Abbildung der motorischen Funktion bei gleichzeitiger Fokussierung auf TTE-Analysen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Der pU reduziert die Anzahl der Endpunkte zur motorischen Funktion. Er plant jeweils die Erhebung der Zeit bis zum Erreichen der Meilensteine Sitzen, Stehen und Gehen ohne Unterstützung sowie die Zeit von der maximalen Ausprägung eines dieser Meilensteine bis zum vollständigen Verlust eines Meilensteins. Damit bildet er das Erreichen der Meilensteine und den Erhalt der Funktion ab. Darüber hinaus beschränkt er die Erhebung der motorischen Funktion durch Instrumente von 4 auf 2 (CHOP-INTEND und den HINE). Diese Planung stellt eine Verbesserung im Vergleich zum ursprünglich geplanten Protokoll dar. Die Interpretation

der Effekte bezüglich motorischen Funktion sollte unter Berücksichtigung der Gesamtheit der Endpunkte erfolgen.

Der pU weist im Studienprotokoll darauf hin, dass Unsicherheiten bezüglich der Dokumentation des Alters bei Erreichen eines Meilensteins der motorischen Funktion (Bericht der Eltern oder Beurteilung durch Neuropädiaterinnen oder Neuropädiater) bestehen und darüber hinaus eine weitere Verzerrung durch die unterschiedliche Häufigkeit von Visiten unter den verschiedenen Therapieoptionen entstehen könnte. Solche Verzerrungen sollten im Studienverlauf wo möglich vermieden werden. Um die Beurteilung der Relevanz dieser potenziellen Verzerrungen zu ermöglichen, müssen die angesprochenen Faktoren dokumentiert werden und Teil der Berichterstattung über die Studie sein.

2.2.3 Fragestellung gemäß PICO: Outcome (Nebenwirkungen)

Position des G-BA

Die Grenzwerte für die Erhebung der im Beschluss genannten spezifischen UE sollten vor Studienbeginn definiert und präspezifiziert sein.

Als Annäherung an die Erhebung von SUE sollte ein kombinierter Endpunkt aus UE, die zum Tod führen und UE, die zu einer Hospitalisierung führen, ausgewertet werden.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Ergänzung eines kombinierten Endpunkts aus UE mit Hospitalisierung und Tod jeglicher Ursache, da in SMARtCARE keine spezifische Erfassung von UE, die zum Tod führen, erfolgt.

Ergänzung spezifischer UE auf Grundlage der mit SMARtCARE abgestimmten und durch SMARtCARE angestoßenen, notwendigen Anpassungen des CRFs.

Für eine detaillierte Darstellung des Sachverhalts wird auf Abschnitte 5.2.2 und 5.2.3 des Studienprotokolls verwiesen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Spezifische UE

In der ersten Version des Protokolls hatte der pU den Verzicht auf die Erhebung spezifischer UE, die auf Anforderungen der EMA aus Risk-Management Plänen zu Onasemnogen-Abeprarvovec und Nusinersen zurückgehen, damit begründet, dass für die Erhebung dieser UE keine klinisch relevanten Grenzwerte definiert seien. Die Forderung nach solchen Grenzwerten lässt der pU in der überarbeiteten Version des Protokolls fallen. Stattdessen sollen diese UE durch konkrete Abfrage im CRF des SMARtCARE-Registers erhoben werden. Dieses Vorgehen erscheint in der vorliegenden Situation sachgerecht.

Eines der aus Risk-Management Plänen identifizierten spezifischen UE wurde auch vor der Protokollanpassung bereits erhoben (Hydrozephalus). Für die weiteren spezifischen UE (Hepatotoxizität, Thrombozytopenie, kardiale Ereignisse, Entzündung der

Spinalganglionzellen, renale Toxizität) erfolgt die aktive Abfrage erst durch die Protokollanpassung. Der pU weist jedoch darauf hin, dass diese UE für den Zeitraum der Datenerhebung vor der Protokollanpassung aus der unspezifischen Erhebung der UE auf Basis der MedDRA-kodierten Auswertung abgeleitet werden können. Die potenziell unterschiedlichen Häufigkeiten von UE, die auf einer generellen Abfrage basieren, und solchen, die aktiv abgefragt werden, sollten bei dieser Auswertung berücksichtigt werden.

Annäherung an die Erhebung von SUE

Die Auswertung von SUE dient der Charakterisierung schwerwiegender Nebenwirkungen der Therapie. Innerhalb des SMARtCARE-Registers wird dokumentiert, ob ein UE zur Hospitalisierung oder zur Verlängerung einer Hospitalisierung führt, die weiteren Kriterien für SUEs (UEs, die zum Tod führen, lebensbedrohliche UEs, UEs, die zu einer dauerhaften oder schwerwiegenden Behinderung führen sowie Entwicklungen von kongenitalen Anomalien oder Geburtsfehlern) werden nicht erhoben. Als Annäherung an die Erhebung von SUE hat der G-BA deshalb einen kombinierten Endpunkt aus UEs, die zur Hospitalisierung führen, und UEs, die zum Tod führen, angefordert.

Der pU beschreibt, dass er den Vorschlag des G-BA nicht umsetzen kann. Der Grund sei, dass im SMARtCARE-Register keine Daten zu UEs, die zum Tod führen, erhoben werden. Der pU plant stattdessen, als Annäherung an SUEs die Auswertung eines kombinierten Endpunkts aus UEs, die zu einer Hospitalisierung führen, und Todesfällen jeglicher Ursache. Dieses Vorgehen erscheint nicht sachgerecht.

Die vom pU vorgeschlagene Kombination von UEs, die zur Hospitalisierung führen mit Todesfällen jeglicher Art ist in der vorliegenden Indikation nicht geeignet, Nebenwirkungen der Therapie abzubilden. Da die SMA bei einem relevanten Anteil der Patientinnen und Patienten zum Tod führt, würde ein solcher kombinierter Endpunkt nicht zwischen Nebenwirkungen und Fortschreiten der Erkrankung differenzieren. Potenziell bestehende Unterschiede der Nebenwirkungen der Therapiealternativen könnten überdeckt werden.

Um die Annäherung an SUEs gemäß Vorschlag des G-BA umzusetzen, wäre es nicht notwendig, für alle UEs die potenzielle Folge Tod zu erheben. Es müsste aber für alle Todesfälle dokumentiert werden, ob diese auf UEs zurückgehen. Sollte dies nicht möglich sein, ist die Beschränkung der Auswertung auf UEs, die zur Hospitalisierung führen, sinnvoller als die Gesamtheit der Todesfälle mit einzubeziehen.

