GBA
Gemeinsamer
Bundesausschuss

Nutzenbewertung

von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V

Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 11 i. V. m. Kapitel 5 § 12 Nr. 1 Satz 2 VerfO

Wirkstoff: Doxecitin/Doxribtimin

Datum der Veröffentlichung: 16. November 2026
Version 1.0

Inhaltsverzeichnis

Tabellenverzeichnis 3
Abbildungsverzeichnis 4
Abkürzungsverzeichnis 5
Hintergrund 7
1 Fragestellung 8
2 Auswahl und Methodik der relevanten Studien 9
2.1 Studienbasis für die Nutzenbewertung 9
2.2 Beschreibung der Studien 13
2.2.1 Charakterisierung der Einzelstudien 13
2.2.2 Integrierte Wirksamkeitsanalyse (ISE) 28
2.2.3 Integrierte Sicherheitsauswertung (ISS) 40
2.3 Endpunkte 45
2.3.1 Mortalität 46
2.3.2 Morbidität 48
2.3.3 Lebensqualität 55
2.3.4 Sicherheit 55
2.3.5 Erhebungszeitpunkte 57
2.4 Statistische Methoden 58
2.5 Verzerrungspotential auf Studien- und Endpunktebene 62
2.6 Indirekter Vergleich 62
3 Ergebnisse der eingeschlossenen Studien 63
3.1 Studiencharakteristika und Studienmedikation 63
3.2 Mortalität 63
3.3 Morbidität 63
3.4 Lebensqualität 63
3.5 Sicherheit 63
4 Diskussion der Methodik und Ergebnisse 64
4.1 Design und Methodik der studienübergreifenden Auswertungen (ISE und ISS) 64
4.2 Zulassungsstatus und Zulassungspopulation 70
4.3 Zusammenfassende Beurteilung zum Verzerrungspotential und zur Ergebnissicherheit 71
5 Anforderung an eine qualitätsgesicherte Anwendung 72
6 Zusammenfassung der Nutzenbewertung 73
Referenzen 74
7 Anhang 76
7.1 Studiencharakteristika und Studienmedikation 76
7.2 Mortalität 83
7.3 Morbidität 84
7.4 Sicherheit 90
7.5 Indirekter Vergleich zum Gesamtüberleben 93

Tabellenverzeichnis

Tabelle 1: Empfohlenes Dosierungsschema von Doxecitin/Doxribtimin g emäß Fachinformation 8
Tabelle 2: Übersicht über die Studienbasis 9
Tabelle 3: Charakterisierung der Studien 101, 102, 107 und 114 14
Tabelle 4: Charakterisierung der Intervention der Studien 101, 102, 107 und 114 24
Tabelle 5: Zusammenfassung der Endpunktbewertung der ISE und ISS 46
Tabelle 6: Charakterisierung der Studienpopulation; ISE-/ISS-Kernpool, Datenschnitt: 15.03.2024 78
Tabelle 7: Charakterisierung der Studienpopulation mit Informationen zur Diagnose; ISE-/ISS-Kernpool, Datenschnitt: 15.03.2024 81
Tabelle 8: Angaben zur Exposition mit der Studienmedikation; ISE-/ISS-Kernpool, Datenschnitt: 15.03.2024 83
Tabelle 9: Ergebnisse zu „Anzahl wiedererlangter motorischer Meilensteine nach Behandlungsbeginn“ und „Wiedererlangen des motorischen Meilensteins nach Behandlungsbeginn“; ISE-Kernpool, Datenschnitt: 14.03.2025 85
Tabelle 10: Ergebnisse zu „Anzahl verlorener motorischer Meilensteine nach Behandlungsbeginn“ und „Verlust des motorischen Meilensteins nach Behandlungsbeginn“; ISE-Kernpool, Datenschnitt: 14.03.2025 86
Tabelle 11: Ergebnisse zum Endpunkt „Künstliche Beatmung“; ISE-Kernpool, Datenschnitt: 14.03.2025 87
Tabelle 12: Ergebnisse zum Endpunkt „Künstliche Ernährung“; ISE-Kernpool, Datenschnitt: 14.03.2025 89
Tabelle 13: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE; ISS-Kernpool, Datenschnitt: 15.03.2024 91
Tabelle 14: UE mit Inzidenz ≥ 10 % nach Preferred Term; ISS-Kernpool, Datenschnitt: 15.03.2024 91
Tabelle 15: SUE, die bei ≥ 5 % der Personen aufgetreten sind, nach Preferred Term; ISS-Kernpool, Datenschnitt: 15.03.2024 92
Tabelle 16: Durch den pU identifizierte und bewertete Confounder 97
Tabelle 17: Datenverfügbarkeit von als relevant bewerteten Confoundern 98

Abbildungsverzeichnis

Abbildung 1: Patientenfluss der ISE (Eigene Darstellung)..... 29
Abbildung 2: Patientenfluss der ISS (Eigene Darstellung)..... 41
Abbildung 3: Teilpopulation der ISS und Nenner je UE-Kategorie (Eigene Darstellung) 41

Abkürzungsverzeichnis

AESI	Unerwünschtes Ereignis von besonderem Interesse (Adverse Event of Special Interest)
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
AST	Aspartat-Aminotransferase
ATE	Average Treatment Effect
ATO	Average Treatment Effect in the Overlap
BiPAP	Bilevel Positive Airway Pressure
CHMP	Committee for Medicinal Products for Human Use
COVID-19	Coronavirus-Krankheit-2019 (Coronavirus Disease 2019)
CPAP	Continuous Positive Airway Pressure
CTCAE	Common Terminology Criteria for Adverse Events
CUP	Compassionate Use Program
dC	Desoxycytidin
DC	Doxecitin
dCMP	Desoxycytidinmonophosphat
dT	Desoxythymidin
DT	Doxribtimin
dTMP	Desoxythymidinmonophosphat
EAP	Expanded Access Program
EMA	European Medicines Agency
EP	Evaluable Population
EPAR	European Public Assessment Report
FVC	Forcierte Vitalkapazität (Forced Vital Capacity)
GBA	Gemeinsamer Bundesausschuss
GGT	Gamma-Glutamyltransferase
GKV	Gesetzliche Krankenversicherung
GMP	Good Manufacturing Practice
IPTW	Inverse Probability of Treatment Weighting
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ISE	Integrierte Wirksamkeitsanalyse (Integrated Summary of Efficacy)
ISS	Integrierte Sicherheitsauswertung (Integrated Summary of Safety)
KI	Konfidenzintervall
MedDRA	Medical Dictionary for Regulatory Activities
mtDNA	mitochondriale DNA
MUPD	Modified Untreated Patient Database
N	Anzahl
PS	Propensity Score
PT	Preferred Term
pU	pharmazeutischer Unternehmer
SAP	Statistischer Analyseplan
SGB	Sozialgesetzbuch

SITA	Strongly Ignorable Treatment Assignment
SLR	Systematische Literaturrecherche
SMD	Standardisierte Mittelwertdifferenz (Standardized Mean Difference)
SMQ	Standardised MedDRA Query
SOC	Systemorganklasse
SUE	Schwerwiegendes unerwünschtes Ereignis
TK2	Thymidinkinase 2
TK2d	Thymidinkinase-2-Defizienz
TTE	Target Trial Emulation
UE	Unerwünschtes Ereignis
ULN	Obere Grenze des Normbereichs (Upper Limit of Normal)
UPD	Untreated Patient Database
VerfO	Verfahrensordnung des GBA
WHO	World Health Organization

Hintergrund

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

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

Bei Orphan Drugs erfolgt eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann, wenn der Umsatz des betreffenden Arzneimittels die Umsatzschwelle gemäß § 35a Absatz 1 Satz 12 SGB V überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt.

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

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Doxecitin/Doxribtimin in seiner Sitzung am 10. November 2026 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 17. August 2026 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pU in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

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

Die Nutzenbewertung wird am 16. November 2026 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des GBA (http://www.g-ba.de) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der GBA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Fragestellung

Doxecitin/Doxribtimin (KYGEVVI®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Absatz 1 Satz 11 1. Halbs. SGB V gilt für Orphan Drugs der medizinische Zusatznutzen bereits durch die Zulassung als belegt. Der GBA bestimmt bei Orphan Drugs, die einen Umsatz von 30 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien. Die Bewertung bezieht sich auf folgende Indikation [13]:

> Behandlung von Kindern und Erwachsenen mit genetisch bestätigter Thymidinkinase-2-Defizienz (TK2d) mit einem Alter bei Symptombeginn bis zur Vollendung des 12. Lebensjahres.

Gemäß Fachinformation richtet sich die Dosierung nach dem Gewicht der Patientinnen und Patienten und wird auf der Grundlage der individuellen Verträglichkeit titriert und dosiert, wobei die maximale empfohlene Erhaltungsdosis 400 mg/kg/Tag Doxecitin und 400 mg/kg/Tag Doxribtimin beträgt (siehe Tabelle 1).

Doxecitin/Doxribtimin sollte gemäß Fachinformationen täglich in drei gleichen Dosen zusammen mit Nahrung verabreicht werden.

Tabelle 1: Empfohlenes Dosierungsschema von Doxecitin/Doxribtimin[1)] gemäß Fachinformation [13]

Zeitpunkt	Dosierung
Anfangsdosis	130 mg/kg/Tag Doxecitin und 130 mg/kg/Tag Doxribtimin
Tag 14, Zwischendosis	260 mg/kg/Tag Doxecitin und 260 mg/kg/Tag Doxribtimin
Tag 28, Erhaltungsdosis	400 mg/kg/Tag Doxecitin und 400 mg/kg/Tag Doxribtimin

[1)] Für Personen mit mäßiger oder schwerer Nierenfunktionsstörung sollte eine langsamere Titration (mit einem Abstand von mind. 4 Wochen zwischen den jeweiligen Dosiserhöhungen) erfolgen.

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 2: Übersicht über die Studienbasis

Studienname (Nummer)	Mit Dossier vom pU eingereicht	Vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Relevant für die Nutzenbewertung	Hintergrund
Einzelstudien zu Doxecitin/Doxribtimin				
MT-1621-101[1)]	Ja	Nein	Nein[2)]	Retrospektiver, nicht interventioneller[3)] Chart Review bereits behandelter Personen ohne Behandlungszuweisung und ohne einheitlichen Untersuchungsplan. Die Studienpopulation geht in studienübergreifende Auswertungen ein (siehe gepoolte Studienauswertungen).
TK0102[1)]	Ja	Nein	Nein[2)]	Offene, einarmige Fortsetzungsstudie (Weiterbehandlung von 35 der 38 Personen aus MT-1621-101 sowie von 12 weiteren, außerhalb der Studien vorbehandelten Personen) in einer bereits behandelten Population[4)]. Die Studienpopulation geht in studienübergreifende Auswertungen ein (siehe gepoolte Studienauswertungen).
MT-1621-107[5)6)]	Ja	Nein	Nein	Retrospektiver, nicht interventioneller[3)] Chart Review ohne Behandlungszuweisung und ohne einheitlichen Untersuchungsplan. Die Studienpopulation geht in studienübergreifende Auswertungen ein (siehe gepoolte Studienauswertungen).
TK0114[5)]	Ja	Nein	Nein	Retrospektive EAP-Datenerhebung ohne Behandlungszuweisung und ohne einheitlichen Untersuchungsplan[3)]. Die Studienpopulation geht in studienübergreifende Auswertungen ein (siehe gepoolte Studienauswertungen).
Gepoolte Studienauswertungen				
ISE[7)] mit Daten aus: • MT-1621-101[1)] • TK0102[1)] • MT-1621-107 • TK0114	Ja	Ja	Nein	Wird nicht herangezogen (nähere Informationen siehe Kapitel 2.2.2).
ISS[8)] mit Daten aus: • MT-1621-101[1)] • TK0102[1)] • MT-1621-107	Ja	Ja	Nein	Wird nicht herangezogen (nähere Informationen siehe Kapitel 2.2.3).
ISE-/ISS-Kernpool[9)] mit Daten aus: • MT-1621-101[1)] • TK0102[1)]	Ja	Nein	Nein	Wird im Anhang dargestellt (nähere Informationen siehe Kapitel 2.2.2 und 2.2.3).

Studienname (Nummer)	Mit Dossier vom pU eingereicht	Vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Relevant für die Nutzenbewertung	Hintergrund
Studien zu externen Kontrollen				
UPD / Updated-UPD[10)]	Ja	Ja	Nein	Datenbank unbehandelter Personen, überwiegend aus publizierten Fallberichten und Fallserien; Grundlage der externen Kontrollpopulationen ISE-UPD und ISE-MUPD für die historischen Vergleiche zum Gesamtüberleben und zu den motorischen Meilensteinen (siehe Kapitel 2.6).
MT-1621-107[6)11)]	Ja	Ja	Nein	Von den 43 in Studie 107 dokumentierten unbehandelten Personen ergänzen 40[12)] die Updated-UPD (siehe Kapitel 2.6).

[1)] Zulassungsrelevante Studien gemäß EPAR [2].

[2)] Studie geht als Datenquelle in den ISE-/ISS-Kernpool ein, dessen Ergebnisse im Anhang dargestellt werden (siehe gepoolte Studienauswertungen; Begründung in Kapitel 2.2.2 und 2.2.3).

[3)] Nicht interventionell bedeutet, dass die Studie selbst keine Behandlung zuwies oder vorgab: Die Behandlung mit Pyrimidin-Nukleos(t)iden erfolgte vor der Zulassung außerhalb der Studie (CUP, prüfarztinitiierte Programme bzw. EAP) nach ärztlicher Entscheidung; die Studie wertete die vorhandenen Krankenakten bzw. Programmformulare nachträglich aus. Dies gilt entsprechend für MT-1621-107 und TK0114.

[4)] Studie 102 erhebt Daten zwar prospektiv, zentrales Einschlusskriterium war jedoch eine laufende Behandlung mit Pyrimidin-Nukleos(t)iden. Alle Personen waren bei Studieneintritt bereits für unterschiedlich lange, teils unbekannte Zeiträume mit Pyrimidin-Nukleos(t)iden behandelt; Voraussetzung für die Aufnahme in Studie 102 war, dass sie diese Zeit überlebt hatten und die Behandlung noch fortführten (Immortal Time Bias, Selektionsbias). Ein Ausgangsstatus vor Behandlungsbeginn liegt in Studie 102 nicht vor (siehe Kapitel 2.2.2).

[5)] Supportive Studien gemäß EPAR [2].

[6)] Neben 18 behandelten Personen umfasst dieser retrospektive Chart Review auch 43 unbehandelte Personen. MT-1621-107 hat damit zwei Funktionen: Datenquelle für behandelte Personen im ISE-Gesamtpool bzw. ISS-Gesamtpool und für unbehandelte Personen in der historischen Vergleichspopulation der ISE (ISE-UPD). Die Rolle dieses Studienanteils in der ISE-UPD ist in den Fußnoten 10 und 12 erläutert.

[7)] Die ISE ist die vom pU vorgelegte integrierte Wirksamkeitsanalyse. Für behandelte Personen werden Daten aus MT-1621-101, TK0102, MT-1621-107 und TK0114 nach personenbezogener Deduplizierung im ISE-Gesamtpool zusammengeführt (siehe Kapitel 2.2.2).

[8)] Die ISS ist die vom pU vorgelegte integrierte Sicherheitsauswertung. Der ISS-Gesamtpool umfasst ausschließlich die Studien MT-1621-101, TK0102 und MT-1621-107; TK0114 ist nicht Teil der ISS (siehe Kapitel 2.2.3).

[9)] Zulassungsrelevante gepoolte Studienauswertung (ISE-/ISS-Kernpool aus MT-1621-101 und TK0102) gemäß EPAR. [2]

[10)] Für die historischen Überlebensvergleiche wurden zwei externe Kontrolldatenbestände herangezogen. Die UPD ist eine aus publizierten Einzelfalldaten zusammengestellte Datenbank unbehandelter Personen mit TK2d; im Studienbericht MT-1621-101 wird sie als „101-UPD“ bezeichnet, weil sie als externe Kontrolle für diese Studie erstellt wurde – sie enthält keine Teilnehmenden der Studie 101. UPD und Updated-UPD beruhen auf patientenindividuellen Literaturdaten unbehandelter Personen mit TK2d (verfügbare Daten bis einschl. 17.06.2019 bzw. 31.10.2021). Aus der 101-UPD (N = 103) wurde die 101-MUPD (N = 69) abgeleitet, die ausschließlich als Kontrolle für MT-1621-101 diente. Die Updated-UPD (N = 128) wurde selbst nicht in Analysen verwendet: 113 individuelle unbehandelte Personen der Updated-UPD wurden mit 40 unbehandelten Personen aus MT-1621-107 zur ISE-UPD zusammengefasst (N = 153). Nach Ausschluss von 39 Personen für die Überlebenszeitanalyse (Alter bei Symptombeginn unbekannt: n = 9; letztes bekanntes Lebensalter unbekannt: n = 17; vor dem niedrigsten Alter bei Behandlungsbeginn in der ISE verstorben oder zensiert: n = 13) ergab sich die ISE-MUPD (N = 114), die als Kontrolle für die behandelten Populationen der ISE diente. Mit Symptombeginn im Alter von ≤ 12 Jahren umfassen die ISE-UPD 116 und die ISE-MUPD 92 Personen. Näheres zur Bildung der Populationen und zu den Analysen siehe Kapitel 2.6.

[11] Studie MT-1621-107 ist eine retrospektive, nicht interventionelle, multizentrische und multinationale Auswertung von Krankenakten (Chart Review) bei behandelten und unbehandelten Personen mit TK2d.