2.2.4 Studiendesign: prospektive / retrospektive Datenerhebung

Position des G-BA

Die Nutzung bereits erhobener Daten zu Nusinersen und Onasemnogen- Abeparvovec (aus dem Register SMARtCARE und ggf. weiteren Registern) sollte, sofern diese den genannten Anforderungen an die Datenqualität im Beschluss zur AbD zu Onasemnogen-Abeparvovec entsprechen, für die Registerstudie eingeplant werden

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Heranziehen nicht-paralleler Daten zu Nusinersen bei Erfüllung des präspezifizierten Kriteriums zur Überlappung der Propensity Score Verteilungen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Der pU beschreibt im Studienprotokoll die Datenerhebung im SMARTCARE-Register vor Beginn der AbD (aktuell geplant im Januar 2022). Aus dieser Beschreibung geht hervor, dass Daten zu Nusinersen seit Juli 2018 erhoben werden, Daten zu Onasemnogen-Abeparvovec seit Juli 2020 bzw. in geringem Umfang auch vorher. Damit liegen 2 Arten von retrospektiven Daten für die AbD vor. Es handelt sich zum einen um Daten zu Nusinersen und Onasemnogen-Abeparvovec, die parallel erhoben wurden (mindestens seit Juli 2020), zum anderen um Daten zu Nusinersen, die nicht parallel zur Verfügbarkeit von Onasemnogen-Abeparvovec vor dessen Zulassung im Juli 2020 erhoben wurden.

Der pU folgt der Anforderung des G-BA nur teilweise. Er plant grundsätzlich die Berücksichtigung von retrospektiven Daten (die vor dem Start der Studie im SMARTCARE-Register erhoben wurden) zu Nusinersen in der AbD ein. Der pU weist darauf hin, dass eine Berücksichtigung der retrospektiven Daten unter dem Vorbehalt der Eignung und Qualität dieser Daten steht. Das erscheint sachgerecht.

Eine Berücksichtigung von retrospektiven Daten zu Onasemnogen-Abeparvovec ist dagegen gemäß Protokoll nicht vorgesehen. Die Einschlusskriterien für Patientinnen und Patienten verlangen weiterhin, dass der Behandlungsbeginn nach Studienstart für die AbD liegen muss, und geben an, dass dieses Kriterium ausschließlich für Nusinersen aufgehoben werden kann, wenn auf historische Daten zurückgegriffen werden soll. Damit ist der Einschluss von Patientinnen und Patientinnen, die vor Beginn der AbD (geplant Januar 2022) mit Onasemnogen-Abeparvovec behandelt wurden, nach wie vor ausgeschlossen. Der pU begründet nicht, warum er die retrospektiven Daten zu Onasemnogen-Abeparvovec (parallel erhoben zu Nusinersen) nicht in der AbD berücksichtigen möchte. Die Beschränkung der Berücksichtigung retrospektiver Daten auf Nusinersen entspricht nicht den Anforderungen des G-BA und erscheint nicht sachgerecht.

Die Mängel sollten in einer weiteren Revision des Studienprotokolls und des SAP behoben werden.

2.2.5 Studiendesign: Auswahl von Confoundern**Position des G-BA**

Die Liste der Confounder sollte an die im Beschluss genannten Patientenpopulationen und an die für die Registerstudie genutzten Datenquellen (siehe nachfolgende Punkte) angepasst werden.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Zuordnung der Confounder zu den jeweils relevanten Studienpopulationen. Für eine Anpassung an weitere Datenquellen wird auf Punkt 6 verwiesen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Der pU passt die Liste der jeweils relevanten Confounder an die verschiedenen Subpopulationen der Gesamtstudienpopulation an. Dieses Vorgehen und die jeweils zugewiesenen Confounder erscheinen sachgerecht.

2.2.6 Datenquelle 1: Einbindung des RESTORE-Registers**Position des G-BA**

Der pharmazeutische Unternehmer sollte an dem selbst geführten RESTORE-Register die notwendigen Anpassungen gemäß finalem Studienprotokoll und SAP für die AbD vornehmen, um Auswertungen auf Basis des RESTORE-Registers gemeinsam mit der vorliegenden Registerstudie z. B. in Form einer Metaanalyse für die AbD nutzen zu können.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Eine Präspezifikation sämtlicher Analysen für das RESTORE- Register sowie etwaiger struktureller Anpassungen war in der vom G-BA gewährten Frist von acht Wochen nicht umsetzbar. Darüber hinaus stehen nach Auffassung von Novartis Gene Therapies die Bestimmungen des G-BA- Beschlusses vom 4. Februar 2021 hinsichtlich der Eignung internationaler Datenquellen vor dem Hintergrund der Übertragbarkeit auf den deutschen Versorgungskontext bisher einem Heranziehen von Ergebnissen des RESTORE-Registers bisher entgegen.

Novartis Gene Therapies wird während der Prüfung der angepassten Studiendokumente durch G-BA und IQWiG die Arbeit zur Präspezifikation sämtlicher Analysen für das RESTORE-Register fortsetzen. Sollte der G-BA im Rahmen eines Änderungsbeschlusses mit Blick auf Übertragbarkeitsanforderungen ein Heranziehen von Ergebnissen aus dem RESTORE-Register für eine erneute Nutzenbewertung im Jahr 2027 ermöglichen, kann Novartis Gene Therapies die entsprechend präspezifizierten Analysen im Rahmen eines Amendments im ersten Quartal 2022 dem G-BA bereitstellen. In diesem Falle würde ebenfalls eine Berücksichtigung des RESTORE-Registers im ersten Zwischenbericht an den G-BA (Einreichung am 4. August 2022) erfolgen.

Bewertung der vom pU vorgenommenen Änderungen

Der pU postuliert, dass für eine mögliche Einbindung der Ergebnisse des RESTORE-Registers vom G-BA nur eine Frist von acht Wochen gewährt worden sei. Dies ist sachlich falsch.

Bereits im Konzept des IQWiG für die AbD zu Onasemnogen-Abeparvovec wurden Voraussetzungen für die Einbindung von Ergebnissen internationaler Register beschrieben und diese Option zur Erweiterung der Fallzahl eröffnet [3]. Im Fachaustausch zur Diskussion des

Konzepts am 23.11.2020, an dem der pU teilgenommen hat, wurden in diesem Zusammenhang einige notwendige strukturelle Anpassungen für das RESTORE-Register, für das der pU Sponsor ist, erörtert [8]. Diese betrafen unter anderem Vorgaben zur Datenerhebung für die Zentren analog dem Vorgehen im SMARtCARE-Register, Maßnahmen zur Vermeidung großer Mengen fehlender Werte, insbesondere auch für mit Nusinersen behandelte Patientinnen und Patienten, sowie eine Source Data Verification. Der pU hatte in diesem Zusammenhang im Fachgespräch angegeben, „flexibel für Diskussionen bezüglich technischer Änderungen“ des RESTORE-Registers zu sein. Die Möglichkeit und die Voraussetzungen für die Einbindung von Ergebnissen des RESTORE-Registers sind dem pU daher seit mehr als 1 Jahr bekannt.

Die mögliche Einbindung internationaler Register wurden im Beschluss des G-BA vom 04.02.2021 sowie in beiden Beratungsgesprächen zur AbD am 29.06.2021 sowie am 11.08.2021 erneut vom G-BA adressiert [1,6,7]. Entsprechend wurde auch in der Prüfung der ersten Version des Studienprotokolls zur AbD die mögliche Einbindung des RESTORE-Registers beschrieben (Addendum A21-107).