[12] Von 43 in MT-1621-107 als unbehandelt dokumentierten Personen wurden 3, die auch in TK0114 als behandelt erfasst waren, vor der Bildung der ISE-UPD ausgeschlossen; 40 unbehandelte Personen aus MT-1621-107 ergänzten damit die Updated-UPD in der ISE-UPD.

Abkürzungen: CUP: Compassionate Use Program; EAP: Expanded Access Program; EPAR: European Public Assessment Report; ISE: Integrierte Wirksamkeitsanalyse; ISS: Integrierte Sicherheitsauswertung; MUPD: Modified Untreated Patient Database; pU: pharmazeutischer Unternehmer; TK2d: Thymidinkinase-2-Defizienz; UPD: Untreated Patient Database.

Der pU legt in Modul 4 zur Ableitung des Zusatznutzens zwei studienübergreifende Auswertungen vor: eine integrierte Wirksamkeitsanalyse (ISE) und eine integrierte Sicherheitsauswertung (ISS). Die ISE fasst die Wirksamkeitsdaten aller behandelten Personen der Studien MT-1621-101 (im Folgenden Studie 101), TK0102 (Studie 102), MT-1621-107 (Studie 107) und TK0114 (Studie 114) zusammen (N = 104, davon N = 82 mit Symptombeginn im Alter von ≤ 12 Jahren). Die ISS umfasst die Sicherheitsdaten der Studien 101, 102 und 107 (N = 67, davon N = 50 mit Symptombeginn im Alter von ≤ 12 Jahren); Studie 114 ist nicht Bestandteil der ISS. Der pU begründet dies mit einer begrenzten Datenverfügbarkeit zur Sicherheit in der Studie 114 (siehe Kapitel 2.2.3). Ergebnisse auf Einzelstudienebene stellt der pU ausschließlich für die Sicherheit der Studie 114 deskriptiv dar.

Davon abzugrenzen ist der ISE-/ISS-Kernpool, der ausschließlich die beiden pivotalen Studien 101 und 102 umfasst (ISE-/ISS-Kernpool gesamt: N = 50, davon N = 39 mit Symptombeginn im Alter von ≤ 12 Jahren); der Kernpool umfasst damit nur Personen, die die Aufnahmebedingungen dieser Studien erfüllten – Behandlung sowie Überleben bis zur maßgeblichen Visite der Studie 101 bzw. laufende Behandlung bei Aufnahme in Studie 102 (siehe Kapitel 2.2.2). Studie 101 ist abgeschlossen; die noch laufende Studie 102 geht mit zwei Datenschnitten ein (14.03.2025 für die Morbiditätsendpunkte der ISE; 15.03.2024 für Gesamtüberleben, Patientencharakteristika und Exposition sowie für die ISS; siehe Kapitel 2.2.2 und 2.4). Dieser Kernpool umfasst die zulassungsbegründenden Studien, wird vom pU jedoch nicht zur Ableitung des Zusatznutzens in Modul 4 herangezogen.

Bei einer seltenen Erkrankung wie der TK2d können gepoolte Auswertungen einarmiger Datenquellen grundsätzlich sinnvoll sein, um mehr Personen in eine gemeinsame Auswertung einzubeziehen, als in den einzelnen Studien zur Verfügung stehen. Voraussetzung ist jedoch, dass die zugrunde liegenden Einzelstudien eine hinreichend valide Datengrundlage bieten. Die vier Einzelstudien (101, 102, 107 und 114) weisen bereits für sich genommen maßgebliche designbedingte Limitationen auf. Die Studien 101 und 107 sind retrospektive Auswertungen von Krankenakten; die Behandlung in Studie 101 erfolgte im Rahmen von „Compassionate Use Programs“ (CUP) bzw. „Expanded Access Programs“ (EAP). Studie 114 ist eine retrospektive Auswertung der Aufnahme- und Nachsorgeformulare des globalen EAP 113. Ein einheitlicher Untersuchungsplan lag diesen drei Studien nicht zugrunde: Art und Zeitpunkt der Untersuchungen richteten sich nach der klinischen Routine, sodass sich Umfang und Vollständigkeit der Daten zwischen Personen und Zentren unterscheiden können. Studie 102 ist eine einarmige Fortsetzungsstudie in einer bereits mit Pyrimidin-Nukleos(t)iden behandelten Population. Pyrimidin-Nukleos(t)ide bezeichnet hier die Wirkstoffe von Doxecitin (DC) und Doxribtimin (DT) – Desoxycytidin (dC) und Desoxythymidin (dT) – in nicht „Good Manufacturing Practice“ (GMP)-konformen Zubereitungen sowie deren Monophosphate dCMP/dTMP, die zwar nicht Bestandteil des zugelassenen Arzneimittels sind, deren Ergebnisse sich aus Sicht der „European Medicines Agency“ (EMA) jedoch auf das zugelassene Arzneimittel übertragen lassen (siehe Kapitel 2.2.1, Abschnitt „Charakterisierung der Intervention“).

Für keine der drei retrospektiven Studien ist der Patientenfluss bis zum Einschluss beschrieben: Es ist nicht dargelegt, nach welchen Kriterien die Zentren ausgewählt, wie viele Personen bzw.

Krankenakten gesichtet und aus welchen Gründen Personen gegebenenfalls nicht eingeschlossen wurden. Auch für Studie 102 ist die Auswahl vor Studienbeginn maßgeblich: Aufgenommen werden konnten ausschließlich Personen, die bereits mit Pyrimidin-Nukleos(t)iden (nicht GMP-konforme Formulierungen oder DC/DT) behandelt wurden. Von den 38 Personen der Studie 101 wurden 35 in Studie 102 übernommen; aus welchen Gründen die übrigen 3 Personen nicht übernommen wurden, ist nicht beschrieben.

Zudem konnten in drei der vier Studien nur Personen eingeschlossen werden, die zu einem bestimmten Zeitpunkt nach Behandlungsbeginn noch behandelt wurden bzw. noch als in Behandlung dokumentiert waren: Studie 101 setzte eine letzte dokumentierte Visite zwischen dem 01.06.2018 und dem 15.12.2018 voraus, Studie 102 eine laufende Behandlung mit Pyrimidin-Nukleos(t)iden bei Studieneintritt und Studie 114 eine Behandlung mit Nukleos(t)iden im EAP bis zum Stichtag 01.03.2024 (dabei ist unklar, ob diese zum Stichtag noch laufend sein musste). In den Studien 101 und 102 mussten die Personen damit die Zeit vom Behandlungsbeginn bis zum spezifizierten Visitenzeitfenster bzw. bis zur Aufnahme in Studie 102 überleben (erste Aufnahme am 10.07.2019; dies galt auch für die 12 neu aufgenommenen Personen); in Studie 114 galt dies zumindest für Personen, deren Behandlung bereits vor der Aufnahme in das EAP begonnen hatte. Früh verstorbene Personen und Personen mit Behandlungsabbruch dürften in den Studien damit systematisch unterrepräsentiert sein (Selektionsbias, Immortal Time Bias; siehe Kapitel 2.2.2). Zu erwarten ist eine Überschätzung des Behandlungseffekts, deren Ausmaß nicht abschätzbar ist. In Studie 107 war der Einschluss dagegen nicht an eine Visite oder eine laufende Behandlung zu einem bestimmten Zeitpunkt gebunden; eingeschlossen wurden alle Personen, für die geeignete Krankenakten oder zumindest eine Angabe zum Vitalstatus vorlagen, also auch bereits verstorbene Personen. Da jedoch auch hier die Auswahl der Personen nicht nachvollziehbar ist, ist ein Selektionsbias in keiner der vier Studien auszuschließen.

Studie 102 ist zwar die einzige Studie mit prospektiver Datenerhebung, kann die genannten Limitationen jedoch nicht ausgleichen, da die prospektive Erhebung erst nach der beschriebenen Selektion einsetzt (Überleben und fortgesetzte Behandlung bis zur Studienaufnahme). Alle 47 Personen waren bei Studieneintritt bereits vorbehandelt, und zwar unterschiedlich lange und teils über unbekannte Zeiträume: 35 Personen hatten zuvor an Studie 101 teilgenommen; für die 12 Personen ohne vorherige Teilnahme an Studie 101 sind Art, Dosis und Dauer der Vorbehandlung nicht je Person berichtet. Die als Baseline festgelegte Erhebung an Studientag 1, dem ersten Tag der Gabe von DC/DT innerhalb der Studie 102, bildet daher keinen unbehandelten Ausgangszustand ab, sondern einen Zustand unter laufender Vorbehandlung; Veränderungen gegenüber dieser Baseline beschreiben nur den weiteren Verlauf unter fortgesetzter Behandlung.

Aus den genannten Gründen werden die Einzelstudien nicht zur Nutzenbewertung herangezogen. Bei vorliegender Informationslage eignen sich auch die darauf basierenden gepoolten Auswertungen ISE, ISS und ISE-/ISS-Kernpool nicht für die Nutzenbewertung (siehe auch Kapitel 2.2.2 und 2.2.3). Der Kernpool besteht aus den vom pU gesponserten Studien 101 und 102, welche von der EMA als die zulassungsbegründenden Hauptstudien eingestuft wurden. Die Ergebnisse des ISE-Kernpools (Morbiditätsendpunkte: Datenschnitt 14.03.2025; Gesamtüberleben: Datenschnitt 15.03.2024) und des ISS-Kernpools (Datenschnitt 15.03.2024) für Personen mit Symptombeginn im Alter von ≤ 12 Jahren (N = 39) werden im Anhang dargestellt.

Neben den einarmigen studienübergreifenden Auswertungen legt der pU für das Gesamtüberleben verschiedene historische Vergleiche mit unbehandelten Personen vor. Da bereits für die behandelten Personen ein wesentlicher Selektionsbias nicht ausgeschlossen werden kann, werden auch diese indirekten Vergleiche nicht zur Nutzenbewertung herangezogen. Weitere Ausführungen finden sich in Kapitel 2.6.

Zur Nutzenbewertung für Doxecitin/Doxribtimin herangezogene Unterlagen

- Dossier des pU zu Doxecitin/Doxribtimin [10]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [2]
- Studienberichte inkl. Studienprotokolle und statistische Analysepläne (SAP) der Studien MT-1621-101 [16], TK0102 [14], MT-1621-107 [15] und TK0114 [9]
- Studienunterlagen zur ISE [11] und ISS [12]
- Fachinformation zu Doxecitin/Doxribtimin [13]

2.2 Beschreibung der Studien

Gepoolte Auswertungen einarmiger Einzelstudien können mehr Personen in eine gemeinsame Auswertung einbeziehen, als in den einzelnen Studien zur Verfügung stehen, setzen jedoch voraus, dass die zugrunde liegenden Einzelstudien eine hinreichend valide Datengrundlage bieten. Die Gründe, aus denen die vier Einzelstudien 101, 102, 107 und 114 sowie die darauf aufbauenden gepoolten Auswertungen nicht zur Nutzenbewertung herangezogen werden, sind in Kapitel 2.1 zusammengefasst und werden im Folgenden im Einzelnen dargelegt: In Kapitel 2.2.1 werden zunächst die vier Einzelstudien charakterisiert, einschließlich ihrer Datengrundlage und der Behandlungsexposition. Darauf aufbauend werden in den Kapiteln 2.2.2 und 2.2.3 die Bildung der gepoolten Auswertungen ISE und ISS beschrieben und bewertet.

2.2.1 Charakterisierung der Einzelstudien

Die 4 Einzelstudien sind in Tabelle 3 charakterisiert. Bewertungsrelevante Protokolländerungen werden im Abschnitt „Protokolländerungen“ beschrieben. Die Charakterisierung der Intervention bzw. der retrospektiv dokumentierten Behandlungsexposition ist in Tabelle 4 dargestellt.

Tabelle 3: Charakterisierung der Studien 101, 102, 107 und 114

Charakteristikum	Beschreibung			
	Pivotale Studien (ISE-/ISS-Kernpool)		Supportive Studien	
	Studie 101	Studie 102	Studie 107	Studie 114
Design und Studienablauf	**Design** **Retrospektiver**, einarmiger, multizentrischer Chart Review bei behandelten Personen mit TK2d. Die Behandlung erfolgte im Rahmen von CUP bzw. EAP mit nicht GMP-konformen Pyrimidin-Nukleos(t)iden[1)2)]. **Studienziele** Retrospektive Beschreibung von Sicherheit, Verträglichkeit und Wirksamkeit der Pyrimidin-Nukleos(t)id-Behandlung bei Personen mit TK2d; Untersuchung der Auswirkungen der Pyrimidin-Nukleos(t)ide auf TK2d-assoziierte klinische Parameter; Beschreibung des prätherapeutischen Krankheitsverlaufs von Symptombeginn bis Beginn der Pyrimidin-Nukleos(t)id-Behandlung; Identifizierung potentieller Zielgrößen für klinische Studien zu TK2d; Identifizierung von Erkrankungsmerkmalen, die während therapeutischer Interventionen überwacht werden sollten.	**Design** **Prospektive**, offene, einarmige, multizentrische interventionelle Phase-II-Fortsetzungsstudie bei zuvor behandelten Personen mit TK2d. Die Studienbehandlung erfolgte mit DC/DT. Zusätzlich konnten Personen ohne vorherige Teilnahme an Studie 101 mit Sponsorfreigabe in Studie 102 aufgenommen werden, sofern sie bereits mit nicht GMP-konformen Nukleos(t)iden oder mit DC/DT behandelt wurden. **Studienziele** Charakterisierung von Sicherheit und Verträglichkeit der fortgesetzten Behandlung mit DC/DT bei Personen mit TK2d, die bereits mit dC/dT, dCMP/dTMP oder DC/DT vorbehandelt waren; Untersuchung der fortgesetzten Wirksamkeit von DC/DT.	**Design** **Retrospektive**, multizentrische, multinationale, nicht interventionelle Chart-Review-Studie bei behandelten und unbehandelten Personen mit TK2d. **Studienziele** Retrospektive Erfassung des Vitalstatus bei unbehandelten Personen mit TK2d und bei Personen, die mit Pyrimidin-Nukleos(t)iden behandelt worden waren; Charakterisierung des klinischen Verlaufs; Beschreibung von Sicherheit und Verträglichkeit.	**Design** **Retrospektiver**, nicht interventioneller, multizentrischer Chart Review bei behandelten Personen mit TK2d. Die Behandlung mit DC/DT erfolgte im Rahmen des globalen EAP (Studie 113)[4)]. **Studienziele** Retrospektive Erfassung von Behandlungsexposition, Überleben und Verträglichkeit.

Charakteristikum	Beschreibung			
	Pivotale Studien (ISE-/ISS-Kernpool)		Supportive Studien	
	Studie 101	Studie 102	Studie 107	Studie 114
	Studienablauf Bereits vorhandene Angaben aus den Krankenakten der behandelten Personen wurden retrospektiv erhoben und in anonymisierter Form in eine elektronische Datenbank übertragen. Berücksichtigt wurden, soweit dokumentiert: • Vor Behandlungsbeginn: Krankheitsverlauf von Symptombeginn bis Beginn der Pyrimidin-Nukleos(t)id-Behandlung. • Nach Behandlungsbeginn: weiterer klinischer Verlauf sowie Angaben zu Sicherheit, Verträglichkeit und Wirksamkeit der Behandlung. Studie 101 selbst stellte kein Studienmedikament bereit und verabreichte keine Behandlung. Die individuellen Behandlungsregime wurden durch die behandelnden Ärztinnen und Ärzte festgelegt. Der Zeitraum der retrospektiven Datenerhebung bezeichnet die Aktenauswertung und Datenübertragung und ist von den individuellen Beobachtungs- und Behandlungszeiträumen zu unterscheiden.	**Studienablauf** Zuvor in Studie 101 erfasste Personen sowie weitere bereits vorbehandelte Personen aus CUP wurden aufgenommen. Bei Personen ohne vorherige Teilnahme an Studie 101 war eine Sponsorfreigabe erforderlich, um sicherzustellen, dass vor Behandlungsbeginn erhobene klinische und funktionelle Angaben für die Beurteilung von Sicherheit und Wirksamkeit ausreichend waren. Nach Studienaufnahme wurde DC/DT als Studienintervention verabreicht. Personen mit nicht GMP-konformer Vorbehandlung wurden auf DC/DT umgestellt; bereits mit DC/DT behandelte Personen setzten die Behandlung fort. Studienaufnahme und erstmaliger Beginn einer Nukleos(t)id-Behandlung sind daher nicht gleichzusetzen. Daten wurden während der planmäßigen Studienvisiten prospektiv dokumentiert. Die Auswahl spezifischer Untersuchungen richtete sich nach vorliegenden Voruntersuchungen, klinischem und funktionellem Status und wurde mit dem Sponsor und den Prüfärztinnen bzw. Prüfärzten abgestimmt.	**Studienablauf** Prüf- bzw. Studienpersonal identifizierte an teilnehmenden Zentren Krankenakten von Personen, die die Einschlusskriterien erfüllten. Relevante Angaben wurden aus den vorhandenen Krankenakten in die Studiendatenbank übertragen und retrospektiv ausgewertet. Erfasst wurden Personen, die bereits in der klinischen Versorgung mit Pyrimidin-Nukleos(t)iden behandelt worden waren, sowie Personen, die in den Krankenakten als unbehandelt dokumentiert waren[3)]. Studie 107 wies keiner Person eine Behandlung oder Nichtbehandlung zu, initiierte keine Therapie und regelte keine Dosisanpassung.	**Studienablauf** Personen des globalen EAP 113, die mit DC/DT behandelt worden waren und bis zum maßgeblichen Stichtag eine Behandlung mit Pyrimidin-Nukleos(t)iden erhalten hatten, wurden in die retrospektive Datenerhebung einbezogen. Behandlung, Überwachung und Umgang mit UE erfolgten im EAP nach ärztlichem Ermessen, lokaler Versorgungspraxis und regulatorischen Vorgaben. Für Studie 114 wurden die relevanten Angaben retrospektiv den bereits verfügbaren Aufnahme- und Nachsorgeformularen des globalen EAP entnommen und ausgewertet. Studie 114 wies keine Studienmedikation zu und initiierte keine Behandlung. Behandlung und etwaige Dosisanpassungen erfolgten im Rahmen des globalen EAP 113.