Im vorliegenden Studienprotokoll beschreibt der pU primär jedoch lediglich bereits bekannte Eigenschaften des RESTORE-Registers. Daraus geht nicht hervor, dass der pU die notwendigen strukturellen und organisatorischen Änderungen des RESTORE-Registers herbeigeführt hat. Im Gegenteil findet sich z. B. zum Kriterium „Source Data Verification“ weiterhin die Angabe „No“, obwohl dieser Punkt explizit im Fachaustausch am 23.11.2020 adressiert wurde und der pU selbst dieses Kriterium, u. a. auf Basis des Beratungsgesprächs am 11.8.2021, im Zusammenhang mit dem SMARtCARE-Register als zwingend erforderlich anerkannt hat.

Der pU beschreibt schließlich, dass aus seiner Sicht ein Amendment zum Studienprotokoll und SAP zur Einbindung des RESTORE-Registers erforderlich sei. Es erscheint jedoch ausreichend, die im Beschluss des G-BA beschriebenen Anforderungen zur Einbindung weiterer Register (siehe Abschnitt 1.2.3 des G-BA-Beschlusses vom 04.02.2021) im Studienprotokoll aufzuführen und die notwendigen Bezüge zu den jeweiligen Ausführungen im Studienprotokoll zur AbD, z. B. zur Source Data Verification, herzustellen. Studienprotokoll und SAP zur Datenerhebung im SMARtCARE-Register können dann, wie bereits im Konzept A20-61 des IQWiG zur AbD von Onasemnogen-Abeparvovec vom 01.10.2020 beschrieben, Ausgangspunkt für die Einbindung weiterer internationaler Register einschließlich des RESTORE-Registers sein [3].

Ob und inwieweit der pU dann das RESTORE-Register entsprechend adaptiert und dadurch ermöglicht, dass Ergebnisse aus dem RESTORE-Register eingebunden und gemeinsam mit der finalen Analyse der AbD aus dem SMARtCARE-Register im Dossier zur Nutzenbewertung vorgelegt und sinnvoll interpretiert werden können, liegt im Verantwortungsbereich des pU.

2.2.7 Datenquelle 2: SMArtCARE-Zentren außerhalb Deutschlands

Position des G-BA

SMArtCARE-Zentren außerhalb Deutschlands sollten nicht grundsätzlich als Datenquelle ausgeschlossen werden, da diese u. a. auch prospektiv Daten für symptomatische Patienten liefern können.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Erweiterung auf österreichische Studienzentren innerhalb von SMArtCARE und Ergänzung entsprechender Subgruppenanalyse nach Region (Deutschland, Österreich).

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Die vom pU vorgenommene Erweiterung auf Studienzentren in Österreich erscheint in Verbindung mit den Zentrums-bezogenen Einschlusskriterien (siehe Abschnitt 2.2.8) sachgerecht.

2.2.8 Datenquelle 3: Einschluss von Studienzentren, die die Qualitätssicherungsrichtlinie des G-BA nicht erfüllen

Position des G-BA

Es sollte keine ausschließliche Beschränkung auf Zentren, die die Qualitätssicherungs-Richtlinie (QS-RL) des G-BA für die Anwendung von Onasemnogen-Abeparvovec erfüllen, vorgenommen werden. Vielmehr sollte die Entscheidung, ob ein Zentrum eingeschlossen wird oder nicht, von der tatsächlich in diesem Zentrum umgesetzten Qualität bzw. Versorgung abhängen.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Heranziehen von Daten aus Studienzentren unabhängig von der Verfügbarkeit beider Studieninterventionen oder der Erfüllung der QS-RL des G-BA.

Ableitung von notwendigen Patientenzahlen pro Zentrum zur Gewährleistung eines Mindestmaßes an Erfahrung und Vergleichbarkeit in der Behandlung und Dokumentation von Confoundern und Endpunkten als Approximation der vom G-BA nicht näher definierten tatsächlichen Qualität auf Grundlage von SMArtCARE-Daten.

Für eine detaillierte Darstellung des Sachverhalts wird auf Abschnitt 6.3 des Studienprotokolls verwiesen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Der pU beschreibt in Anlehnung an die QS-RL des G-BA zu Onasemnogen-Abeparvovec, dass für die AbD im Wesentlichen 2 Kriterien angewendet werden sollten, um eine ausreichende Versorgungsqualität der an der AbD beteiligten Studienzentren sicherzustellen:

1) **Ausreichende Erfahrung in der Behandlung von Menschen mit SMA:**

Zur Operationalisierung einer ausreichenden Erfahrung in der Behandlung von Menschen mit SMA sieht der pU vor, dass in dem jeweiligen Zentrum mindestens 10 Kinder und Jugendliche im Alter unter 18 Jahren, darunter mindestens 5 Kinder im Alter unter 10 Jahren, in den letzten 3 Jahren vor Beginn der AbD medikamentös behandelt wurden. Diese Fallzahl liegt zwar unter der in der QS-RL für Onasemnogen-Abeparvovec angegebenen Fallzahl (mindestens 15 Behandlungsfälle insgesamt, darunter mindestens 5 bei Kindern unter 1 Lebensjahr). Die vom pU eingesetzten Fallzahlen erscheinen jedoch als Kompromiss zwischen einer Annäherung an die Anforderungen der QS-RL-Anforderungen einerseits und einer Erweiterung der Fallzahl für die AbD andererseits sachgerecht.

2) **Ausreichende Erfahrung insbesondere der Physiotherapeutinnen und Physiotherapeuten in der Erhebung der motorischen Meilensteine anhand der im SMARtCARE-Register verwendeten Skalen:**

Eine ausreichende Erfahrung in der Erhebung der motorischen Meilensteine sieht der pU durch eine Teilnahme der Zentren am SMARtCARE-Register als ausreichend gewährleistet an, da entsprechende Vorgaben und Schulungen Bestandteil des SMARtCARE-Registers sind. Dies erscheint ebenfalls sachgerecht.

Die Definition der beiden Kriterien Behandlungserfahrung und Erfahrung in der Erhebung motorischer Meilensteine als Zentrums-bezogene Einschlusskriterien für die AbD erscheint zusammenfassend sachgerecht, ebenso wie die konkrete Operationalisierung im Studienprotokoll.

2.2.9 Auswertung der Datenerhebung; Fallzahlplanung

Position des G-BA

Die Beschreibung der Rekalkulation der Fallzahlplanung (36-Monats-Analyse) im SAP sollte deutlich detaillierter erfolgen, darüber hinaus sollte die genaue Verwendung des Maß R^2 und dessen genaue Definition ergänzt werden. Die Beschreibung der Rekalkulation sollte auf Basis einer verschobenen Hypothesengrenze für die Beurteilung der Effekte erfolgen.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Konkretisierung der verwendeten Annahmen und Methoden für die Fallzahlplanung sowie konkordante Anpassungen resultierend aus Stratifikationsanforderungen des G-BA. Klarstellung der Anwendung des beschriebenen methodischen Vorgehens auch im Rahmen der Rekalkulation von Fallzahlen zu 18 und 36 Monaten.