Charakteristikum	Beschreibung			
	Pivotale Studien (ISE-/ISS-Kernpool)		Supportive Studien	
	Studie 101	Studie 102	Studie 107	Studie 114
		Ergänzend wurden Angaben aus der Zeit zwischen dem Ende von Studie 101 und dem Beginn von Studie 102 retrospektiv erhoben.		
	Studienstatus Studie ist abgeschlossen (Ende der Datenerhebung: 28.03.2019)	**Studienstatus** Studie ist laufend (geplanter finaler Datenschnitt: August 2026)	**Studienstatus** Studie ist abgeschlossen (letzte Studienbeobachtung: 21.01.2022)	**Studienstatus** n. a.
Population	**Wesentliche Einschlusskriterien** • Bestätigte Mutation im TK2-Gen. • Verfügbare medizinische Aufzeichnungen seit Symptombeginn. • Einnahme von dC/dT und/oder dCMP/dTMP zur TK2d-Behandlung. • Eine letzte Visite mit Wirksamkeits- und/oder Sicherheitsparametern im Zeitraum vom 01.06.2018 bis 15.12.2018.	**Wesentliche Einschlusskriterien** • Bestätigte Mutation im TK2-Gen. • Laufende Behandlung mit Pyrimidin-Nukleos(t)iden; Teilnahme an Studie 101 oder Behandlung außerhalb dieser Studie möglich; Personen ohne vorherige Teilnahme an Studie 101 nur mit Zustimmung des Sponsors. • Bereitschaft zur Protokoll- und Arzneimitteladhärenz; stabiles Behandlungs- und Bewegungsregime.	**Wesentliche Einschlusskriterien** • Bestätigte Mutation im TK2-Gen. • Für die Datenerhebung geeignete Krankenakten oder mind. eine Information zum Vitalstatus. • Behandelte Personen: Einnahme von Pyrimidin-Nukleos(t)iden (dCMP/dTMP, dC/dT und/oder DC/DT) zur Behandlung der TK2d; zusätzlich wurden unbehandelte Personen erfasst.	**Wesentliche Einschlusskriterien** Eingeschlossen wurden Personen, die im globalen EAP 113[5)] mit DC/DT aufgenommen worden waren und bis zum Stichtag 01.03.2024 eine Behandlung mit Nukleos(t)iden erhalten hatten[6)].
	Wesentliche Ausschlusskriterien Andere genetische oder polygenetische Erkrankung.	**Wesentliche Ausschlusskriterien** • Lebererkrankung in der Anamnese oder ALT, AST oder Gesamtbilirubin ≥ 2 × ULN ohne vorherige Sponsorfreigabe. • Erkrankungen, die den Krankheitsverlauf wesentlich beeinflussen können.	**Wesentliche Ausschlusskriterien** Keine.	**Wesentliche Ausschlusskriterien** Keine.[7)]

Charakteristikum	Beschreibung			
	Pivotale Studien (ISE-/ISS-Kernpool)		Supportive Studien	
	Studie 101	Studie 102	Studie 107	Studie 114
Intervention und Zahl der Patientinnen und Patienten	**Behandelte Studienpopulation** • Gesamtzahl der zuvor geprüften Personen/Krankenakten sowie Nichteinschlüsse mit Gründen: k. A. • Eingeschlossen: N = 38[8)] • Davon mit Symptombeginn ≤ 12 Jahre: **N = 29**[9)] **Unbehandelte Population** Keine.[2)]	**Behandelte Studienpopulation** • Eingeschlossen: N = 47[10)] • Davon mit Symptombeginn ≤ 12 Jahre: **N = 38**[8)] **Unbehandelte Population** Keine.	**Behandelte Studienpopulation** • Gesamtzahl der zuvor geprüften Personen/Krankenakten sowie Nichteinschlüsse mit Gründen: k. A. • Eingeschlossen: N = 18[11)] • Davon mit Symptombeginn ≤ 12 Jahre: **N = 13**[9)] **Unbehandelte Studienpopulation[12)]** N = 43[13)]	**Behandelte Studienpopulation** • Gesamtzahl der zuvor geprüften Personen/Krankenakten sowie Nichteinschlüsse mit Gründen: k. A. • Eingeschlossen: N = 43[14)] • Davon mit Symptombeginn ≤ 12 Jahre: **N = 29**[9)] **Unbehandelte Studienpopulation** Keine.
Ort und Zeitraum der Durchführung; Datenschnitte	**Studienzentren** 8 Studienzentren: USA (1), Spanien (4), Israel (3) **Studienzeitraum** • Erstbehandlung: k. A. • Letzte berücksichtigte Visite: k. A. (laut Einschlusskriterium zwischen 01.06.2018 und 15.12.2018) • Zeitraum der retrospektiven Datenerhebung (Aktenauswertung und Datenübertragung): 30.10.2018 bis 28.03.2019[15)] **Datenschnitt** Ende der retrospektiven Datenerhebung: 28.03.2019	**Studienzentren** 7 Studienzentren: USA (1), Spanien (4), Israel (2) **Studienzeitraum** • Erste Aufnahme: 10.07.2019 • Letzte Visite der letzten Person: k. A. **Datenschnitt** • Es liegen Interimsanalysen zu den Datenschnitten 15.03.2024	**Studienzentren** 20 Studienzentren: USA, Spanien, Mexiko, Russland, Italien, Türkei, Vereinigtes Königreich **Studienzeitraum** • Erstbehandlung: k. A. • Letzte berücksichtigte Visite: k. A. • Zeitraum der retrospektiven Datenaufbereitung: 22.09.2021 bis 21.01.2022 **Datenschnitt** Ende der retrospektiven Datenerhebung: 21.01.2022	**Studienzentren** Studienzentren: k. A. (Wohnorte der Personen: Argentinien, Brasilien, Costa Rica, Deutschland, Frankreich, Ghana, Indien, Irland, Israel, Italien, Mexiko, Türkei, Ukraine, Vereinigtes Königreich) **Studienzeitraum** • Erstbehandlung: k. A. • Letzte berücksichtigte Visite: k. A. • Zeitraum der retrospektiven Datenaufbereitung: k. A. **Datenschnitt** 01.03.2024 (Stichtag der für den Bericht berücksichtigten Daten)[17)]

Charakteristikum	Beschreibung			
	Pivotale Studien (ISE-/ISS-Kernpool)		Supportive Studien	
	Studie 101	Studie 102	Studie 107	Studie 114
		(ISS) und 14.03.2025 (ISE) vor[16]; letzterer wurde im Zulassungsverfahren von der EMA angefordert. • Geplanter finaler Datenschnitt: August 2026		
Endpunkte gemäß Studienprotokoll	**Erhobene Datenvariablen** Mortalität: Vitalstatus sowie Alter/Datum bei Tod oder beim letzten bekannten Lebenszeichen. Morbidität • Motorische Meilensteine. • Motorische Funktion und Mobilität: 6MWT, MFM-20/ MFM-32, CHOP INTEND, HFMSE, NSAA, Egen-Klassifikation (EK1/EK2), RULM und MRC-Muskelkraftskala. • Aktivitäten des täglichen Lebens u. Hilfsmittel: Ankleiden/Baden sowie Nutzung unterstützender Hilfsmittel; Ankleiden/Baden wurde separat erfasst und nicht als Entwicklungsmeilenstein gewertet. • Respiratorischer Status: Beatmungsmodus einschl. invasiver / nicht invasiver Unterstützung und Tracheostoma, Beginn/Ende, Dauer und tägliche Stunden; Lungenfunktion (FVC, FEV1, MIP, SNIP, MEP, PEF, PCF),	**Primärer Endpunkt** Sicherheit. **Sekundäre Endpunkte** • Motorische Meilensteine. • Motorische Funktionstests: Egen-Klassifikation, HFMSE, CHOP INTEND, 6MWT, NSAA, RULM, 10MWRTT, 100MWRTT, EK-R, PROMIS-Items zur körperlichen Funktionsfähigkeit (Mobilität und obere Extremität). • Respiration und Ernährung: invasive / nicht invasive Beatmung; FEV1, FVC und weitere Atemparameter; kardiopulmonaler Belastungstest (CPET), nächtliche Kapnographie, Apnoe-Hypopnoe-Index, S_aO_2 und CO_2; Sondenernährung und Anteil der supplementären Ernährung. • Wachstum/BMI/Z-Scores, Schluckfunktion, Kreatinkinase/ Laktat/Cystatin C und weitere Biomarker, respiratorische	**Erhobene Datenvariablen** • Überlebensstatus. • Motorische Meilensteine. • Künstliche Beatmung und künstliche Ernährung: bei Verfügbarkeit Beginn, Ende, Modus, Tracheostoma, tägliche Beatmungsdauer und nächtliche Pulsoxymetrie bzw. Sondentyp. • CGI-S und CGI-I. • Sicherheit.	**Erhobene Datenvariablen** • Vitalstatus (lebend/verstorben), Datum des letzten Kontakts, Todesursache. • Sicherheit: keine eigene Erhebung; UE aus Sicherheitsdatenbank des Sponsors, soweit im EAP gemeldet (verpflichtend SUE und Schwangerschaften, weitere UE freiwillig).

Charakteristikum	Beschreibung			
	Pivotale Studien (ISE-/ISS-Kernpool)		Supportive Studien	
	Studie 101	Studie 102	Studie 107	Studie 114
	Schlafuntersuchungen, Thoraxbildgebung und respiratorische Infektionen. • Ernährungsstatus: nasogastrale Sonde oder Gastrostomie einschl. Einlage/Entfernung und Anteil supplementärer Ernährung. • Weitere klinische Parameter: Körpergröße, Körpergewicht, BMI, Kreatinkinase, Hospitalisierungen, chirurgische Eingriffe, körperlich/neurologische Befunde, Schlucktests und Biomarker. <u>Lebensqualität:</u> INQoL, EQ-5D-5L, PedsQL, Neuro-QOL Fatigue Questionnaire. <u>Sicherheit:</u> UE, SUE, TEAE/TESAE, Laborwerte, Vitalzeichen, EKG.	Infektionen, Hospitalisierungen und Operationen. • INQoL, CGI-I/PGI-I.		

[1)] Die in Studie 101 eingesetzten nicht GMP-konformen Pyrimidin-Nukleos(t)id-Formulierungen waren nicht mit dem zugelassenen Kygevvi-Fertigarzneimittel produkt- und darreichungsformidentisch. dC/dT entsprechen auf Wirkstoffebene DC/DT; dCMP/dTMP sind phosphorylierte Nukleotid-Monophosphate und nicht Bestandteil von Kygevvi. Die EMA erwartete aus den Unterschieden zwischen den historischen und kommerziellen Formulierungen keinen klinisch relevanten Einfluss und betrachtete die Daten als auf DC/DT extrapolierbar [2].

[2)] Für die Studie-101-spezifischen Überlebensanalysen wurde eine externe historische Vergleichspopulation aus publizierten Natural-History-Daten unbehandelter Personen herangezogen (101-UPD bzw. daraus abgeleitete 101-MUPD). Die externe historische Vergleichspopulation umfasste in der 101-UPD 103 Personen (publizierte Individualdaten, Literatursuche bis 17.06.2019). Die daraus abgeleitete 101-MUPD umfasste nach Ausschluss von Personen mit fehlendem Alter bei Symptombeginn (n = 4), fehlendem Alter bei Tod bzw. letztem Lebenszeichen (n = 13) oder einem letzten Lebensalter vor dem frühesten Behandlungsalter (n = 17) 69 Personen, davon 56 mit Symptombeginn im Alter von ≤ 12 Jahren. Ein gemeinsamer Datenschnitt für Studie 101 und 101-MUPD ist nicht dokumentiert Die Bewertung der historischen Vergleiche erfolgt in Kapitel 2.6.

[3)] Die Datenerhebung von Studie 107 war zunächst auf behandelte Personen begrenzt. Mit Amendment 2 vom 05.04.2021 wurde sie auf unbehandelte Personen erweitert.

[4)] EAP 113 war das globale Expanded-Access-Programm, in dessen Rahmen die Behandlung bereitgestellt wurde. Studie 114 ist die davon getrennte retrospektive Auswertung der in den Aufnahme- und Nachsorgeformularen des EAP verfügbaren Daten; TK0114 umfasste selbst keine Behandlung.

[5] Die medizinischen Voraussetzungen für die vorherige Aufnahme in das globale EAP TK0113 umfassten eine bestätigte TK2-Mutation, klinische Zeichen oder Symptome einer TK2d, ein Risiko schwerer funktioneller Beeinträchtigung oder Mortalität sowie das Fehlen einer geeigneten Möglichkeit zur Teilnahme an einer laufenden Sponsor-Studie. Diese Voraussetzungen waren Kriterien der EAP-Aufnahme und keine nachträglich festgelegten Einschlusskriterien von Studie 114.

[6] Unklar ist, ob die eingeschlossenen Personen zum Stichtag noch unter laufender Behandlung stehen mussten.

[7] Studie 114 führte keine erneute medizinische Selektion mit eigenen Ausschlusskriterien durch. Die inhaltlich relevanten Ausschlusskriterien galten für die vorherige Aufnahme in TK0113, insbesondere andere genetische Erkrankungen, relevante Leber- oder Nierenerkrankungen sowie ein vorangegangener behandlungsbedingter Abbruch einer Nukleos(t)id-Therapie. Sie sind daher nicht als studienspezifische Ausschlusskriterien von TK0114 zu interpretieren.

[8] Von den 38 behandelten Personen in Studie 101 nahmen 35 später auch an Studie 102 teil, darunter 28 Personen mit Symptombeginn im Alter von ≤ 12 Jahren.

[9] Die in Modul 4 je Studie ausgewiesenen Anzahlen der Personen mit Symptombeginn im Alter von ≤ 12 Jahren (29 + 38 + 13 + 29 = 109) beschreiben die Einzelstudien. Da Personen in mehreren Studien erfasst sein können, ist die Summe keine Anzahl eindeutiger Personen. Die anwendungsgebietsbezogene ISE-Population umfasst nach personenbezogener Deduplizierung 82 Personen; welche der 109 studienbezogen gezählten Personen in sie eingingen, weist der pU explizit aus; rechnerisch aus den ausgewiesenen Populationsgrößen (ISE-Kernpool N = 39, anwendungsgebietsbezogene ISS-Population N = 50, anwendungsgebietsbezogene ISE-Population N = 82) verteilen sich die 82 Personen nach der Studie der Ersterfassung auf die Studien 101 (29), 102 (10), 107 (11) und 114 (32); Herleitung siehe Kapitel 2.2.2, Schritt 3). Für Studie 114 bezieht sich die Angabe (29) auf die 34 ausschließlich in Studie 114 erfassten Personen; für die 9 auch in Studie 107 erfassten Personen wurde das Alter bei Symptombeginn in Studie 114 nicht erhoben, sondern aus Studie 107 übernommen (Studienbericht TK0114, Tabelle 10-2). Zusammen mit den 3 in Studie 107 als unbehandelt geführten Personen ergeben sich 32 erstmals in Studie 114 erfasste Personen im Anwendungsgebiet; hinzu kommen die auch in Studie 107 behandelten Personen mit Symptombeginn im Alter von ≤ 12 Jahren, deren Zahl nicht ausgewiesen ist.

[10] Von den 47 behandelten Personen in Studie 102 hatten 35 zuvor an Studie 101 teilgenommen; 12 Personen wurden ohne vorherige Teilnahme an Studie 101 neu in Studie 102 aufgenommen. Von den Personen mit Symptombeginn im Alter von ≤ 12 Jahren hatten laut ISE 28 zuvor an Studie 101 teilgenommen und 10 nicht (siehe Kapitel 2.2.2).

[11] 9 Personen wurden sowohl in Studie 107 als auch in Studie 114 erfasst. Diese Überlappung ist nur für die jeweilige Gesamtpopulation dokumentiert; eine getrennte Angabe für Personen mit Symptombeginn ≤ 12 Jahre liegt nicht vor.

[12] Mit Amendment 2 vom 05.04.2021 wurde die retrospektive Datenerhebung auf unbehandelte Personen erweitert; zuvor waren ausschließlich behandelte Personen vorgesehen.

[13] Studie 107 erfasste 43 unbehandelte Personen. Sie bilden keine Studienpopulation im Sinne der Intervention; 40 von ihnen gingen in die historische Vergleichspopulation der ISE (ISE-UPD) ein, deren Bewertung in Kapitel 2.6 erfolgt.

[14] Die 9 in beiden Studien erfassten Personen waren in Studie 107 entweder als behandelte oder als unbehandelte Personen dokumentiert. Demographische Angaben und Krankheitsmerkmale zu Studienbeginn wurden im Studienbericht zu 107 berichtet. Studie 114 ergänzte für diese Personen nur Datenvariablen, die in Studie 107 nicht erhoben worden waren.