Ergänzung einer Fallzahlplanung mit verschobener Hypothesengrenze.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

In den überarbeiteten Versionen des Protokolls und des SAP werden Konkretisierungen der verwendeten Annahmen und Methoden für die Fallzahlplanung sowie Anpassungen resultierend aus Stratifikationsanforderungen des G-BA vorgenommen. Es wird zusätzlich auch eine Fallzahlplanung für eine verschobene Nullhypothese ergänzt. Zudem wird eine weitere Fallzahlplanung ohne verschobene Nullhypothese vorgenommen mit dem Ziel, die Situation eines dramatischen Effekts abzubilden.

Die Fallzahlplanung ohne verschobene Nullhypothese für die Situation eines dramatischen Effekts erscheint nicht sachgerecht, da diese Berechnung nur 1 Kriterium für einen dramatischen Effekt beinhaltet, nämlich das reduzierte Signifikanzniveau von $\alpha=0,01$. Ein dramatischer Effekt ist aber weiterhin dadurch charakterisiert, dass orientierend ein beobachteter Effekt von $RR=10$ vorliegen muss. Dieses Kriterium bleibt in der Fallzahlplanung unberücksichtigt.

Die vorgelegte Fallzahlplanung für eine verschobene Nullhypothese wird vereinfachend mit der Farrington-Manning-Formel für die Datensituation einer Vierfeldertafel durchgeführt, obwohl ein Endpunkt mit zensierten Daten betrachtet wird. Da im vorliegenden Fall die Fallzahlplanung nur orientierenden Charakter hat, erscheint die Anwendung dieser Methode durchaus sachgerecht. Allerdings wird die Berechnung nur zu einem Zeitpunkt durchgeführt. Es ist unklar, ob diese Fallzahlplanung sich auf die Studiendauer von 36 Monaten bezieht.

Zudem erscheint die Fallzahlkalkulation für den Endpunkt „Sitting without support to month 18“ fehlerhaft (siehe Tabellen 11 und 12 im SAP bzw. 31 und 32 im Studienprotokoll). Für die verschobene Hypothese wird eine Hypothesengrenze von 2, für den tatsächlichen Effekt jedoch mit 1,86 ein Wert kleiner als 2 angelegt. Statt eine relevante Überlegenheit zur Grenze 2 statistisch zu testen, wird hier fälschlicherweise die Fallzahl für einen Test auf Nichtunterlegenheit mit Grenze 2 berechnet. Da hier Hypothesengrenze und tatsächlicher Effekt mit 2 bzw. 1,86 sehr dicht beieinanderliegen, resultiert entsprechend die sehr große Fallzahl von fast 14.000.

Unverständlich ist zudem die Kritik an der Fallzahlplanung des IQWiG, dass im Addendum A21-107 die angewendete Methode nicht angegeben werde und widersprüchliche Angaben beinhalte, da in den beiden Szenarien trotz deutlich unterschiedlicher Beobachtungsdauer die gleiche Anzahl von Ereignissen verwendet worden sei. Im Addendum A21-107 wurde aber gar keine Fallzahlplanung vorgenommen. Vermutlich ist die Fallzahlplanung im Rapid Report A20-61 [3] gemeint. Hier wird aber sehr wohl die Methode genannt (Cox-Modell bei verschobener Nullhypothese) und es wird angegeben, dass in den beiden Szenarien mit deutlich unterschiedlichen Beobachtungsdauern auch deutlich unterschiedliche Anzahlen von Ereignissen auftreten [3].

Von den vorgelegten Fallzahlplanungen erscheint lediglich diejenige für eine verschobene Nullhypothese methodisch sachgerecht, und das auch nur für den Endpunkt EFS. Für den Endpunkt „Sitting without support to month 18“ ist die Testrichtung falsch. Beim Endpunkt EFS ist jedoch unklar, auf welche Studiendauer sich die Fallzahlplanung bezieht und woher die Annahmen über die Ereigniswahrscheinlichkeiten in der Nusinersengruppe kommen. Zudem ist es nicht nachvollziehbar, warum in einer Situation mit eingeschränkter Fallzahl eine Power von 90 % gefordert wird, was den benötigten Stichprobenumfang unnötig in die Höhe treibt. Es sollte eine grobe Fallzahlabeschätzung für den Zeitraum 36 Monate mit einer Power von 80 % unter Berücksichtigung einer verschobenen Nullhypothese durchgeführt werden. Diese hat in diesem Projekt aber nur orientierenden Charakter.

2.2.10 Auswertung der Datenerhebung: Confounderadjustierung 1

Position des G-BA

Die Aufteilung der Patienten in die vorgeschlagenen „Behandlungsgruppen“ für die Confounderadjustierung sollte geändert werden. Eine Aufteilung der Patienten muss durch Informationen erfolgen, die zu Studienbeginn vorliegen.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Ergänzung des G-BA-Ansatzes unter Zuordnung sämtlicher Patienten innerhalb der betreffenden Studienpopulationen zur Erstbehandlung.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Die geforderte Zuordnung aller therapie-naiven Patienten zur Erstbehandlung wird in den überarbeiteten Versionen des Protokolls und des SAP vorgenommen. Dieses Vorgehen erscheint sachgerecht.

2.2.11 Auswertung der Datenerhebung: Confounderadjustierung 2

Position des G-BA

Fehlende Details für die Propensity-Score-Analyse sollten ergänzt werden (Überprüfung der Güte, konkrete Kriterien für eine ausreichende Überlappung und Balanciertheit).

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Explizite Erwähnung der Überprüfung der Güte, Ergänzung prä-spezifizierter Kriterien für die Überlappung und Balance sowie die Entscheidungsstruktur zur Modellauswahl.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

In der überarbeiteten Version des SAP werden fehlende Details für die Propensity-Score-Analyse (Überprüfung der Güte, konkrete Kriterien für eine ausreichende Überlappung und Balanciertheit) ergänzt. Diese Ergänzungen sind aber unvollständig, nur zum Teil sachgerecht und insgesamt widersprüchlich. Im Einzelnen liegen folgende Mängel vor:

- Das Kriterium für eine ausreichende Überlappung ist so wie formuliert unlogisch und nicht sinnvoll. Es wird angegeben, dass eine ausreichende Überlappung vorliegt, wenn in einer Behandlungsgruppe für 50% der Patientinnen und Patienten nicht gilt $PS < 0,3$ und in der anderen Behandlungsgruppe für 50% der Patientinnen und Patienten gilt $PS > 0,7$. Die Anwendung dieses Kriteriums würde bedeuten, dass zum einen Situationen mit 0 % Überlappung als ausreichend überlappend gelten (z. B. eine PS-Verteilung im 1. Arm zwischen 0,2 und 0,4, sodass die 1. Bedingung gilt und eine PS-Verteilung im 2. Arm zwischen 0,6 und 0,8, sodass die 2. Bedingung gilt), aber andererseits Situationen mit 100 % Überlappung als nicht ausreichend überlappend angesehen werden (z. B. eine PS-Verteilung in beiden Armen zwischen 0,4 und 0,6, sodass beide Bedingungen nicht gelten). Es bleibt offen, ob es sich hierbei um ein Versehen des pU handelt, zumal sich der Fehler an mehreren Stellen im SAP findet.
- Zur Beurteilung der Balanciertheit wird zunächst folgendes Kriterium für die standardisierten Mittelwertdifferenzen (SMDs) aller Confounder zwischen den Behandlungsgruppen nach Gewichtung angegeben:
 - $abs(SMD) < 0,1$: adäquate Balanciertheit
 - $abs(SMD)$ zwischen 0,1 und 0,2: keine schwerwiegende Unbalanciertheit
 - $abs(SMD) > 0,2$: schwerwiegende Unbalanciertheit

Dieses Kriterium wäre sachgerecht. Allerdings wird das Kriterium dadurch abgeschwächt, dass Ergebnisse mit $abs(SMD) > 0,1$ (ohne weitere Einschränkungen) in Abhängigkeit vom Stichprobenumfang und möglichen Konvergenzproblemen bei der PS-Analyse doch als ausreichend balanciert angesehen werden. Die oben beschriebenen Kriterien werden damit nicht angewendet. Was das konkret bedeutet bleibt unklar. Zudem sieht der Entscheidungsalgorithmus zur Modellwahl (Abb. 3 im SAP) vor, dass die Gewichtungsart gewählt wird, für die die Summe der Werte für $abs(SMD)$ aller Confounder am geringsten ausfällt. Es wird nicht angegeben, dass keine PS-Analyse durchgeführt wird, wenn für einen der Confounder eine schwerwiegende Unbalanciertheit festgestellt wird.

- Es fehlt die Angabe, dass die Zielpopulation, für die der in der PS-Analyse (nach Trimming und Gewichtung) letztlich geschätzte Behandlungseffekt gilt, genau zu beschreiben ist und dass zu begründen ist, dass diese Zielpopulation für die Ausgangsfragestellung angemessen ist.

Zusammenfassend sind somit die Details für die Propensity-Score-Analyse (Überprüfung der Güte, konkrete Kriterien für eine ausreichende Überlappung und Balanciertheit) unvollständig, nur zum Teil sachgerecht, und widersprüchlich.

Die Mängel sollten in einer weiteren Revision des Protokolls und des SAP behoben werden.

2.2.12 Auswertung der Datenerhebung: Confounderadjustierung 3

Position des G-BA

Eine Beschreibung eines Entscheidungsalgorithmus zur Anpassung der Propensity-Score-Analyse bei fehlender Überlappung und Balanciertheit nach Anwendung des ersten Verfahrens sollte ergänzt werden. Gleichfalls sollte die korrekte Konsequenz benannt werden, wenn kein Propensity-Score-Verfahren gefunden werden kann, mit dem eine ausreichende Überlappung und eine ausreichende Balanciertheit der zu vergleichenden Gruppen erreicht werden kann.

In einem solchen Fall ist der Versuch einer Effektschätzung weder mithilfe von Propensity Scores noch mithilfe von Regressionsmodellen sinnvoll.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Ergänzung und Visualisierung der Entscheidungsstruktur sowie entsprechenden Kriterien zur Modellauswahl (u.a. Überlappung und Balanciertheit).

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

In der überarbeiteten Version des SAP wird ein Entscheidungsalgorithmus zur Anpassung der Propensity-Score-Analyse bei fehlender Überlappung und Balanciertheit nach Anwendung des ersten Verfahrens ergänzt (Abb. 3 im SAP). Dieser Entscheidungsalgorithmus erscheint allerdings nicht sachgerecht. Im Einzelnen liegen folgende Mängel vor:

- Die Kriterien zur Modellauswahl (Überlappung und Balanciertheit) erscheinen wie oben dargestellt nicht sachgerecht.
- Es fehlt eine konkrete Angabe, wie das im Entscheidungsalgorithmus angegebene Trimming durchgeführt werden soll.
- Der Entscheidungsalgorithmus enthält auch einen Ansatz über Matching, bei dem es ausreichend ist, wenn nur mindestens 50 % der Confounder berücksichtigt werden. Dieser Ansatz erscheint per se nicht sachgerecht.

Der ergänzte Entscheidungsalgorithmus erscheint damit insgesamt nicht sachgerecht.

Die Mängel sollten in einer weiteren Revision des Protokolls und des SAP behoben werden.

2.2.13 Auswertung der Datenerhebung: Analyse der Endpunkte 1

Position des G-BA

Die Modelle für die Effektschätzung sollten im Detail dargestellt werden. In die Analyse sollte das Zentrum weder als zufälliger noch als fester Effekt eingehen. Ein möglicher Zentrumsseffekt sollte in einer Sensitivitätsanalyse untersucht werden.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Detaillierte Beschreibung der Modelle zur Effektschätzung pro Endpunktart und Analyse-Ansatz unter Streichung eines Zentrumeffekts. Hinzufügen einer Sensitivitätsanalyse für möglichen Zentrumeffekt.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

In der überarbeiteten Version des SAP werden die geplanten Modelle ausreichend detailliert dargestellt unter Streichung des Zentrumeffekts. Es werden Sensitivitätsanalysen für mögliche Zentrumeffekte geplant. Dieses Vorgehen erscheint sachgerecht.

2.2.14 Auswertung der Datenerhebung: Analyse der Endpunkte 2**Position des G-BA**

Im SAP sollte im Detail beschrieben werden, in welcher Form die Confounder als feste Effekte in das jeweilige Endpunkt-Modell eingehen sollen.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Konkretisierung des Eingangs von Counfoundern in die Adjustierungsmodelle.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

In der überarbeiteten Version des SAP wird beschrieben in welcher Form die Confounder in die PS-Analyse eingehen. Dieses Vorgehen erscheint sachgerecht. Es bleibt jedoch die Option ungenutzt, im Fall unzureichender Überlappung und Balanciertheit durch die Betrachtung von Wechselwirkungen möglicherweise eine ausreichende Überlappung und Balanciertheit zu erreichen.

2.2.15 Auswertung der Datenerhebung: Analyse der Endpunkte 3**Position des G-BA**

Angaben, wie überprüft werden soll, ob zeitlich parallele und nicht parallele Daten bzw. Daten aus unterschiedlichen Datenquellen für gepoolte Analysen herangezogen werden können, fehlen und sollten ergänzt werden.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Präspezifikation der Überprüfung der Eignung von nicht- parallelen Daten zu Nusinersen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Es wird in der überarbeiteten Version des SAP ergänzt, in welcher Form die Eignung von nicht-parallelen Daten zu Nusinersen überprüft wird. Im Entscheidungsalgorithmus (Abb. 3 im SAP) wird angegeben, dass bei unzureichender Überlappung für die kombinierte Stichprobe überprüft wird, ob die Verwendung nur der parallelen Daten zu Nusinersen zu einer ausreichenden Überlappung führt. Neben dem Problem des unlogischen und nicht sinnvollen Kriteriums zur Überprüfung der Überlappung (siehe Abschnitt 2.2.11) erscheint der Wechsel

von der kombinierten Stichprobe zu der Stichprobe mit ausschließlich parallelen Daten im Entscheidungsalgorithmus zu früh. Es sollten zunächst die anderen Verfahren, die zu einer verbesserten Überlappung und Balanciertheit führen können (Trimming, Gewichtungsmethode), angewendet werden. Der Wechsel von der kombinierten Stichprobe zu der Stichprobe mit ausschließlich parallelen Daten sollte erst vollzogen werden, wenn alle anderen Optionen gescheitert sind. Zudem sollten die Methoden, die bei einem gänzlichen Scheitern der PS-Analyse geplant sind (Matching, naiver Vergleich), auch zunächst mit der kombinierten Stichprobe durchgeführt werden. Es erscheint nicht sachgerecht, sofort nach Feststellung einer nicht ausreichenden Überlappung im 1. Schritt der PS-Analyse in allen weiteren Schritten nur noch die Stichprobe mit ausschließlich parallelen Daten zu verwenden.