[15] Der 15.12.2018 war laut Protokoll das Ende des Zeitfensters für die letzte Visite und zugleich das geplante Studienende; die retrospektive Datenerhebung (Aktenauswertung und Datenübertragung in die Datenbank) wurde wegen der späten Aufnahme mehrerer Personen in den Chart Review erst am 28.03.2019 abgeschlossen. Der 28.03.2019 ist damit kein klinischer Datenschnitt, sondern das Ende der retrospektiven Datenerhebung.

[16] Für Studie 102 liegen zwei Interimsdatenschnitte vor. Ob der Datenschnitt vom 15.03.2024 präspezifiziert war, ist in den vorgelegten Unterlagen nicht dargelegt. Der Interimsdatenschnitt vom 14.03.2025 wurde im Zulassungsverfahren von der EMA für die Wirksamkeitsendpunkte angefordert und war damit nicht a priori geplant. Beide Datenschnitte sind daher nicht als präspezifiziert anzusehen.

[17] Ein einheitlicher Zeitpunkt der letzten Datenerhebung bzw. letzten Studienbeobachtung ist aufgrund des retrospektiven Studiendesigns nicht nachvollziehbar. Als letztes Beobachtungsdatum lebender Personen unter Behandlung gilt der Stichtag vom 01.03.2024.

Abkürzungen: 6MWT: 6-Minute Walk Test; 10MWRTT: 10-Meter-Walk/Run-Timed-Test; 100MWRTT: 100-Meter-Walk/Run-Timed-Test; ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; BMI: Body-Mass-Index; CGI-I: Clinical Global Impressions scale – Improvement; CGI-S: Clinical Global Impressions scale – Severity; CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CPET: Cardiopulmonary Exercise Testing; CUP: Compassionate Use Program;

dC: Desoxycytidin; DC: Doxecitin; dCMP: Desoxycytidinmonophosphat; dT: Desoxythymidin; DT: Doxribtimin; dTMP: Desoxythymidinmonophosphat; EAP: Expanded Access Program; EK1/EK2: Egen-Klassifikation 1/2; EK-R: Egen-Klassifikation, revidiert; EKG: Elektrokardiogramm; EMA: European Medicines Agency; EQ-5D-5L: EuroQoL 5-Dimension 5-Level; FEV1: Forciertes exspiratorisches Volumen in 1 Sekunde; FVC: Forcierte Vitalkapazität; GMP: Good Manufacturing Practice; HFMSE: Hammersmith Functional Motor Scale Expanded; INQoL: Individualized Neuromuscular Quality of Life Questionnaire; ISE: Integrierte Wirksamkeitsanalyse; ISS: Integrierte Sicherheitsauswertung; k. A.: keine Angabe; MEP: Maximaler exspiratorischer Druck; MFM-20: Motor Function Measure scale-Short Form; MFM-32: Motor Function Measure – 32 items; MIP: Maximaler inspiratorischer Druck; MRC: Medical Research Council; MUPD: Modified Untreated Patient Database; n. a.: nicht anwendbar; Neuro-QOL: Quality of Life in Neurological Disorders; NSAA: North Star Ambulatory Assessment; PCF: Peak Cough Flow; PedsQL: Pediatric Quality of Life Inventory; PEF: Peak Expiratory Flow; PGI-I: Patient Global Impressions scale – Improvement; PROMIS: Patient-Reported Outcomes Measurement Information System; pU: pharmazeutischer Unternehmer; RULM: Revised Upper Limb Module; S_aO_2: Arterielle Sauerstoffsättigung; SNIP: Sniff Nasal Inspiratory Pressure; SUE: Schwerwiegendes unerwünschtes Ereignis; TEAE: Treatment-Emergent Adverse Event; TESAE: Treatment-Emergent Serious Adverse Event; TK2: Thymidinkinase 2; TK2d: Thymidinkinase-2-Defizienz; UE: Unerwünschtes Ereignis; ULN: Obere Grenze des Normbereichs; UPD: Untreated Patient Database.

Protokolländerungen

Zu allen Änderungen liegen lediglich zusammenfassende Angaben in der Änderungshistorie vor.

Studie 101

Das Originalprotokoll der Studie 101 wurde am 15.08.2018 finalisiert (keine nachträglichen Protokolländerungen).

Studie 102

Das am 07.01.2019 finalisierte Originalprotokoll der Studie 102 wurde im Studienverlauf durch fünf Amendments geändert. Für keine der vorliegenden Amendment-Fassungen ist dokumentiert, wie viele Personen zum jeweiligen Änderungsdatum eingeschlossen wurden. Daher ist nicht nachvollziehbar, welche Personen unter welcher Protokollfassung behandelt oder hinsichtlich neu eingeführter Erhebungen beurteilt wurden. Die erste Person wurde am 10.07.2019 in die Studie aufgenommen. Mit Amendment 3 vom 29.04.2021 wurde die Einnahme von DC/DT zulassungskonform mit Nahrung vorgesehen [13]. Außerdem wurden die „Clinical Global Impressions scale – Improvement" (CGI-I) und die „Patient Global Impressions scale – Improvement" (PGI-I) als Wirksamkeitserhebungen sowie Leistungskriterien für die motorischen Meilensteine ergänzt. Mit Amendment 4 vom 11.04.2022 wurden leberbezogene Kriterien für das Absetzen der Studienmedikation sowie Vorgaben für eine mögliche erneute Behandlung nach dem Absetzen ergänzt. Amendment 5 vom 18.08.2023 umfasste unter anderem Anpassungen der Sicherheitsdokumentation und der Verfahren zur Erkennung von Sicherheitssignalen, die Ergänzung einer Definition des Studienendes (globales Studienende, sobald die letzte Person im letzten Land die letzte Visite abgeschlossen hat; nach der Zulassung kann die Studie in Regionen mit kommerzieller Verfügbarkeit beendet werden) sowie Hinweise zu möglichen Dosisanpassungen bei Gewichtsveränderungen zwischen Studienvisiten. Zudem wurde die Erhebung von Biomarkern nach Monat 36 gestrichen. Die EMA wertete die Einnahme mit Nahrung und die Dosisanpassung bei Gewichtsveränderungen als Änderungen, die die Studienergebnisse potentiell beeinflussen; da danach nur wenige pharmakokinetische Proben erhoben wurden und keine quantifizierte Expositions-Wirkungs-Beziehung vorliegt, lässt sich ihr Einfluss auf Wirksamkeit und Sicherheit nicht abschätzen [2]. Diese Änderungen betreffen den Umfang der Sicherheits- und Verlaufserhebungen sowie die Vorgaben zur weiteren Behandlung und sind daher bei der Interpretation der Ergebnisse zu berücksichtigen.

Studie 107

Das Originalprotokoll der Studie 107 vom 12.10.2020 wurde durch zwei Amendments geändert. Beide Änderungen erfolgten vor Beginn der retrospektiven Datenerhebung am 22.09.2021. Mit Amendment 1 vom 14.12.2020 wurde die ursprünglich vorgesehene prospektive fünfjährige Nachbeobachtung gestrichen. Eine andere feste Nachbeobachtungsdauer als Ersatz wird im Studienbericht nicht genannt. Die durchgeführte Studie beschränkte sich auf die retrospektive Erfassung bereits vorhandener Angaben aus Krankenakten. Mit Amendment 2 vom 05.04.2021 wurde diese Datenerhebung auf unbehandelte Personen erweitert; zuvor waren ausschließlich behandelte Personen vorgesehen. Laut Studienbericht wurden 3 Personen unter Amendment 1 und die übrigen unter Amendment 2 eingeschlossen. Die konkreten Einschlussdaten und die Zeitpunkte, zu denen die jeweilige Protokollfassung an den einzelnen Zentren umgesetzt wurde, sind in den vorliegenden Angaben nicht ausgewiesen. Daher lässt sich nicht nachvollziehen, ob die 3 Personen bereits vor Beginn der Datenerhebung eingeschlossen wurden oder Amendment 1 bei ihrem Einschluss noch die am jeweiligen Zentrum angewendete Fassung war.

Studie 114

Für Studie 114 sind keine Änderungen eines Interventionsprotokolls, sondern Aktualisierungen der retrospektiven Datenerhebungs- und Datenmanagementgrundlage dokumentiert. Die Datenerhebung beruhte auf dem Methoden-Dokument vom 30.05.2023, dem SAP vom 19.07.2023 und dem „Data Management and Data Entry Plan“. Dieser wurde bis Version 6 vom 22.04.2024 mehrfach überarbeitet.

Mit Version 2 vom 11.08.2023 wurden Regeln zur Verarbeitung von Dosierungsangaben ergänzt. Dies betrifft die retrospektive Beschreibung der Behandlungsexposition, insbesondere Angaben zu Behandlungszeitraum, Dosis, Dosierungsfrequenz und Dosisänderungen. Wie viele Personen betroffen sind, ist nicht berichtet; der Einfluss auf die Expositionsangaben ist daher nicht beurteilbar (siehe Kapitel 2.2.2, Abschnitt „Behandlung und Exposition auf Studienebene“). Mit Version 3 vom 18.09.2023 wurden Regeln für unvollständige Angaben zum letzten Kontakt und zum Alter bei Symptombeginn ergänzt. Diese Angaben sind insbesondere für die Interpretation des Vitalstatus, der Nachbeobachtung und der Zuordnung zur anwendungsgebietsbezogenen ISE-Population relevant.

Charakterisierung der Intervention

Die in den Einzelstudien eingesetzte Intervention entsprach nicht durchgängig der später zugelassenen Arzneiform von DC/DT und dem in der Fachinformation vorgesehenen Dosierungsregime [13]. In Studie 101 wurden nicht GMP-konforme dC/dT- oder dCMP/dTMP-Präparationen verwendet. Studie 102 verabreichte GMP-konformes DC/DT nach Prüfplan; nicht GMP-konform vorbehandelte Personen wurden dabei auf DC/DT umgestellt. Studie 107 dokumentierte retrospektiv unterschiedliche Präparationen (dC/dT, dCMP/dTMP und/oder DC/DT) und Dosen außerhalb unternehmensgesponserter Studien. Studie 114 erfasste retrospektiv Personen, die im globalen EAP 113 mit DC/DT behandelt wurden, teils nach einer Vorbehandlung mit Pyrimidin-Nukleos(t)iden vor der EAP-Aufnahme. Somit war in keiner der Studien für alle Teilnehmenden sichergestellt, dass die Behandlung mit der zugelassenen Formulierung (DC/DT) und der in der Fachinformation empfohlenen Dosistitration erfolgte.

Die EMA erwartete von den Unterschieden zwischen den nicht kommerziellen Präparationen und der kommerziellen Formulierung jedoch keinen Einfluss auf die klinischen Ergebnisse und hielt die Daten von Personen mit Pyrimidin-Nukleos(t)id-Behandlung hinsichtlich der Formulierung für übertragbar auf das in Studie 102 verwendete kommerzielle DC/DT [2]. Auch die Unterschiede zwischen der GMP-konformen Zwei-Beutel- und der Ein-Beutel-Formulierung bewertete sie als geringfügig; zur Überbrückung beider Formulierungen lagen pharmakokinetische Studien vor. Dies ist relevant, da in Studie 102 bei allen Personen ein Wechsel der Darreichungsform dokumentiert ist. Von der Formulierung zu trennen ist nach Einschätzung der EMA die Evidenz zur Dosierung: Da weder Dosisfindungs- noch Dosis-Wirkungs-Studien und keine quantifizierte Expositions-Wirkungs-Beziehung beim Menschen vorliegen, bleiben die optimale Dosis und der Einfluss einzelner Dosisanpassungen unsicher [2]. Inwieweit die Personen der vier Studien vergleichbar behandelt wurden, wird in Kapitel 2.2.2 eingeordnet.

Tabelle 4: Charakterisierung der Intervention der Studien 101, 102, 107 und 114

Intervention
Studie 101
Wirkstoff: Nicht GMP-konforme Pyrimidin-Nukleos(t)id-Formulierungen (dC/dT oder dCMP/dTMP) • dC/dT entsprechen den Wirkstoffen Doxecitin/Doxribtimin auf Wirkstoffebene, nicht jedoch der zugelassenen kommerziellen GMP-Fertigarzneimittelform[1)]. • dCMP/dTMP sind Nukleotide und keine Wirkstoffe des zugelassenen KYGEVVI®; sie wurden in frühen, nicht GMP-konformen Formulierungen eingesetzt und später durch die Nukleoside dC/dT ersetzt[2)]. **Behandlungsschema/ Dosierung** Die Behandlungsschemata wurden individuell durch die behandelnden Ärztinnen und Ärzte festgelegt[3)]. Eine studienspezifische Dosierung oder ein studienspezifisches Behandlungsschema waren nicht vorgegeben. **Regeln zur Dosisunterbrechung, Dosisreduktion und Wiederaufnahme, Behandlungsabbruch und Begleitmedikationen** • Keine studienspezifischen Vorgaben zu Dosisunterbrechung, Dosisreduktion oder Wiederaufnahme. • Keine studienspezifischen Abbruchkriterien und keine Vorgaben zur Begleitmedikation[4)].
Studie 102
Wirkstoff und Verabreichung • Studienmedikation war die GMP-konforme 1:1-Kombination aus DC und DT. • Die benötigte Menge Pulver sollte zur Einnahme in Wasser oder Apfelsaft gelöst werden. Die zubereitete Lösung war laut Studienbericht 24 Stunden haltbar; verbleibende Lösung sollte verworfen werden. • Dosierung (siehe auch „Überleitung nach Vorbehandlung“) Alle eingeschlossenen Personen waren unter laufender Vorbehandlung mit nicht GMP-konformem dC/dT bzw. dCMP/dTMP oder mit DC/DT. o Für Personen mit vorausgegangener Behandlung mit nicht GMP-konformen dC/dT- oder dCMP/dTMP-Formulierungen war die Überleitung auf DC/DT vorgesehen; Personen mit bereits bestehender Behandlung mit DC/DT sollten diese grundsätzlich fortführen. o Die Tagesdosis war in 3 gleiche Gaben zu teilen. o Die Anwendung konnte oral oder über eine Ernährungssonde erfolgen[5)]; die Gaben sollten im Abstand von etwa 6 (± 2) Stunden verabreicht werden. • Überleitung nach Vorbehandlung o Bei stabiler Vorbehandlung mit jeweils 400 mg/kg/Tag dC oder dCMP und dT oder dTMP bzw. mit DC/DT wurde die Behandlung mit derselben Dosis DC/DT (400/400 mg/kg/Tag) fortgeführt. o Bei einer Vorbehandlung mit weniger als jeweils 400 mg/kg/Tag der beiden Komponenten von dC/dT, dCMP/dTMP oder DC/DT war die Überleitung auf die der bisherigen Dosis nächstgelegene Dosisstufe vorgesehen: 130/130, 260/260 oder 400/400 mg/kg/Tag DC/DT. Nach der Überleitung sollte der Prüfarzt die Behandlung mit der gewählten Dosis fortführen und die Dosis im weiteren Verlauf titrieren. **Dosisanpassungen, Unterbrechung, Absetzen und Wiederaufnahme der Studienmedikation** • Während der Auftitration konnte die Dosis abhängig von der individuellen Verträglichkeit angepasst werden; aus Verträglichkeitsgründen war eine Dosisreduktion möglich. • Für die Auftitration nach Überleitung auf eine niedrigere Dosisstufe bzw. bei erneuter Titration nach einer Unterbrechung (siehe unten) waren folgende Dosisstufen gemäß Titrationsschema des Studienprotokolls vorgesehen: Anfangsdosis 130/130 mg/kg/Tag DC/DT, Zwischendosis 260/260 mg/kg/Tag DC/DT, Erhaltungsdosis 400/400 mg/kg/Tag DC/DT. Die Dosissteigerungen sollten nur erfolgen, wenn die vorausgegangene Dosis sicher und verträglich war. Die Dosis war anhand des bei jeder Visite erhobenen Körpergewichts festzulegen, auf 3 gleiche Tagesgaben zu verteilen und ab Amendment 3 vom 29.04.2021 mit Nahrung einzunehmen.