Zudem sollten die beiden Stichproben der zeitlich parallel und nicht parallel erhobenen Daten deskriptiv verglichen werden und bei zentralen Auswertungen der kombinierten bzw. der Teilstichprobe die jeweils andere Stichprobe für Sensitivitätsanalysen verwendet werden.

Die Mängel sollten in einer weiteren Revision des Protokolls und des SAP behoben werden.

2.2.16 Auswertung der Datenerhebung: Berücksichtigung verschobener Hypothesengrenzen

Position des G-BA

Die Berücksichtigung einer verschobenen Hypothesengrenze bei der Auswertung der Daten fehlt und sollte ergänzt werden. Diese Ergänzungen könnten beispielsweise bei der (bisher fehlenden) Formulierung einer Hypothese erfolgen.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Die Bewertung der Ergebnisse dieser nicht- interventionellen Beobachtungsstudie obliegt dem G-BA im Rahmen einer erneuten Nutzenbewertung im Jahr 2027. Es erfolgt eine umfassende Berichterstattung über sämtliche in Studienprotokoll und SAP vorgesehenen Analysen zu dem im G-BA-Beschluss vom 4. Februar 2021 vorgesehenen Zeitpunkten.

Die Anwendung der vom IQWiG vorgeschlagenen Hypothesengrenzen wird von Novartis Gene Therapies dabei insb. mit Blick auf die Konsistenz zur methodischen Herleitung aus den Kriterien des „dramatischen Effekt“ sowie der praktischen Umsetzbarkeit in seltenen Erkrankungen – dem maßgeblichen Regelungsbereich des § 35a Abs. 3b SGB V – als nicht sachgerecht eingeschätzt. Insbesondere in der spezifischen Verfahrenskonstellation zu Onasemnogen-Abeparvovec, in welcher durch den G-BA bereits vor Beginn der Anwendungsbegleitenden Datenerhebung kein demonstrierter Zusatznutzen beschieden wurde, erscheint es sinnvoll, im Rahmen einer erneuten Nutzenbewertung auf Grundlage der besten verfügbaren Evidenz ausgeprägte Vorteile, die mit einer 99-prozentigen statistischen Sicherheit festgestellt werden können (vgl. Kriterien zum „dramatischen Effekt“), zu berücksichtigen.

Gemäß den Bestimmungen des G-BA-Beschlusses vom 4. Februar 2021, welcher die Anwendung einer „verschobenen Nullhypothese“ explizit auf die Fallzahlplanung der Studie

bezieht, erfolgte ergänzend eine Fallzahlplanung unter Anwendung der vom IQWiG vorgeschlagenen Hypothesengrenzen.

Für eine detaillierte Darstellung des Sachverhalts wird auf Abschnitt 8.2.2.2 des Studienprotokolls verwiesen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Die Berücksichtigung einer verschobenen Hypothesengrenze bei der Auswertung der Daten wird in den überarbeiteten Versionen des Protokolls und des SAP nicht ergänzt. Das erscheint nicht sachgerecht, und Studienprotokoll und SAP sollten entsprechend ergänzt werden. Unabhängig davon wird auf Basis der üblichen Präsentation der Ergebnisse in Form von Effektschätzungen mit Konfidenzintervall in der späteren Nutzenbewertung bei der Interpretation der Ergebnisse dieser Ansatz zur Anwendung kommen.

2.2.17 Auswertung der Datenerhebung: Subgruppenanalysen

Position des G-BA

Aufgrund der zu erwartenden geringen Fallzahlen wird vorgeschlagen, alle relevanten Subgruppenanalysen ohne die Anforderung einer statistisch signifikanten Interaktion zu rechnen und darzustellen.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Entfernung des Kriteriums einer statistisch signifikanten Interaktion für die Berichterstattung von Subgruppenergebnissen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

In den überarbeiteten Versionen des Protokolls und des SAP wird die Anforderung einer statistisch signifikanten Interaktion für die relevanten Subgruppenanalysen entfernt. Dieses Vorgehen erscheint sachgerecht.

2.2.18 Auswertung der Datenerhebung: Umgang mit fehlenden Confoundern

Position des G-BA

Für die Berücksichtigung von Daten sollten die entsprechenden Register/ Datensätze grundsätzlich Informationen zu allen relevanten Baseline-Confoundern enthalten. Ein Ausschluss von einzelnen Personen mit verbleibenden fehlenden Daten aus allen Analysen, die diese Confounder berücksichtigen, erscheint in Anbetracht geringer Fallzahlen jedoch nicht sachgerecht.

Es wird vorgeschlagen, verbleibende fehlende Werte bei einzelnen Personen durch den Ansatz der Multiplen Imputation zu ersetzen. Darüber hinaus sollten Angaben, in welchem Umfang bzw. aus welchen Gründen fehlende Daten zu erwarten sind, und Angaben zum Umgang mit unplausiblen Daten bzw. Ausreißern ergänzt werden.

Des Weiteren sollte eine Beschreibung der Anteile fehlender Daten vorgesehen werden.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Präspezifikation einer Imputation von Fehlwerten mittels multipler Imputation für Confounder bzw. Baseline- Charakteristika.

Hinsichtlich des Umgangs mit Fehlwerten bei historischen Daten wird auf die Niederschrift zum Beratungstermin am 11. August 2021 verwiesen (Frage 7). Rückfragen zur Kongruenz der Positionen hinsichtlich einer Fehlwerts substitution einerseits und Anforderungen an die Eignung historischer Daten andererseits wurden durch Novartis Gene Therapies mündlich im und schriftlich nach dem Beratungstermin vorgetragen. Mangels näherer Ausführungen des G-BA erfolgte eine bestmögliche Umsetzung der G-BA-Positionen ohne weitere Diskussion von Konsistenzaspekten.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

In der überarbeiteten Version des SAP wird ergänzt, dass beim Auftreten von fehlenden Daten eine multiple Imputation vorgenommen wird. Es wird zudem ergänzt, wie mit unplausiblen Daten umgegangen werden soll. Dieses Vorgehen erscheint sachgerecht unter der Voraussetzung, dass die Angaben zu fehlenden Werten ausreichend detailliert sind und insbesondere Umfang der fehlenden Daten, Gründe für die fehlenden Daten und Anteile fehlender Daten beschrieben werden.

2.2.19 Auswertung der Datenerhebung: Umgang mit Behandlungswechseln 1

Position des G-BA

Die Aufteilung der Patienten in die vorgeschlagenen „Behandlungsgruppen“ sollte geändert werden, da eine adäquate Aufteilung der Patienten durch Informationen erfolgen muss, die zu Studienbeginn vorliegen.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Allokation von Patienten in zwei Behandlungsarme anhand der Ersttherapie im Rahmen des G-BA-Ansatzes.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Die geforderte Zuordnung aller therapie-naiven Patienten zur Erstbehandlung wird in den überarbeiteten Versionen des Protokolls und des SAP vorgenommen. Dieses Vorgehen erscheint sachgerecht.

2.2.20 Auswertung der Datenerhebung: Umgang mit Behandlungswechseln 2

Position des G-BA

Ein Cox-Modell mit zeitabhängigen Kovariablen wird im vorliegenden Fall nicht als adäquate Methode für den Umgang mit Behandlungswechseln erachtet.

Es wird eine Zuordnung therapienaiver Patientinnen und Patienten zur jeweiligen Erstbehandlung (New-User-Design) empfohlen. Als Sensitivitätsanalyse sollten ergänzende Auswertungen mit Zensierungen bei Behandlungswechseln erfolgen, wobei der Zeitpunkt der Zensierung variiert werden sollte, um „Carry-over“-Effekte für die vorherige Behandlung zu berücksichtigen.

Sofern die Ausgangsfragestellung aufgrund eines zu hohen Anteils an Behandlungswechseln nicht mehr beantwortet werden kann, kann ggf. alternativ ein Prevalent-New-User-Design für die Auswertung genutzt werden. Ob diese Option herangezogen werden sollte, kann jeweils nach Übermittlung von Daten zum Verlauf der AbD (siehe nachfolgenden Punkt) an den G-BA entschieden und in einem Amendment zum Protokoll und SAP implementiert werden

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Abbildung einer Zuordnung von Patienten zur Erstbehandlung ohne Berücksichtigung von Therapiewechseln im Rahmen der Hauptanalyse und unter alleiniger Betrachtung von Therapiewechseln mittels Zensierung im Rahmen von Sensitivitätsanalysen im Rahmen des G-BA-Ansatzes.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Die geforderte Zuordnung aller therapie-naiven Patienten zur Erstbehandlung wird in den überarbeiteten Versionen des Protokolls und des SAP vorgenommen. Als Sensitivitätsanalyse werden ergänzende Auswertungen mit Zensierungen bei Behandlungswechseln geplant, sowie eine Variation des Zeitpunkts der Zensierung, um „Carry-over“-Effekte zu berücksichtigen. Dieses Vorgehen erscheint sachgerecht.

2.2.21 Auswertung der Datenerhebung: Umgang mit Behandlungswechseln 3

Position des G-BA

Angaben zur Anzahl von Patienten, die die Behandlung wechseln einschließlich der jeweils vorliegenden Zeiten unter den verschiedenen Behandlungen, sollten Bestandteil der regelmäßig dem G-BA vorzulegenden Angaben zum Verlauf der AbD sein.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Berichterstattung über Therapiewechsel und -zeitpunkte mit jeder Einreichung beim G-BA.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Der pU beabsichtigt gemäß Studienprotokoll, zu den unterschiedlichen Berichtszeitpunkten (siehe auch Abschnitt 2.2.22) jeweils auch Informationen zum Therapiewechsel bereitzustellen („extend of treatment switching on a study level“). Das Vorgehen erscheint sachgerecht unter der Voraussetzung, dass diese Angaben ausreichend detailliert sind und beispielsweise auch Angaben zur Beobachtungsdauer (Mittelwert, Median, Minimum, Maximum, Quartile) enthalten.

2.2.22 Auswertung der Datenerhebung: Geplante Analysen

Position des G-BA

Die geplanten Zeitpunkte für die Interimsanalysen und die finale Analyse weichen von den im Beschluss genannten Zeitpunkten ab.

Die vorzulegenden Analysen sollten in Relation zum Beschlussdatum, nicht in Relation zum Studienstart geplant und entsprechend der Angaben im Beschluss durchgeführt werden. Zu jeder Zwischenanalyse sollte entsprechend auch eine Prüfung auf Abbruch wegen Vergeblichkeit vorgenommen werden.

Anmerkung des pU (aus seiner Zusammenfassung der vorgenommenen Änderungen)

Anpassung der Analyse- und Berichtzeitpunkte an die Vorgaben des G-BA-Beschlusses vom 4. Februar 2021 mit Verweis auf die in beiden Beratungsterminen dargelegten Auswirkungen auf die Datenverfügbarkeit und daraus resultierende Belastbarkeit von Zwischenergebnissen.

Prüfung eines Abbruchs wegen Vergeblichkeit mit zweiter und dritter Zwischenanalyse. Deskriptive Diskussion von Patientenzahlen und Ergebnissen der aktualisierten Fallzahlplanungen im Rahmen der ersten Zwischenanalyse, jedoch keine formale Abbruchentscheidung aufgrund erheblicher Unsicherheit von Zwischenergebnissen auf Grundlage einer erwartbar eingeschränkten Datengrundlage.

Für eine detaillierte Darstellung des Sachverhalts wird auf Abschnitt 8.2.2.4 und 8.4 des Studienprotokolls verwiesen.

Bewertung der vom pU vorgenommenen Änderungen durch das IQWiG

Zur Beschreibung der geplanten Analysen hat der pU das Studienprotokoll und den SAP weitreichend geändert und die Abschnitte 8.4 und 8.5 im Studienprotokoll sowie das Kapitel 6 im SAP grundlegend überarbeitet. Die Änderungen umfassen neben den vom G-BA adressierten Punkten weitere Aspekte. Die im Zusammenhang mit der AbD relevanten und über die in den vorhergehenden Abschnitten des vorliegenden Dokuments hinausgehend beschriebenen Aspekte werden nachfolgend bewertet.

Analyse- und Berichtszeitpunkte

Der pU hat in der neuen Version des Studienprotokolls wie gefordert die einzelnen Berichts- und Analysezeitpunkte in Relation zum Datum des G-BA-Beschlusses (04.02.2021) und nicht mehr in Relation zum Studienstart geplant. Dies erscheint grundsätzlich sachgerecht.

Da im vorliegenden Fall das Studienprotokoll und der SAP aufgrund weitreichender Mängel der ersten Versionen erneut geprüft werden müssen, liegen zwischen möglichem Beginn der AbD auf Basis eines finalen Studienprotokolls und SAPs sowie erstem Berichtszeitpunkt nur wenige Monate (Freigabe der Dokumente nicht vor Anfang 2022, erster Berichtszeitpunkt August 2022 geplant). Der pU beschreibt, dass die erste Zwischenanalyse mit Berichtszeitpunkt August 2022 (18 Monate nach Beschlussfassung) daher keine aussagekräftigen Ergebnisse

enthalten werde. Er schlägt daher vor, dass zu diesem Zeitpunkt entgegen der Beschlussfassung des G-BA auf eine Prüfung auf Vergeblichkeit verzichtet wird.