Intervention
• Bei Unterbrechung oder Dosisänderung aus Verträglichkeitsgründen war kein Ausschleichen erforderlich. • Nach einer allgemeinen Unterbrechung konnte die Behandlung nach Entscheidung des Prüfarztes entweder in der vorherigen Dosis oder mit erneuter Titration fortgesetzt werden; maßgeblich waren die Dauer der Unterbrechung und die individuelle Verträglichkeit. • Ab Amendment 4 vom 11.04.2022 sollte die Studienmedikation unterbrochen werden, wenn eines der folgenden Kriterien erfüllt war: ALT oder AST > 8 × ULN; ALT oder AST > 5 × ULN über mehr als 2 Wochen; ALT oder AST > 3 × ULN zusammen mit Gesamtbilirubin > 2 × ULN oder INR > 1,5; oder ALT oder AST > 3 × ULN zusammen mit neu auftretendem Ikterus, Fatigue, Übelkeit, Erbrechen, rechtsseitigen Oberbauchschmerzen oder Druckschmerz, Fieber, Hautausschlag und/oder Eosinophilie (> 5 %). **Maßnahmen nach leberwertbedingter Unterbrechung und Wiederaufnahme** • Nach leberwertbedingter Unterbrechung wurde der Sponsor informiert. Die Person sollte, bevorzugt innerhalb von 48 bis 72 Stunden nach Bekanntwerden der auffälligen Werte, untersucht und eine örtliche hepatologische Fachperson hinzugezogen werden. Bis zum Abklingen der Symptome und/oder zur Stabilisierung der Leberwerte auf das Ausgangsniveau bzw. den ursprünglichen Abweichungsgrad waren engmaschige Kontrollen vorgesehen, bspw. anfangs zwei- bis dreimal pro Woche und bei Stabilisierung und Symptomfreiheit seltener. Nach hepatologischer Beurteilung sollten Sponsor und Prüfarzt unter Berücksichtigung der Krankengeschichte, anderer möglicher Ursachen und des individuellen Nutzen-Risiko-Verhältnisses über endgültiges Absetzen oder Wiederaufnahme entscheiden. Eine Wiederaufnahme war erst nach Rückkehr der Leberfunktionswerte auf das Ausgangsniveau, Rücksprache mit dem Medical Monitor und Einwilligung nach Risikoaufklärung vorgesehen. • Bei Wiederaufnahme nach einer leberwertbedingten Unterbrechung[6)] sollte die Anfangsdosis ungefähr der Hälfte der unmittelbar vor dem Absetzen erhaltenen Tagesdosis entsprechen, verteilt auf 3 gleiche Gaben. Die anschließende Aufdosierung sollte langsam in 4-Wochen-Abständen bis zur Erhaltungsdosis von 400/400 mg/kg/Tag DC/DT oder der höchsten verträglichen Dosis erfolgen. Bei einer vorausgegangenen Dosis von 400/400 mg/kg/Tag DC/DT war bspw. folgendes Vorgehen vorgesehen: Anfangsdosis der Wiederaufnahme an Tag 1: 200/200 mg/kg/Tag DC/DT; Zwischendosis an Tag 29: 260/260 mg/kg/Tag DC/DT; Erhaltungsdosis an Tag 57: 400/400 mg/kg/Tag DC/DT[7)]. Jede Stufe war auf 3 gleiche Tagesgaben zu verteilen; die Steigerungen setzten Sicherheit und Verträglichkeit der vorausgegangenen Stufe voraus. **Beendigung der Studienmedikation und Studienteilnahme** • Ein Abbruch der Studienmedikation und/oder Studienteilnahme war unter anderem bei Widerruf der Einwilligung, medizinischer Entscheidung des Prüfarztes, nicht akzeptabler Toxizität, Schwangerschaft, Nichteinhaltung des Protokolls oder Beendigung der Studie durch den Sponsor möglich. • Bei erneut auftretenden klinisch relevanten Auffälligkeiten nach Wiederaufnahme, bspw. ALT oder AST > 3 × ULN, sollte der Sponsor unverzüglich zur weiteren Beurteilung informiert werden; die Studienmedikation konnte erneut unterbrochen oder dauerhaft abgesetzt werden. **Erlaubte und nicht erlaubte Begleitbehandlung** • Begleitmedikationen einschl. sogenannter „mitochondrialer Cocktails“ (rezeptfreie und/oder verschreibungspflichtige Ergänzungspräparate) sollten während der Studienteilnahme in unveränderter Dosis und mit unverändertem Einnahmeschema fortgeführt werden. • Nach Umstellung auf DC/DT waren nicht GMP-konforme Nukleos(t)ide nicht zulässig. Andere Prüfpräparate zur Behandlung der TK2d oder anderer Indikationen waren während der Studie ebenfalls nicht erlaubt.

Intervention

Studie 107

Einordnung der dokumentierten Behandlung
Studie 107 verabreichte keine Studienmedikation und enthielt keine protokollierten Vorgaben zur Intervention. Retrospektiv wurden Angaben zu Personen erhoben, die in der klinischen Versorgung mit dCMP/dTMP, dC/dT und/oder DC/DT behandelt worden waren.

Dosierung, Dosisänderungen und Dosisabbruch, erlaubte und nicht erlaubte Begleitmedikationen
Keine studienspezifischen Vorgaben.

Studie 114 und EAP 113

EAP 113 war das EAP/CUP mit Behandlungsvorgaben für DC/DT. Studie 114 diente der retrospektiven Erhebung vorhandener Angaben aus dem globalen EAP 113. Die gemeinsame Übersicht beschreibt die Behandlungsvorgaben von 113 und getrennt davon die Auswahl und Datenerhebung von Studie 114; Studie 114 legte keine eigene Behandlung fest.

Retrospektive Datenerhebung 114

- In Studie 114 wurden Personen berücksichtigt, die in das globale EAP 113 aufgenommen worden waren, eine Einwilligung unterzeichnet und bis zum Stichtag 01.03.2024 Pyrimidin-Nukleos(t)ide erhalten hatten. Eine zusätzliche Auswahl nach Dosis, Ansprechen oder Verträglichkeit ist in den beschriebenen Kriterien nicht angegeben. Studie 114 erhob bereits vorhandene Angaben aus Aufnahme- und Nachbeobachtungsformularen und wies keine Behandlung oder zusätzliche Dosierung zu. Die Behandlung im globalen EAP erfolgte nach Entscheidung der behandelnden Ärztin bzw. des behandelnden Arztes.
- Wirkstoff und Darreichungsform: DC/DT als Pulver zur Herstellung einer Lösung zum Einnehmen; der Studienbericht zu Studie 114 beschreibt einen Beutel mit 2 g DC und 2 g DT. Die nachfolgenden Vorgaben zur Anwendung und Dosierung sind dem Behandlungsprogramm EAP 113 zugeordnet.
- Behandlungsdauer (EAP 113): Das Programm sollte beendet werden, wenn DC/DT im jeweiligen Land kommerziell verfügbar war oder der Sponsor das Programm beendete. Bei erstattungsfähiger Verfügbarkeit und einem nach ärztlicher Beurteilung bestehenden Nutzen sollte auf die kommerziell verfügbare Behandlung gewechselt werden.

Dosierung und Titration (EAP 113)

- Anfangsdosis: 130/130 mg/kg/Tag DC/DT.
- Zwischendosis: Bei ausreichender Verträglichkeit war nach 2 Wochen (Tag 15) eine Steigerung auf 260/260 mg/kg/Tag DC/DT vorgesehen.
- Erhaltungsdosis: Bei weiterhin ausreichender Verträglichkeit war nach weiteren 2 Wochen (Tag 29) eine Steigerung auf die Erhaltungsdosis von 400/400 mg/kg/Tag DC/DT vorgesehen.
- Die Dosierung sollte nach dem Körpergewicht bestimmt werden. Die Entscheidung über eine weitere Dosiserhöhung unter Berücksichtigung der Verträglichkeit lag bei der behandelnden Ärztin bzw. dem behandelnden Arzt.
- Die Tagesdosis sollte in 3 gleich großen Tagesgaben verabreicht werden.
- Die Lösung sollte in Wasser zubereitet und oral oder über eine enterale Ernährungssonde verabreicht werden. Die Gaben sollten in Abständen von etwa 6 (± 2) Stunden erfolgen und mit Nahrung eingenommen werden.
- Die behandelnde Ärztin bzw. der behandelnde Arzt sollte die Patientin bzw. den Patienten oder die Betreuungsperson zur Anwendung des Produkts anleiten.

Umgang mit einer ausgelassenen Dosis (EAP 113)

- Eine ausgelassene Dosis sollte möglichst zeitnah eingenommen werden. Eine doppelte oder zusätzliche Dosis zum Ausgleich war nicht vorgesehen.
- Lag der Zeitpunkt innerhalb der 2 Stunden vor der nächsten regulären Dosis, sollte die ausgelassene Dosis nicht nachgeholt werden; die nächste Dosis sollte zur üblichen Zeit eingenommen werden.

Dosisanpassungen (EAP 113)
Für die Aufdosierung war eine verträglichkeitsabhängige Anpassung vorgesehen; die behandelnde Ärztin bzw. der behandelnde Arzt entschied, ob eine höhere Dosis toleriert werden konnte.

Intervention
Dosisunterbrechung (EAP 113) • Die Behandlung sollte bei ALT oder AST > 8 × ULN unterbrochen werden. • Eine Unterbrechung war ebenfalls vorgesehen bei ALT oder AST > 5 × ULN über mehr als 2 Wochen. • Weitere Kriterien waren ALT oder AST > 3 × ULN zusammen mit Gesamtbilirubin > 2 × ULN oder INR > 1,5. • Eine Unterbrechung war außerdem vorgesehen bei ALT oder AST > 3 × ULN zusammen mit klinischen Zeichen einer möglichen Leberschädigung, bspw. Fatigue, Übelkeit, Erbrechen, Schmerzen oder Druckempfindlichkeit im rechten Oberbauch, Fieber, Hautausschlag oder Eosinophilie > 5 %. **Wiederaufnahme nach leberwertbedingter Unterbrechung (EAP 113)** • Eine Wiederaufnahme sollte erst nach Rücksprache mit der behandelnden Ärztin bzw. dem behandelnden Arzt und nach Rückkehr der Leberfunktionswerte auf das Ausgangsniveau erfolgen. • Bei Wiederaufnahme sollte die Anfangsdosis ungefähr der Hälfte der unmittelbar vor dem Absetzen erhaltenen Tagesdosis entsprechen, aufgeteilt auf 3 gleiche Gaben. Anschließend war eine langsame Aufdosierung in 4-Wochen-Abständen bis zur Erhaltungsdosis von 400/400 mg/kg/Tag DC/DT oder bis zur höchsten verträglichen Dosis vorgesehen[8)]. • Bei einer vorausgegangenen Dosis von 400/400 mg/kg/Tag DC/DT war bspw. folgendes Vorgehen vorgesehen: Anfangsdosis der Wiederaufnahme an Tag 1: 200/200 mg/kg/Tag DC/DT[9)]; Zwischendosis an Tag 29: 260/260 mg/kg/Tag DC/DT; Erhaltungsdosis an Tag 57: 400/400 mg/kg/Tag DC/DT. Die Tagesdosis war jeweils auf 3 gleiche Gaben zu verteilen; eine Steigerung setzte voraus, dass die vorherige Dosis sicher und verträglich war. • Nach Wiederaufnahme und nach jeder Dosissteigerung sollten die Leberwerte mind. alle 2 Wochen kontrolliert werden. Erst wenn nach 4 Wochen unter der endgültigen Zieldosis keine Auffälligkeiten vorlagen und die Person symptomfrei war, konnten die Kontrollen wieder im regulären Untersuchungsintervall erfolgen. **Dauerhafte Beendigung und Unterbrechung des Programms (EAP 113)** • Die Patientin bzw. der Patient, die Betreuungsperson oder die gesetzliche Vertretung konnte die Einwilligung und Programmteilnahme jederzeit ohne Nachteile für die weitere Behandlung widerrufen. • Eine Beendigung war nach Ermessen der behandelnden Ärztin bzw. des behandelnden Arztes unter anderem bei Schwangerschaft, Nichteinhaltung des Behandlungsprotokolls, Verlust der Nachbeobachtbarkeit, im Sicherheits- oder medizinischen Interesse der Person, Unverträglichkeit, Tod oder Beendigung des Programms bzw. der Wirkstoffentwicklung durch den Sponsor vorgesehen. • Der Sponsor konnte die Behandlung zudem aussetzen oder dauerhaft beenden, bspw. bei Lieferengpässen, Ablehnung einer Zulassung oder Sicherheitsbedenken. • Bei dauerhafter Beendigung der Behandlung war eine Abbruchvisite empfohlen. **Begleitmedikation, Impfungen und nicht zulässige Behandlungen (EAP 113)** • Begleitmedikationen (einschl. sogenannter „mitochondrialer Cocktails“) sollten im Aufnahmeformular dokumentiert werden. Eine weitere Erfassung nach Aufnahme ist im Behandlungsplan nicht beschrieben. In Studie 114 wurden Begleitmedikationen nicht erfasst. • Keine relevanten studienspezifischen Vorgaben zu Begleitmedikationen und nicht zulässigen Behandlungen.

[1)] 32 der 38 Personen begannen die Behandlung mit nicht GMP-konformem dC/dT. Weitere 5 Personen wechselten im Verlauf von dCMP/dTMP auf dC/dT. Ein einheitliches Anfangs-, Zwischen- und Erhaltungsdosisschema lässt sich aus den Studienunterlagen nicht ableiten.

[2)] 6 der 38 Personen begannen die Behandlung mit nicht GMP-konformem dCMP/dTMP. Anfangsdosen von 100/100 bzw. 200/200 mg/kg/Tag dCMP/dTMP und unterschiedliche weitere Dosissteigerungen wurden dokumentiert. 5 Personen wechselten während oder nach der Dosissteigerung auf dC/dT. 1 Person erhielt bis zum Ende der Datenerhebung weiterhin dCMP/dTMP in einer zuletzt dokumentierten Dosis von 400/400 mg/kg/Tag.

[3)] Die individuell festgelegten Dosierungen liegen nicht vor und könnten vom empfohlenen Dosierungsschema der Fachinformation (Anfangsdosis 130/130, Zwischendosis 260/260 und Erhaltungsdosis 400/400 mg/kg/Tag DC/DT) abweichen [13].

[4)] Erfasst wurden, soweit in den Krankenakten verfügbar, Medikamente zur Behandlung von Manifestationen der TK2d, während Hospitalisierungen verabreichte Medikamente sowie Medikamente zur Behandlung relevanter Erkrankungen oder von UE im Zusammenhang mit dC/dT bzw. dCMP/dTMP. Dies umfasste verschreibungspflichtige und nicht verschreibungspflichtige Arzneimittel, pflanzliche Präparate sowie mitochondriale Ergänzungspräparate („mitochondriale Cocktails"). Als Begleitmedikation galten Medikamente, die bei Beginn der außerhalb der Studie eingeleiteten Pyrimidin-Nukleos(t)id-Behandlung bereits angewendet oder während dieser Behandlung neu begonnen wurden.

[5)] Vor Amendment 3 konnte die Einnahme mit oder ohne Nahrung erfolgen. Seit Amendment 3 vom 29.04.2021 war die Einnahme mit Nahrung vorgegeben.

[6)] Nach Wiederaufnahme und nach jeder weiteren Dosissteigerung sollten die Leberwerte mind. alle 2 Wochen kontrolliert werden. Soweit möglich, war eine begleitende pharmakokinetische Probenentnahme vorgesehen. Erst wenn nach 4 Wochen unter der endgültigen Zieldosis keine Auffälligkeiten vorlagen und die Person symptomfrei war, konnten die Leberwertkontrollen wieder nach dem regulären Untersuchungsplan erfolgen und die zugehörigen pharmakokinetischen Probenentnahmen entfallen.

[7)] Das beschriebene Wiederaufnahmeschema betrifft die erneute Behandlung nach einer leberwertbedingten Unterbrechung. Die Anfangsdosis von 200/200 mg/kg/Tag DC/DT gilt für das Beispiel einer vorausgegangenen Dosis von 400/400 mg/kg/Tag DC/DT; allgemein war ungefähr die Hälfte der unmittelbar vor der Unterbrechung erhaltenen Tagesdosis vorgesehen. Anfangsdosis und Zeitabstände der Dosissteigerungen unterscheiden sich vom empfohlenen Dosierungsschema der Fachinformation für den Behandlungsbeginn; ein Schema für die Wiederaufnahme nach leberwertbedingter Unterbrechung enthält die Fachinformation nicht[13].

[8)] Das Wiederaufnahmeschema des EAP 113 entspricht dem der Studie 102. Die Anfangsdosis von 200/200 mg/kg/Tag DC/DT (bei vorausgegangener Dosis von 400/400 mg/kg/Tag) und die 4-Wochen-Abständen der Dosissteigerungen unterscheiden sich vom Dosierungsschema der Fachinformation für den Behandlungsbeginn (Anfangsdosis 130/130 mg/kg/Tag DC/DT, Steigerungen an Tag 14 und Tag 28).

[9)] Anfangsdosis 200/200 mg/kg/Tag DC/DT entspricht nicht der empfohlenen Anfangsdosis gemäß Fachinformation (130/130 mg/kg/Tag DC/DT) [13].

Abkürzungen: ALT: Alanin-Aminotransferase; AST: Aspartat-Aminotransferase; CUP: Compassionate Use Program; dC: Desoxycytidin; DC: Doxecitin; dCMP: Desoxycytidinmonophosphat; dT: Desoxythymidin; DT: Doxribtimin; dTMP: Desoxythymidinmonophosphat; EAP: Expanded Access Program; GMP: Good Manufacturing Practice; INR: International Normalized Ratio; TK2d: Thymidinkinase-2-Defizienz; UE: Unerwünschte Ereignis; ULN: Obere Grenze des Normbereichs.

2.2.2 Integrierte Wirksamkeitsanalyse (ISE)

Die ISE führt Individualdaten behandelter Personen aus vier Studien zusammen (Pooling): Aus zwei abgeschlossenen retrospektiven, nicht interventionellen Krankenaktenauswertungen (Studien 101 und 107), aus der prospektiven, offenen Phase-II-Studie 102 (zum Zeitpunkt der Dossiereinreichung noch laufend) sowie aus der retrospektiven, nicht interventionellen Datenerhebung der Studie 114 zu Personen, die im globalen EAP 113 behandelt wurden (Stichtag: 01.03.2024). Der nachträglich studienübergreifend gebildete und um Mehrfacherfassungen bereinigte ISE-Gesamtpool (N = 104) ist keine prospektiv geplante Kohorte und auch keine Metaanalyse: Die Daten werden ohne Berücksichtigung der Studienzugehörigkeit zusammengelegt (eine nach Studien getrennte Darstellung fehlt) und die einarmigen, nach Design und Datenerhebung unterschiedlichen Studien bilden keine metaanalytisch auswertbare Studienbasis. Ziel der ISE ist es, die Wirksamkeit an einer größeren Personenzahl als in den Einzelstudien zu untersuchen. Daraus folgen weder eine klinische Homogenität der vier Studien noch vollständige Daten zu den einzelnen Endpunkten.

Für die nachfolgende Darstellung gelten die Bezeichnungen aus Abbildung 1: Der ISE-Gesamtpool (N = 104) umfasst alle behandelten Personen der vier Studien 101, 102, 107 und 114 unabhängig vom Alter bei Symptombeginn. Die anwendungsgebietsbezogene ISE-Population (N = 82) ist die Teilpopulation des ISE-Gesamtpools mit Symptombeginn im Alter von ≤ 12 Jahren. Der ISE-Kernpool gesamt (N = 50) umfasst nur die behandelten Personen der Studien 101 und 102. Als ISE-Kernpool (N = 39) wird im Folgenden dessen Teilpopulation mit Symptombeginn im Alter von ≤ 12 Jahren bezeichnet; sie ist zugleich Teil der anwendungsgebietsbezogenen ISE-Population und wird im Anhang dargestellt.

Das Alter bei Symptombeginn ist dabei nur ein zusätzliches Abgrenzungskriterium: In alle Teilpopulationen gingen nur Personen ein, die zuvor die Aufnahmebedingungen der jeweiligen Studie erfüllt hatten. In Studie 101 waren dies eine Behandlung und das Überleben bis zu einer Visite zwischen dem 01.06.2018 und dem 15.12.2018, in Studie 102 eine laufende Behandlung bei Aufnahme und in Studie 114 die Aufnahme in das EAP (für bereits vorbehandelte Personen also das Überleben bis zur EAP-Aufnahme); Studie 107 erfasste auch verstorbene Personen (siehe Abschnitt „Auswahl und Bildung des ISE-Gesamtpools und der Teilpopulationen“). Für die Sicherheit gelten die entsprechenden Bezeichnungen der ISS: ISS-Gesamtpool (N = 67; Studien 101, 102 und 107), anwendungsgebietsbezogene ISS-Population (N = 50), ISS-Kernpool gesamt (N = 50; Studien 101 und 102) und ISS-Kernpool (N = 39; Symptombeginn im Alter von ≤ 12 Jahren). Die Sicherheitsauswertungen werden in Kapitel 2.2.3 beschrieben.

Die Ergebnisse des ISE-Kernpools zu den Morbiditätsendpunkten beruhen für die Studie 102 auf dem Datenschnitt vom 14.03.2025; die Angaben zum Gesamtüberleben sowie zu Patientencharakteristika und Exposition des ISE-Kernpools liegen – wie auch alle Angaben des ISS-Kernpools – nur zum Datenschnitt vom 15.03.2024 vor.

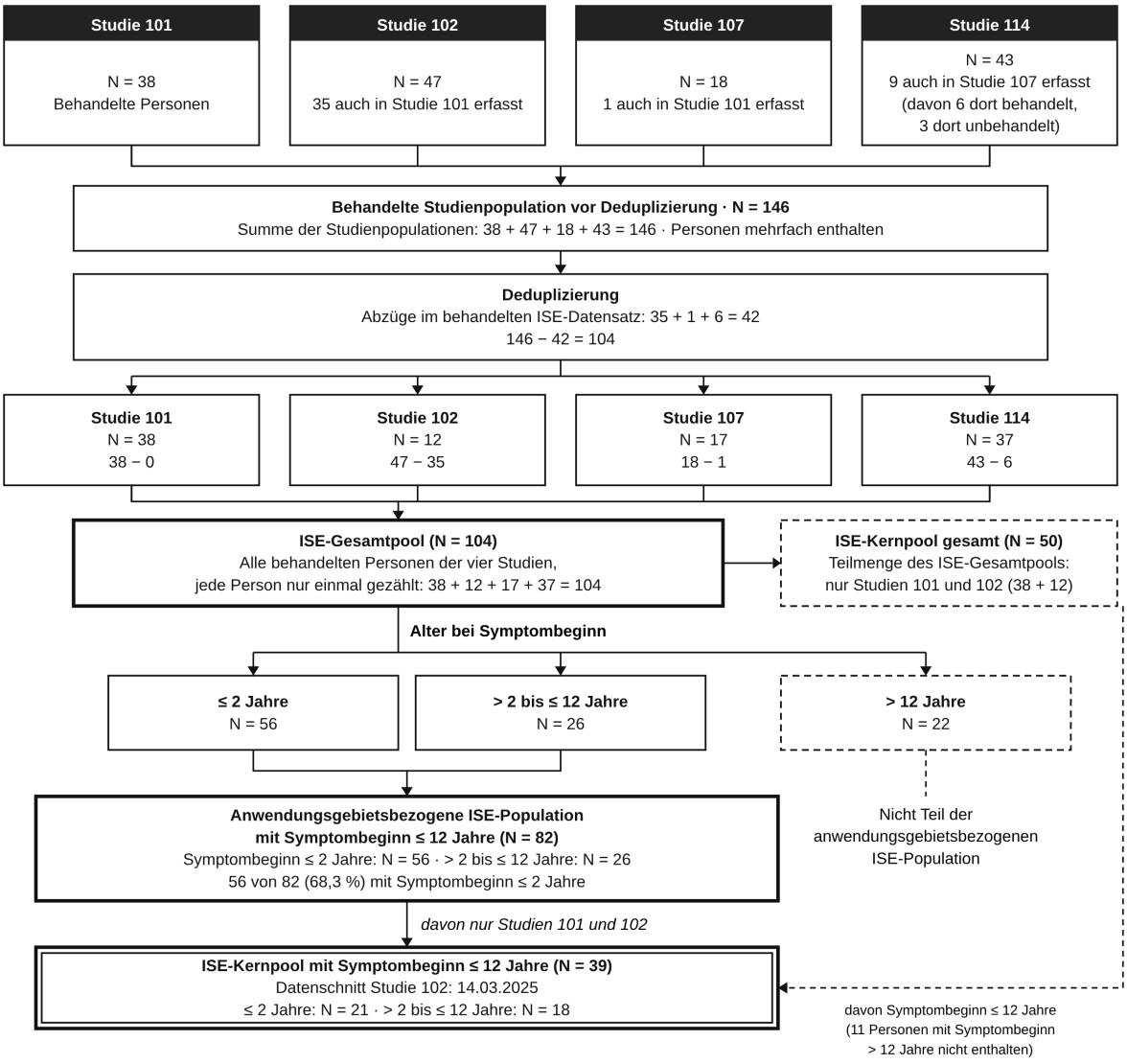

Abbildung 1: Patientenfluss der ISE (Eigene Darstellung) [10, 11]

Die nachfolgende Beschreibung beschränkt sich auf die behandelten Personen der vier Studien. Zunächst werden die Datengrundlagen der Studien sowie Behandlung und Exposition beschrieben. Anschließend werden die Auswahl der Personen, die Bildung des ISE-Gesamtpools (N = 104) und der daraus abgeleiteten Teilpopulationen (anwendungsgebietsbezogene ISE-Population (N = 82), ISE-Kernpool (N = 39)) sowie die zeitliche Vereinheitlichung der Studiendaten dargestellt; eine Bewertung der Zusammenführung schließt den Abschnitt ab. Die Endpunkte werden in Kapitel 2.3, die Vergleiche mit unbehandelten Personen (historische und indirekte Vergleiche) in Kapitel 2.6 bewertet.

Quellstudien und zugrunde liegende Studiendaten

Studie 101 (38 Behandelte): Vom 30.10.2018 bis zum 28.03.2019 wurden Angaben aus den Krankenakten der Zentren in elektronische Prüfbögen (eCRF) übertragen und im Monitoring anhand der Quelldokumente überprüft. Erfasst werden sollte der Krankheitsverlauf von Symptombeginn bis zur letzten dokumentierten Visite. Wann die erste berücksichtigte Datenerhebung bzw. Untersuchung stattgefunden hat und ob die Krankenakten nur Untersuchungsaufzeichnungen oder auch anamnestische Angaben umfassten, die noch vor der ersten berücksichtigten Untersuchung lagen, ist nicht berichtet. Die Untersuchungen erfolgten im Rahmen der Routineversorgung ohne einheitlichen Visitenplan. Art und Häufigkeit der dokumentierten Untersuchungen unterscheiden sich daher zwischen Personen und Zentren; manche Untersuchungen wurden in einzelnen Zentren erst im Verlauf eingeführt. 35 der 38 Personen wurden anschließend in Studie 102 aufgenommen; 2 der übrigen 3 Personen hatten die Behandlung in Studie 101 wegen erhöhter Leberwerte abgebrochen und konnten daher nicht übergehen, für die dritte Person ist der Grund nicht beschrieben.

Studie 102 (47 vorbehandelte Personen): Eine prospektive, offene, einarmige Phase-II-Fortsetzungsstudie mit planmäßigen Visiten. Welche Untersuchungen bei Einschluss erfolgten, legten Sponsor und Prüfpersonal für jede Person individuell fest, abhängig davon, welche Untersuchungen bereits vor Studieneinschluss durchgeführt worden waren. Prospektiv dokumentiert wurde ab Studieneintritt. Die Behandlung zwischen dem Ende der Studie 101 und dem Eintritt in Studie 102 sowie die medizinische Vorgeschichte (Vorerkrankungen und frühere Krankheitsereignisse) der Personen ohne Teilnahme an Studie 101 wurden retrospektiv erfasst. Für die Zeit vor dem Einschluss in Studie 102 liegen keine personenbezogenen Angaben zum Behandlungszeitraum der Teilnehmenden vor Studieneinschluss in Studie 102 vor – für die 35 Personen aus Studie 101 nicht für die Zeit zwischen beiden Studien, für die 12 übrigen Personen nicht für die gesamte Vorbehandlung; für letztere sind auch Art und Dosis der Vorbehandlung nicht je Person ausgewiesen. Während der COVID-19-Pandemie fanden Visiten teilweise als Fernvisiten statt, sodass einzelne Untersuchungen entfielen. Es liegen zwei Datenschnitte vor (15.03.2024 und 14.03.2025). Den späteren Datenschnitt forderte die EMA im Zulassungsverfahren ausschließlich für die Wirksamkeitsauswertungen der Studie 102 an [2]; für die Sicherheit liegt der Datenschnitt vom 15.03.2024 vor (nähere Informationen zur ISS siehe Kapitel 2.2.3). Für Studie 102 gehen daher in die anwendungsgebietsbezogene ISE-Population (N = 82) und den ISE-Kernpool (N = 39) die Daten zum Datenschnitt vom 14.03.2025 ein; für Gesamtüberleben, Patientencharakteristika und Exposition des ISE-Kernpools gilt der Datenschnitt vom 15.03.2024. Gemäß Modul 4 war der finale Datenschnitt der Studie 102 für August 2026 geplant; Ergebnisse hierzu liegen nicht vor.

Studie 107 (18 Behandelte): Ausgewertet wurden Krankenakten von Zentren, die Personen mit TK2d betreuen (erste Einwilligung: 22.09.2021; letzte Beobachtung: 21.01.2022). Laut Prüfplan sollten Angaben ab Symptombeginn erhoben werden; ein fester Erhebungszeitraum war nicht vorgegeben. Ab wann die Aufzeichnungen tatsächlich vorliegen, ist nicht angegeben und der Umfang der Daten hing ausdrücklich von der Verfügbarkeit der Akten ab (für die Aufnahme genügten Angaben zum Vitalstatus). Wie bei Studie 101 ist daher nicht berichtet, wann die erste berücksichtigte Datenerhebung bzw. Untersuchung stattgefunden hat und ob die Krankenakten nur

Untersuchungsaufzeichnungen oder auch anamnestische Angaben umfassten, die noch vor der ersten berücksichtigten Untersuchung lagen. Erhoben wurden Demographie, Krankheitsmerkmale, Exposition und Krankheitsverlauf; bei den Sicherheitsereignissen wurden nur tödliche SUE sowie UE erfasst, die zu Behandlungsabbruch, -unterbrechung oder Dosisreduktion führten. Die medizinische Vorgeschichte, also Vorerkrankungen und frühere Krankheitsereignisse vor Behandlungsbeginn, sowie Begleitmedikamente wurden nicht erhoben. Wie in Studie 101 wurden die erhobenen Daten im Monitoring anhand der Quelldokumente der Zentren geprüft.

Studie 114 (43 Behandelte): Ausgewertet wurden Aufnahme- und Nachsorgeformulare des globalen EAP 113. Ein Erhebungszeitraum ist nicht angegeben. Die Dokumentation beginnt mit der EAP-Aufnahme; Vorbehandlung, Diagnosebestätigung und Alter bei Symptombeginn stammen aus dem Aufnahmeformular, wurden also anamnestisch erhoben. Für die 9 auch in Studie 107 erfassten Personen wurden nur Angaben berichtet, die in Studie 107 nicht erhoben worden waren. Wann die erste berücksichtigte Aufnahme in das EAP stattgefunden hat und wie weit die anamnestischen Angaben vor der Erstaufnahme zurückreichen, ist nicht berichtet. Die Nachsorgeformulare wurden nicht zu festgelegten Zeitpunkten ausgefüllt; außer Exposition und Vitalstatus wurden keine Verlaufsdaten erhoben und Sicherheitsereignisse stammen aus der Sicherheitsdatenbank des Sponsors (verpflichtend SUE und Schwangerschaften, weitere UE freiwillig). Der 01.03.2024 ist in Studie 114 zugleich Einschlussstichtag und Datenschnitt: Eingeschlossen wurden Personen, die im globalen EAP 113 bis zu diesem Tag mit Nukleos(t)iden behandelt worden waren (dabei ist unklar, ob die Behandlung zu diesem Tag noch laufen musste), und Daten gingen bis zu diesem Tag ein. Ein Überleben bis zum Stichtag war keine Voraussetzung; unter den 43 Personen ist eine verstorbene Person, eine weitere beendete die Studie aufgrund eines medizinischen Ereignisses. Personen ohne laufende Dosisdokumentation und ohne dokumentierten Tod oder Behandlungsabbruch wären als Studienabbruch mit unbekanntem Grund gezählt worden; da dies auf keine Person zutraf, standen die übrigen 41 Personen zum Stichtag unter dokumentierter Behandlung. Der Stichtag legt damit weder einen gemeinsamen Beobachtungsbeginn noch eine einheitliche Beobachtungsdauer fest.

Behandlung und Exposition auf Studienebene

Die Fachinformation sieht für DC/DT (1:1) eine gewichtsbezogene, verträglichkeitsgesteuerte Titration vor: 130/130 mg/kg/Tag als Anfangsdosis, 260/260 mg/kg/Tag als Zwischendosis ab Tag 14 und maximal 400/400 mg/kg/Tag als Erhaltungsdosis ab Tag 28, jeweils in drei Dosen mit Nahrung; bei mäßiger oder schwerer Nierenfunktionsstörung langsamer [13].

In den ISE-Gesamtpool (N = 104) gingen alle Personen mit mindestens einer Dosis einer Pyrimidin-Nukleos(t)id-Formulierung ein, also auch Personen, die mit nicht GMP-konformem dC/dT oder dCMP/dTMP behandelt wurden; der pU begründet dies damit, dass alle Formulierungen dC und dT enthalten. In den Studien 101 und 107 war die Behandlung nicht durch ein Studienprotokoll vorgegeben, in Studie 114 nur durch die Vorgaben des EAP. Angaben etwa zu Titrationszeitpunkten, Einnahmehäufigkeit, Einnahme mit Nahrung oder Begleitmedikation liegen daher nur vor, soweit sie dokumentiert wurden; ob die Behandlung den Vorgaben der Fachinformation entsprach, ist nicht überprüfbar.

Hinweis: Die folgenden Angaben beziehen sich, soweit nicht anders angegeben, auf alle behandelten Personen der jeweiligen Studie einschließlich der Personen mit Symptombeginn im Alter von > 12 Jahren, da studienspezifische Angaben für die Personen mit Symptombeginn im Alter von ≤ 12 Jahren (Anwendungsgebiet) nicht durchgängig vorliegen.

Studie 101 (38 Behandelte): Behandelt wurde mit nicht GMP-konformen Nukleos(t)iden im CUP oder in prüfarztinitiierten Programmen. 32 der 38 Personen begannen mit dC/dT, 6 mit dCMP/dTMP, die nicht Bestandteil des zugelassenen Arzneimittels sind; 5 von ihnen wechselten während oder nach der Dosiseskalation auf dC/dT. Die Behandlungsdauer betrug für die 29 Personen mit Symptombeginn im Alter von ≤ 12 Jahren im Median 1,52 Jahre (Spannweite: 0,25–7,17). 22 dieser 29 Personen erreichten die Erhaltungsdosis von 400/400 mg/kg/Tag (dC/dT). Angaben zu Dosisunterbrechungen bzw. Dosiserhöhungen oder Dosisreduktionen liegen für diese Population nicht vor.