Die Argumentation des pU erscheint inhaltlich sachgerecht, betrifft allerdings auch die Zwischenanalyse selbst. Es wird daher im vorliegenden Fall vorgeschlagen, auf die erste Zwischenanalyse inklusive Prüfung auf Vergeblichkeit zum Zeitpunkt 18 Monate nach Beschlussfassung zu verzichten und zu diesem Zeitpunkt lediglich einen Statusreport vorzulegen.

Zur weiteren Vereinfachung könnte darüber hinaus erwogen werden, die verbleibende zweite (ehemals dritte) Zwischenanalyse um 6 Monate vorzuziehen (von Februar 2026 auf August 2025) und dadurch mit dem dritten Statusreport zusammenzulegen. Die Anzahl der Berichtszeitpunkte könnte dadurch von 5 auf 4 ohne relevanten Informationsverlust verringert werden.

Insgesamt werden daher aufgrund des derzeitigen Verlaufs der Protokollerstellung für die AbD zu Onasemnogen-Abeparvovec folgende Berichtszeitpunkte vorgeschlagen:

- August 2022 (18 Monate nach Beschlussdatum): Statusreport
- Februar 2024 (36 Monate nach Beschlussdatum): erste Zwischenanalyse und Prüfung auf Vergeblichkeit, Statusreport
- August 2025 (54 Monate nach Beschlussdatum): zweite Zwischenanalyse und Prüfung auf Vergeblichkeit, Statusreport
- 1. Juli 2027: Finale Analyse

Zeitpunkt des Datenschnitts für Zwischenanalysen

Der pU gibt im Studienprotokoll an, dass der Datenschnitt für die Analysen etwa 6 Monate vor dem jeweiligen Berichtszeitpunkt erfolgen sollte, damit die Zeit für die Berichtserstellung jeweils ausreichend ist.

Dies ist für die finale Analyse nachvollziehbar, da hierfür ein vollständiges Dossier zur Nutzenbewertung zu erstellen ist. Für die weniger aufwändigen Berichte zu Zwischenanalysen erscheint dies jedoch nicht erforderlich und widerspricht auch dem eigenen Vorschlag des pU in den Beratungsgesprächen mit dem G-BA, in denen er 4 Monate als erforderlich Latenzzeit angegeben hatte. Die Datenschnitte sollten für die Zwischenanalysen daher jeweils 4 Monate, nicht 6 Monate, vor dem jeweiligen Berichtszeitpunkt erfolgen.

Berichtsformat und -inhalt

Der pU beschreibt für die Berichte zu Zwischenanalysen und zur Prüfung auf Vergeblichkeit, dass er diese auf Basis des Moduls 4 der Dossiervorlagen für Nutzenbewertungen nach § 35a SGB V erstellen wird. Dies erscheint sachgerecht. Allerdings sollte in Studienprotokoll und SAP auf eine konkrete Benennung zu befüllender Abschnitte der Dossiervorlagen verzichtet

werden. Denn es ist offen, ob die Struktur und der Inhalt der Dossievorlagen zum jeweiligen Berichtszeitpunkt dem heutigen Stand entsprechen werden.

Für die finale Analyse sollte klargestellt werden, dass diese zwar mit dem vollständigen Dossier zur Nutzenbewertung übermittelt wird, das Dossier aber darüber hinaus den Anforderungen an ein Dossier zur Nutzenbewertung nach § 35a SGB V genügen muss (unter anderem vollständige Darstellung der relevanten Evidenz zu Onasemnogen-Abeparvovec im Vergleich zur jeweiligen zweckmäßigen Vergleichstherapie).

Prüfung auf Vergeblichkeit

Im Zusammenhang mit der Prüfung auf Vergeblichkeit gibt der pU an, dass eine nicht ausreichende Fallzahl ggf. bereits für einen einzelnen „key endpoint“ ausreichend ist, um die Beobachtung für die jeweilige Population zu beenden. In einem solchen Fall sollen die Ergebnisse nicht ausgewertet werden. Beides erscheint nicht sachgerecht.

Die Prüfung auf Vergeblichkeit sollte die Gesamtschau aller Daten umfassen. Sie ist daher auch nur sinnvoll möglich, wenn dieser Prüfung eine vollständige Auswertung der Ergebnisse zum jeweiligen Zeitpunkt zugrunde liegt. Die entsprechenden Berichte zu den Zwischenanalysen müssen daher alle bis dahin erhobenen Ergebnisse und die zugehörigen Analysen vollständig enthalten.

Darüber hinaus erscheint es sinnvoll, die Entscheidung für oder gegen eine Fortsetzung der Beobachtung der Population in Abstimmung mit dem G-BA auf Basis des jeweiligen Zwischenberichts zu treffen.

3 Literatur

1. Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über die Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V: Onasemnogen-Abeparvovec (spinale Muskelatrophie); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen [online]. 2021 [Zugriff: 20.08.2021]. URL: https://www.g-ba.de/downloads/39-261-4702/2021-02-04_AM-RL-XII_awD_Onasemnogen-Abeparvovec_D-549_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: Onasemnogen-Abeparvovec (spinale Muskelatrophie); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen [online]. 2021 [Zugriff: 20.08.2021]. URL: https://www.g-ba.de/downloads/40-268-7293/2021-02-04_AM-RL-XII_awD_Onasemnogen-Abeparvovec_D-549_TrG.pdf.
3. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Konzept für eine anwendungsbegleitende Datenerhebung – Onasemnogen-Abeparvovec; Rapid Report [online]. 2020 [Zugriff: 09.02.2021]. URL: https://www.iqwig.de/download/a20-61_anwendungsbegleitende-datenerhebung-onasemnogen-abeparvovec_rapid-report_v1-0.pdf.
4. Novartis Gene Therapies. Routine data collection and evaluations of onasemnogene abeparvovec in Germany - Study Protocol. Protocol Number: COAV101A1DE01. Version: 1.01 (unveröffentlicht). 05.08.2021.
5. Novartis Gene Therapies. Routine data collection and evaluations of onasemnogene abeparvovec in Germany - Statistical Analysis Plan. Protocol Number: COAV101A1DE01. Version: 1.01 (unveröffentlicht). 05.08.2021.
6. Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV vom 11.08.2021, Beratungsanforderung 2021-B-190 der Firma Novartis Gene Therapies zum Wirkstoff Onasemnogen-Abeparvovec (unveröffentlicht). 09.09.2021.
7. Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV vom 29.06.2021, Beratungsanforderung 2021-B-122 der Firma Novartis Gene Therapies zum Wirkstoff Onasemnogen-Abeparvovec (unveröffentlicht). 23.08.2021.
8. Gemeinsamer Bundesausschuss. Zusammenfassende Dokumentation 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: Onasemnogen-Abeparvovec (spinale Muskelatrophie); Forderung einer anwendungsbegleitenden Datenerhebung und von Auswertungen [online]. 2021 [Zugriff: 13.09.2021]. URL: https://www.g-ba.de/downloads/40-268-7294/2021-02-04_AM-RL-XII_awD_Onasemnogen-Abeparvovec_D-549_ZD.pdf.