Studie 102 (47 vorbehandelte Personen): DC/DT wurde nach Protokoll verabreicht; alle 47 Personen standen bei Einschluss bereits unter laufender Behandlung mit Pyrimidin-Nukleos(t)iden (nicht GMP-konforme Formulierungen oder DC/DT). Bei Einschluss wurden nicht GMP-konform vorbehandelte Personen auf DC/DT umgestellt; eine bereits erreichte Erhaltungsdosis wurde fortgeführt, niedrigere Dosen wurden auf die nächstgelegene Dosisstufe gemäß Fachinformation übergeleitet und bei Verträglichkeit bis zur Erhaltungsdosis gesteigert. Für die 38 Personen mit Symptombeginn im Alter von ≤ 12 Jahren betrug die Behandlungsdauer in der Studie zum Datenschnitt vom 15.03.2024 im Median 52,7 Monate (Spannweite: 45,8–55,0) und zum Datenschnitt vom 14.03.2025 im Median 64,7 Monate (Spannweite: 49,3–66,8). 27 dieser 38 Personen (71,1 %) erhielten bereits an Studientag 1 die Erhaltungsdosis von 400 mg/kg/Tag DC/DT (Startdosis im Median 400,0 mg/kg/Tag (Spannweite: 130–400)). Die letzte dokumentierte Dosis lag zu beiden Datenschnitten im Median bei 400,0 mg/kg/Tag (Spannweite: 130–400 zum Datenschnitt vom 15.03.2024 bzw. 86–400 zum Datenschnitt vom 14.03.2025) und entsprach bei 30 Personen (78,9 %) bzw. 31 Personen (81,6 %) der Erhaltungsdosis. Dosiserhöhungen hatten 25 Personen (65,8 %) zum Datenschnitt vom 15.03.2024 bzw. 26 Personen (68,4 %) zum Datenschnitt vom 14.03.2025, Dosisunterbrechungen 20 (52,6 %) bzw. 21 (55,3 %) und Dosisreduktionen jeweils 11 (28,9 %); mehr als fünf Dosiserhöhungen hatten 9 bzw. 11 Personen, mehr als fünf Unterbrechungen 8 bzw. 9 Personen.

Studie 107 (18 Behandelte): Die mediane Behandlungsdauer aller 18 behandelten Personen betrug 0,6 Jahre (Spannweite: 0,1–9,1), die mediane Dosis 316,5 mg/kg/Tag je Wirkstoff (Spannweite: 100,0–600,0). Behandelt wurde mit dC/dT (13 Personen), dCMP/dTMP (4 Personen) und/oder DC/DT (7 Personen); durch Formulierungswechsel sind Mehrfachnennungen möglich. Bei 6 von 18 Personen (33,3 %) kam es zu Dosismodifikationen, ebenfalls 6 Personen (33,3 %) brachen die Behandlung ab, alle aufgrund von UE. Diese Angaben liegen nur für alle 18 behandelten Personen vor; für die Personen mit Symptombeginn im Alter von ≤ 12 Jahren werden sie nicht gesondert berichtet.

Studie 114 (43 Behandelte): Bis zum Stichtag (01.03.2024) betrug die mediane Behandlungsdauer aller 43 eingeschlossenen Personen 1,17 Jahre (Spannweite: 0,0–7,4). Bei einem Teilnehmer wurde eine maximale Behandlungsdauer von 7,4 Jahren verzeichnet. Die zuletzt erfasste Dosis lag zwischen 130 und 400 mg/kg/Tag je Wirkstoff. Die Mehrheit der Personen (34 Personen (79,1 %)) erhielt zuletzt eine Dosis von 400 mg/kg/Tag. Diese Angaben liegen nur für alle 43 eingeschlossenen Personen vor; für die Personen mit Symptombeginn im Alter von ≤ 12 Jahren (29 der 34 ausschließlich in Studie 114 erfassten Personen; siehe Tabelle 3) werden diese nicht gesondert berichtet.

Fazit

Zusammenfassend ist die Behandlung in den vier Studien nicht vergleichbar dokumentiert. Nur in Studie 102 wurde DC/DT während der Studie – jedoch nicht ab Behandlungsbeginn – nach Prüfplan verabreicht und prospektiv dokumentiert; die EMA bewertete diese Expositionsangaben als zuverlässiger als die retrospektiven Angaben aus Studie 101 [2]. Ob gemäß Fachinformation titriert und die Studienmedikation dreimal täglich mit Nahrung eingenommen wurde, ist für keine Studie vollständig sichergestellt. In Studie 102 wurde die Einnahme mit Nahrung erst seit Amendment 3 vom 29.04.2021 vorgeschrieben; ob sie umgesetzt wurde und wie die Einnahme zuvor erfolgte, ist

nicht dokumentiert. Für die Studien 107 und 114 liegen die Angaben zur Behandlung ausschließlich für alle behandelten Personen vor, nicht für die Personen mit Symptombeginn im Alter von ≤ 12 Jahren.

Angaben zu Behandlungsdauer und Exposition für die anwendungsgebietsbezogene ISE-Population (N = 82) und den ISE-Kernpool (N = 39) werden, soweit verfügbar, im Anschluss an die Bildung der Teilpopulationen im folgenden Abschnitt berichtet.

Auswahl und Bildung des ISE-Gesamtpools und der Teilpopulationen

Schritt 1 – Einschluss in die Einzelstudien: Gemeinsam war allen Studien eine genetisch bestätigte TK2d. Ein bestimmtes Alter bei Symptombeginn oder bestimmte Symptome waren nicht Voraussetzung, eine einheitliche Definition des Symptombeginns liegt nicht vor, und in den retrospektiven Studien wurden Symptome und Alter bei Symptombeginn übernommen, soweit dokumentiert. Die Studien unterscheiden sich in der Datenerhebung (nur Studie 102 prospektiv), den Einschlusskriterien und den Ausschlusskriterien. Einschlusskriterien waren in Studie 101 eine Behandlung mit dC/dT und/oder dCMP/dTMP, Krankenakten ab Symptombeginn und eine letzte Visite zwischen dem 01.06.2018 und dem 15.12.2018 (alle Kriterien kumulativ), in Studie 102 eine laufende Behandlung sowie für Personen ohne Teilnahme an Studie 101 ausreichende Vorbefunde und eine Freigabe durch den Sponsor, in Studie 107 verfügbare Krankenakten oder Angaben zum Vitalstatus und in Studie 114 bzw. im EAP mit TK2d vereinbare Zeichen oder Symptome, ein Risiko schwerer funktioneller Beeinträchtigung oder Tod und keine Möglichkeit zur Teilnahme an einer Studie des Sponsors. Ausgeschlossen waren in Studie 101 Personen mit anderen genetischen Erkrankungen, in Studie 102 zusätzlich Personen mit Lebererkrankungen (Leberwerte ≥ 2 × ULN) und mit Erkrankungen, die den Verlauf der TK2d beeinflussen können, in Studie 107 niemand und in Studie 114 bzw. im EAP unter anderem Personen mit anderen genetischen Erkrankungen, mit vorangegangenem Abbruch einer Nukleos(t)id-Behandlung aufgrund von UE, mit Lebererkrankung bzw. Leberwerten ≥ 3 × ULN ohne Freigabe durch den Sponsor oder mit dialysepflichtiger Niereninsuffizienz. Krankheitsschwere, Begleiterkrankungen und Vorbehandlung können sich daher zwischen den Studien unterscheiden. Wie sich diese Merkmale auf die Studien verteilen, ist nicht ausgewiesen; die klinische Ähnlichkeit der Studienpopulationen lässt sich somit nicht beurteilen.

Schritt 2 – ISE-Gesamtpool (N = 104): Die vier Studien umfassen zusammen 146 Studienteilnahmen behandelter Personen (Studie 101: 38; Studie 102: 47; Studie 107: 18; Studie 114: 43). 42 Studienteilnahmen entfielen auf Personen, die bereits in einer anderen Studie erfasst waren (35 in den Studien 101 und 102, 1 in den Studien 101 und 107, 6 behandelte Personen in den Studien 107 und 114) und wurden nur einmal gezählt; es ergeben sich 104 Personen: 38 aus Studie 101, zusätzlich 12 aus Studie 102, 17 aus Studie 107 und 37 aus Studie 114, darunter die 3 Personen, die in Studie 107 als unbehandelt geführt und erst im EAP behandelt wurden. Die im ISE-SAP festgelegten Mindestkriterien (bestätigte TK2-Mutation, keine andere genetische Erkrankung, mindestens eine Angabe zum Überlebensstatus, mindestens eine Dosis) erfüllten alle 104 Personen; der ISE-Gesamtpool entspricht damit allen behandelten Personen der vier Studien. Eine klinische Ähnlichkeit der Studienpopulationen stellen diese Kriterien nicht sicher; die in Schritt 1 beschriebenen Unterschiede der Ein- und Ausschlusskriterien bleiben bestehen. Für die 35 Personen aus den beiden Studien 101 und 102 stammen die Angaben aus aufeinanderfolgenden, nicht unmittelbar anschließenden Zeiträumen (letzte Visite in Studie 101 spätestens am 15.12.2018, erste Aufnahme in Studie 102 am 10.07.2019), für die Personen aus den Studien 107 und 114 überwiegend aus parallelen Erfassungen. Das Alter bei Symptombeginn spielt in diesem Schritt noch keine Rolle, die Eingrenzung auf das Anwendungsgebiet erfolgt erst in Schritt 3.

Schritt 3 – Anwendungsgebietsbezogene ISE-Population (N = 82): Den ISE-Gesamtpool (N = 104) wertet der pU differenziert nach dem Alter bei Symptombeginn aus. Die beiden Gruppen mit Symptombeginn im Alter von ≤ 2 Jahren (N = 56) und von > 2 bis ≤ 12 Jahren (N = 26) bilden zusammen die anwendungsgebietsbezogene ISE-Population (N = 82); die 22 Personen mit Symptombeginn im Alter von > 12 Jahren liegen außerhalb des Anwendungsgebiets. Das Alter bei Symptombeginn ist dabei das einzige Kriterium, das über die Aufnahmebedingungen der Einzelstudien (Schritt 1) hinaus angewandt wird; diese gelten für alle 82 Personen unverändert. Nach der Studie der Ersterfassung verteilen sich die 82 Personen rechnerisch wie folgt: 29 Personen stammen aus Studie 101 (davon gingen 28 anschließend in Studie 102 über; Studienbericht 102) und 10 wurden erstmals in Studie 102 erfasst – zusammen der ISE-Kernpool (N = 39, siehe unten) –; 11 wurden erstmals in Studie 107 erfasst (anwendungsgebietsbezogene ISS-Population (N = 50), abzüglich des Kernpools) und 32 ausschließlich in Studie 114 (82 abzüglich 50; darunter die 29 nur in Studie 114 erfassten Personen und die 3 in Studie 107 als unbehandelt geführten, erst im EAP behandelten Personen). 43 Personen stammen somit ausschließlich aus den Studien 107 und 114. Für Studie 107 weichen die Angaben zur Zahl der Personen mit Symptombeginn im Alter von ≤ 12 Jahren und zu deren Zuordnung zu den Altersgruppen zwischen Studienbericht, ISE und ISS voneinander ab; für 1 Person unterscheidet sich dadurch die Zuordnung zum Anwendungsgebiet (Details in der internen Dokumentation). Beim letzten Kontakt war die Behandlung bei 78 der 82 Personen (95,1 %) nicht beendet; 4 Personen (alle mit Symptombeginn im Alter von ≤ 2 Jahren; 3 wegen UE, 1 wegen Tod) hatten sie abgebrochen, keine davon aus dem ISE-Kernpool. Die Angabe bezieht sich auf den Behandlungsstatus beim letzten Kontakt; wie viele Personen der anwendungsgebietsbezogenen ISE-Population verstorben sind, geht daraus nicht hervor (Studie 107 allein dokumentiert zwei Todesfälle unter Behandlung bei Symptombeginn im Alter von ≤ 2 Jahren; siehe unten).

In der anwendungsgebietsbezogenen ISE-Population (N = 82) wurden die Morbiditätsendpunkte nicht für alle Personen erhoben. Für „Künstliche Beatmung“ und „Künstliche Ernährung“ weist der pU jeweils 29 Personen (35,4 %) aus, bei denen der Endpunkt in der jeweiligen Studie nicht erfasst wurde, bei „Künstliche Ernährung“ zusätzlich 1 Person ohne Daten; Angaben liegen damit zur künstlichen Beatmung für 53 Personen (64,6 %) und zur künstlichen Ernährung für 52 Personen (63,4 %) vor. Die 29 Personen entsprechen rechnerisch den Personen, die ausschließlich über die EAP-Formulare der Studie 114 erfasst wurden (32 erstmals in Studie 114 erfasste Personen abzüglich der 3 bereits in Studie 107 mit Krankenakte dokumentierten); für sie wurden außer dem Vitalstatus keine Wirksamkeitsdaten erhoben. Für „Motorische Meilensteine“ weist der pU keine entsprechende Angabe aus; die von ihm als „Personen mit vorliegenden Daten“ bezeichneten Personen sind diejenigen, die den jeweiligen Meilenstein vor Behandlungsbeginn erreicht hatten und ihn damit verlieren konnten: 49 Personen (59,8 %) für mindestens einen Meilenstein, je Meilenstein 22 (26,8 %; „Treppensteigen ohne Unterstützung“) bis 49 Personen (59,8 %; „Kopf aufrecht halten ohne Unterstützung“). Für ein Wiedererlangen nach Behandlungsbeginn kamen 40 Personen infrage (je Meilenstein 18 bis 24; vorheriger Verlust erforderlich), für einen Verlust nach Behandlungsbeginn 46 Personen (je Meilenstein 21 bis 46; vorheriges Erreichen erforderlich). Für den ISE-Kernpool (N = 39) weist der pU dagegen bei künstlicher Beatmung und künstlicher Ernährung keine Person als „nicht erhoben“ oder ohne Daten aus (jeweils 0 von 39) und bei den motorischen Meilensteinen keine entsprechende Kategorie; ob die Endpunkte tatsächlich bei allen 39 Personen erhoben wurden, lässt sich daraus nicht ableiten (siehe Kapitel 2.3.2). Die fehlenden Angaben der anwendungsgebietsbezogenen ISE-Population betreffen damit ausschließlich Personen aus den Studien 107 und 114 und sind durch die Studienherkunft bedingt – in Studie 114 wurden diese Endpunkte nicht erhoben –, nicht durch zufälliges Fehlen einzelner Werte; die gepoolten Morbiditätsergebnisse beruhen daher zu rund drei Vierteln (39 von 53 bzw. 52 Personen mit Angaben) auf dem ISE-Kernpool. Angaben zur medizinischen Vorgeschichte sowie zu Vor- und Begleitmedikation liegen nur für

die Studien 101 und 102 vor. Wie sich Krankheitsdauer, Krankheitsschwere, Funktionsstatus bei Behandlungsbeginn und Beobachtungsdauer auf die Studien verteilen, ist nicht ausgewiesen.

ISE-Kernpool gesamt (N = 50) und ISE-Kernpool (N = 39): Der ISE-Kernpool gesamt umfasst die 38 Personen aus Studie 101, von denen 35 in Studie 102 weiterbehandelt wurden, sowie die 12 Personen aus Studie 102 ohne vorherige Teilnahme an Studie 101. 39 davon hatten einen Symptombeginn im Alter von ≤ 12 Jahren und bilden den ISE-Kernpool (N = 39; 21 mit ≤ 2 Jahren, 18 mit > 2 bis ≤ 12Jahren): 28 Personen aus beiden Studien, eine nur aus Studie 101 und 10 nur aus Studie 102. Alle 39 Personen hatten damit die Aufnahmebedingungen der Studie 101 (Behandlung mit dC/dT und/oder dCMP/dTMP, letzte Visite zwischen dem 01.06.2018 und dem 15.12.2018) und/oder der Studie 102 (laufende, vertragene Behandlung bei Aufnahme) erfüllt, also bis zu diesen Zeitpunkten überlebt und die Behandlung fortgeführt. Sie sind Teil der anwendungsgebietsbezogenen ISE-Population (N = 82) und keine zusätzliche Population. Beim letzten Kontakt (Datenschnitt der Studie 102: 15.03.2024) war die Behandlung bei allen 39 Personen nicht beendet; die drei Behandlungsabbrüche unter den 50 Personen (zwei wegen UE, einer wegen Tod) betrafen Personen mit Symptombeginn im Alter von > 12 Jahren. Zum Behandlungsstatus zum Datenschnitt vom 14.03.2025 liegen keine Angaben vor. Ihr Alter bei Symptombeginn lag im Median bei 1,89 Jahren (Spannweite: 0,31–11,7), ihr Alter bei Behandlungsbeginn im Median bei 4,67 Jahren (Spannweite: 0,73–35,5); 9 Personen (23,1 %) wurden im Alter von ≤ 2 Jahren erstmals behandelt, 21 (53,8 %) im Alter von > 2 bis ≤ 12 Jahren, 3 (7,7 %) im Alter von > 12 bis < 18 Jahren und 6 (15,4 %) im Alter von ≥ 18 Jahren. Die Zeit von Symptombeginn bis zur Diagnose lag im Median bei 8,9 Monaten (Spannweite: −60 bis 360); einzelne Personen wurden somit vor dem dokumentierten Symptombeginn diagnostiziert (familiäre Fälle), ein präsymptomatischer Behandlungsbeginn ist für sie nicht auszuschließen. Eine Gegenüberstellung von Alter bei Symptombeginn und Alter bei Behandlungsbeginn je Person legt der pU nicht vor. Die Charakteristika der 39 Personen werden im Anhang dargestellt und in den Kapiteln 2.3.2 und 4.1 eingeordnet.

Behandlungsdauer und Exposition in den Teilpopulationen: Für die anwendungsgebietsbezogene ISE-Population (N = 82) liegen zum Datenschnitt vom 14.03.2025 nur über alle Formulierungen zusammengefasste Angaben vor. Die Dauer der Studienteilnahme (Einwilligung bis letzter Kontakt bzw. bei laufender Behandlung bis zur letzten Behandlung) betrug im Median 62,4 Monate (Spannweite: 0,0–177,7). Die Behandlungsdauer betrug im Median 61,9 Monate (Spannweite: 0,0–170,0); 2 Personen wurden weniger als einen Monat, 41 Personen (50,0 %) fünf Jahre oder länger behandelt. Zum Datenschnitt vom 15.03.2024 betrug sie im Median 54,8 Monate (Spannweite: 0,04–157,1), 33 Personen (40,2 %) waren fünf Jahre oder länger behandelt. Zur Dosis ist allein die durchschnittliche Tagesdosis ab Behandlungstag 28 ausgewiesen, die der pU als Dosis der Erhaltungsphase bezeichnet: Für 80 der 82 Personen mit einer Behandlungsdauer von mindestens 28 Tagen betrug sie im Median 394,6 mg/kg/Tag (Spannweite: 154,5–403,5). Die Erhaltungsphase ist damit allein über die Zeit seit Behandlungsbeginn definiert, nicht über das Erreichen der Erhaltungsdosis von 400/400 mg/kg/Tag; wie viele Personen diese erreichten, geht aus der Angabe nicht hervor. Die vom pU als Beobachtungsdauer bezeichnete Zeit vom Symptombeginn bis zum letzten Kontakt bzw. bei laufender Behandlung bis zur letzten Behandlung betrug im Median 106,1 Monate (Spannweite: 9,7–477,2). Sie umfasst damit auch die Zeit vor Behandlungsbeginn und vor Aufnahme in die Studien und entspricht nicht der Beobachtungszeit innerhalb der Studien. Für den ISE-Kernpool mit Symptombeginn im Alter von ≤ 12 Jahren (N = 39) liegen Angaben zur Behandlungsdauer nur zum Datenschnitt vom 15.03.2024 vor: Im Median 78,4 Monate (Spannweite: 53,2–157,1). Alle 39 Personen wurden mindestens zwei Jahre, 31 Personen (79,5 %) fünf Jahre oder länger behandelt. Weder die Verteilung der Formulierungen noch Titration oder der Anteil der Personen mit Erhaltungsdosis sind für die beiden Teilpopulationen ausgewiesen. Welche Behandlung die Personen der beiden bewerteten Populationen tatsächlich erhielten, lässt sich daher nur für den Anteil aus den Studien 101 und 102 näherungsweise

beschreiben. Für den Datenschnitt vom 14.03.2025, auf dem die Morbiditätsergebnisse des ISE-Kernpools (N = 39) beruhen, liegen keine Angaben zur Behandlungsdauer vor.

Zeitliche Vereinheitlichung der Studiendaten

Die Studien erhoben ihre Daten zu unterschiedlichen Zeiten, ab unterschiedlichen Bezugspunkten und in unterschiedlicher Dichte. Angaben zur Beobachtungsdauer liegen weder je Studie noch für die jeweilige Population im Anwendungsgebiet vor. Der ISE-SAP legt daher für alle Auswertungen, also auch für die anwendungsgebietsbezogene ISE-Population (N = 82) und den ISE-Kernpool (N = 39), einheitliche Regeln für Startzeitpunkt, Baseline, widersprüchliche Angaben aus mehreren Studien, Auswertungszeitpunkte und Datenschnitte fest. Diese Regeln harmonisieren die Auswertung, nicht jedoch die zugrunde liegende Datenerhebung.

Startzeitpunkt (ISE-Tag 1): Der ISE-SAP definiert als Tag 1 der gepoolten Auswertung die erste Gabe der Studienmedikation, bei Teilnahme an mehreren Studien das Datum der ersten Gabe unabhängig von der Formulierung. Dieser Tag wird im Folgenden zur Unterscheidung von Studientag 1 der Einzelstudien als ISE-Tag 1 bezeichnet. ISE-Tag 1 ist damit je Person die erste dokumentierte Gabe einer berücksichtigten Formulierung (dC/dT, dCMP/dTMP oder DC/DT), also der Behandlungsbeginn, nicht der Studieneintritt; das Datum stammt für die Studien 101 und 107 aus der retrospektiv dokumentierten Behandlungshistorie, für Studie 114 aus dem EAP-Aufnahmeformular. Für die 35 aus Studie 101 übergegangenen Personen liegt ISE-Tag 1 in Studie 101 und damit in einer Behandlung mit nicht GMP-konformen Formulierungen; für die 12 übrigen Personen der Studie 102 ist er nicht eindeutig festgelegt. Ob die erste dokumentierte Gabe tatsächlich die erste Behandlung war, lässt sich nicht prüfen; Kalenderdaten sind nicht ausgewiesen.

Aufnahme in die Studien relativ zu ISE-Tag 1: In den Studien 101 und 102 lag ISE-Tag 1 vor der Aufnahme in die Studie. Die Personen gehen erst ab der Studienaufnahme in die Kohorte ein, werden laut ISE-SAP jedoch ab ISE-Tag 1 als unter Risiko gezählt. Die Zeit zwischen ISE-Tag 1 und Studienaufnahme (Studie 102) bzw. letzter Visite (Studie 101) ist für die eingeschlossenen Personen zwar retrospektiv dokumentiert, ein Tod oder Behandlungsabbruch konnte in dieser Zeit jedoch definitionsgemäß nicht auftreten, weil die Aufnahme Überleben und laufende Behandlung voraussetzte. Diese Zeitspanne geht damit als ereignisfreie Zeit in die Auswertungen ein (Immortal Time Bias, siehe unten). Wie lange ISE-Tag 1 je Person vor der Studienaufnahme lag, ist nicht ausgewiesen.

Baseline: Der ISE-SAP definiert als Ausgangswert den letzten Wert innerhalb von sechs Monaten vor ISE-Tag 1; diese Regel betrifft gemessene Verlaufsvariablen, die für den ISE-Kernpool nicht bewertet werden. Für die drei bewerteten Morbiditätsendpunkte gilt als Ausgangsstatus („treatment baseline") – in den Ergebnistabellen als Status „zu Behandlungsbeginn" („at treatment start") bezeichnet – der zuletzt vor ISE-Tag 1 dokumentierte Status, der für den Tag des Behandlungsbeginns übernommen wird, und zwar ohne zeitliche Begrenzung (der ISE-SAP nimmt diese drei Endpunkte ausdrücklich vom 6-Monate-Fenster aus, der Status kann also auch deutlich länger als sechs Monate zurückliegen); er dient nicht als Bezugswert einer Veränderungsmessung, sondern legt fest, wer für Verlust bzw. Wiedererlangen eines Meilensteins infrage kommt und wer zu Behandlungsbeginn als beatmet bzw. sondenernährt gilt. Fehlt in Studie 101 ein Status vor Behandlungsbeginn, wird ersatzweise der erste Status nach Behandlungsbeginn verwendet; diese Regel wurde erst nach Abschluss der Datenerhebung festgelegt, ihre Häufigkeit ist nicht berichtet, und ein solcher Status kann bereits durch die Behandlung beeinflusst sein. Für die 12 Personen der Studie 102 ohne Teilnahme an Studie 101 gilt laut ISE-SAP als Ausgangsstatus der Status vor der ersten Gabe der Studienmedikation (ISE-Tag 1), worunter der ISE-SAP jede Behandlung mit Pyrimidin-Nukleos(t)iden fasst, auch die außerhalb der Studien begonnene Vorbehandlung. Da die Erhebung in Studie 102 erst unter laufender Behandlung einsetzte, dürfte ein solcher Status auf bei Studieneintritt retrospektiv erhobenen anamnestischen Angaben beruhen, und ob diese für die drei Endpunkte vorlag, ist nicht berichtet. Für Personen

ausschließlich aus Studie 114 liegt kein Ausgangsstatus vor. Ein einheitlich und vor Behandlungsbeginn erhobener Ausgangsstatus liegt damit für keinen der drei Endpunkte vor; je nachdem, aus welcher Studie eine Person stammt, beziehen sich Veränderungen daher auf unterschiedlich erhobene Ausgangswerte und sind untereinander nicht vergleichbar.

Umgang bei widersprüchlichen Angaben: Liegen Angaben aus mehreren Quellen vor, bei behandelten Personen vor allem für die 35 in den Studien 101 und 102 sowie die 9 in den Studien 107 und 114 erfassten Personen, hat laut ISE-SAP die vollständigste und aktuellste Angabe Vorrang. Eine umfassende Rangfolge für alle vier Studien liegt nicht vor, und wie häufig sich Angaben widersprachen, ist nicht berichtet.

Auswertungszeitpunkte: Für gemessene Verlaufsvariablen ordnet der ISE-SAP Beobachtungen festen Zeitpunkten relativ zu ISE-Tag 1 zu (Monate zu 28 Tagen mit Zeitfenstern; nach Monat 36 Zieltage im Abstand von 168 Tagen). Die im Anhang dargestellten Endpunkte werden dagegen nicht zu einzelnen Auswertungszeitpunkten, sondern über die gesamte Phase nach Behandlungsbeginn ausgewertet: Meilensteine anhand des letzten verfügbaren Status je Phase, Beatmung und Ernährung anhand des Status zu Behandlungsbeginn, zu einem beliebigen Zeitpunkt danach und am Ende der Beobachtung. Eine Begründung hierfür gibt der ISE-SAP nicht; für die retrospektiv erhobenen Daten der Studie 101 fehlen allerdings regelmäßige Erhebungszeitpunkte, die eine zeitpunktbezogene Auswertung erlauben würden. Die Beobachtungsdauer nach Behandlungsbeginn unterscheidet sich zwischen den Personen erheblich, und wie viele Personen je Zeitpunkt aus welcher Studie beitragen, ist nicht ausgewiesen. In der anwendungsgebietsbezogenen ISE-Population (N = 82) dürften späte Zeitpunkte überwiegend mit Personen des ISE-Kernpools (N = 39) besetzt sein, sodass Unterschiede zwischen frühen und späten Zeitpunkten auch auf eine veränderte Zusammensetzung zurückgehen können.

Datenschnitte: Die Daten jeder Studie gehen bis zu ihrem eigenen Stichtag ein: für Studie 101 bis zur letzten dokumentierten Visite (spätestens 15.12.2018), für Studie 107 bis zum 21.01.2022, für Studie 114 bis zum 01.03.2024 und für Studie 102 bis zum 14.03.2025 (Morbiditätsendpunkte der ISE) bzw. 15.03.2024 (ISS sowie Gesamtüberleben, Patientencharakteristika und Exposition des ISE-Kernpools). Zwischen frühestem und spätestem Stichtag liegen für die ISE gut sechs Jahre (15.12.2018 bis 14.03.2025) und für die ISS gut fünf Jahre (15.12.2018 bis 15.03.2024). Ob der Datenschnitt vom 15.03.2024 präspezifiziert war, ist nicht dargelegt; der Datenschnitt vom 14.03.2025 wurde von der EMA angefordert. Die studienübergreifenden Analysepläne entstanden erst 2021 (ISE-SAP vom 08.10.2021, ISS-SAP vom 20.09.2021; siehe Kapitel 2.4), also nach Abschluss der Datenerhebung der Studie 101 und mehr als zwei Jahre nach Beginn der Studie 102, in Kenntnis eines großen Teils ihrer Daten.

Fazit zur ISE

Eine Zusammenführung einarmiger Studien unterschiedlichen Designs setzt voraus, dass die Daten in den Studien vergleichbar erhoben wurden, unbehandelte Ausgangswerte vorliegen und die Regeln der Zusammenführung vor Kenntnis der Daten festgelegt wurden. Eine strenge klinische Homogenität der Studienpopulationen ist dagegen nicht erforderlich, da ohne Vergleichsgruppe keine Konfundierung eines Behandlungsvergleichs droht; die Studienpopulationen müssen jedoch klinisch hinreichend ähnlich und ihre Merkmale so beschrieben sein, dass sich beurteilen lässt, welche Personen die gepoolte Population abbildet. Unterschiedliche Ein- und Ausschlusskriterien der Einzelstudien führen dazu, dass bestimmte Merkmalskombinationen in der gepoolten Population nicht vorkommen können, und wenn sich die Studienpopulationen in relevanten Merkmalen deutlich unterscheiden, bildet der gepoolte Schätzer keine der beteiligten Populationen ab. Die Ergebnisse sollten daher ergänzend nach Studien getrennt dargestellt werden: Ein Zusammenlegen ohne Berücksichtigung der Studienzugehörigkeit kann Unterschiede zwischen den Studienpopulationen verdecken, sodass die gepoolte Population einheitlicher erscheint, als sie ist. Diese Voraussetzungen sind weder für die anwendungsgebietsbezogene ISE-Population (N = 82)

noch für den ISE-Kernpool (N = 39; Datenschnitt 14.03.2025) erfüllt. Die wesentlichen Limitationen lassen sich zwei Ebenen zuordnen:

- Einzelstudienebene: Drei der vier Studien sind retrospektive Auswertungen von Krankenakten bzw. EAP-Formularen ohne einheitlichen Visitenplan, deren Datenumfang von der Routinedokumentation abhängt; die einzige prospektive Studie 102 schloss nur bereits behandelte Personen unter laufender Behandlung ein, sodass auch dort kein unbehandelter Ausgangszustand vorliegt. Für keine der retrospektiven Studien sind Auswahl der Zentren und Patientenfluss bis zum Einschluss beschrieben, und in den Studien 101, 102 und 114 war der Einschluss an Bedingungen geknüpft, die erst nach Behandlungsbeginn erfüllt sein konnten.
- Ebene der Zusammenführung: Die Regeln des ISE-SAP harmonisieren Startzeitpunkt, Baseline, Auswertungszeitpunkte und Datenschnitte, nicht jedoch die zugrunde liegende Datenerhebung. Behandlung, Formulierung und Beobachtung unterscheiden sich zwischen den Studien, und die Abgrenzung des Anwendungsgebiets beruht auf einem nicht definierten Merkmal.

Auf beiden Ebenen ergeben sich vier Problemfelder, die im Folgenden erläutert werden: die Selektion der Studienpopulationen (Selektionsbias), der Immortal Time Bias, die Abgrenzung des Anwendungsgebiets über das Alter bei Symptombeginn sowie Unterschiede in Behandlung und Beobachtung.

Die erste Limitation betrifft den Selektionsbias. Es ist nicht beschrieben, welche EAP und CUP sowie welche Zentren in die retrospektiven Studien einbezogen wurden, nach welchen Kriterien dies erfolgte und ob Programme oder Zentren vorab ausgeschlossen wurden. Für keine der retrospektiven Studien liegt ein Patientenfluss bis zum Einschluss vor, also die Zahl der gesichteten Krankenakten sowie die Zahl der nicht eingeschlossenen Personen und die Gründe dafür. Damit ist nicht bekannt, welcher Anteil der insgesamt behandelten Personen mit TK2d erfasst wurde. Hinzu kommt, dass der Einschluss an Bedingungen geknüpft war, die erst nach Behandlungsbeginn erfüllt sein konnten: In Studie 101 war eine letzte dokumentierte Visite zwischen dem 01.06.2018 und dem 15.12.2018 erforderlich, in Studie 102 eine laufende Behandlung bei Aufnahme (erste Aufnahme am 10.07.2019), und zwar sowohl für die 35 aus Studie 101 übernommenen als auch für die 12 neu aufgenommenen Personen, und in Studie 114 die Aufnahme in das EAP, für bereits vorbehandelte Personen also das Überleben bis zur EAP-Aufnahme (ein Überleben bis zum Stichtag 01.03.2024 war dagegen nicht vorausgesetzt). Personen, die zwischen Behandlungsbeginn und diesen Zeitpunkten verstarben, die Behandlung abbrachen oder nicht mehr nachbeobachtet wurden, konnten nicht eingeschlossen werden. Die Studienpopulationen sind damit auf Personen selektiert, die bis zu diesen Zeitpunkten überlebt und die Behandlung fortgeführt haben. In Studie 102 kommt eine Selektion auf Verträglichkeit hinzu: Personen mit Lebererkrankung oder Leberwerten ≥ 2 × ULN waren ohne Freigabe durch den Sponsor ausgeschlossen, und die 2 Personen, die die Behandlung in Studie 101 wegen erhöhter Leberwerte abgebrochen hatten, konnten nicht übergehen. Dass frühe Todesfälle und Behandlungsabbrüche unter Behandlung auftreten, zeigt Studie 107, die auch verstorbene Personen einschloss: Unter den 18 behandelten Personen sind fünf Todesfälle dokumentiert, darunter 2 Personen mit Symptombeginn im Alter von ≤ 2 Jahren nach 23 bzw. 44 Behandlungstagen. Im ISE-Kernpool (N = 39) sind solche Verläufe durch die Einschlussbedingungen ausgeschlossen; zum Datenschnitt vom 15.03.2024 standen alle 39 Personen unter Behandlung. Wie viele Personen insgesamt behandelt wurden, aber vor den genannten Zeitpunkten verstarben oder die Behandlung abbrachen, ist nicht bekannt; das Ausmaß der Selektion lässt sich daher nicht quantifizieren. Da die Selektion durch die Einschlussbedingungen der Einzelstudien selbst entsteht, kann sie durch weitere Auswertungen der eingeschlossenen Personen nicht ausgeglichen werden.

Eng damit verbunden ist der Immortal Time Bias. ISE-Tag 1 ist die erste dokumentierte Gabe einer berücksichtigten Formulierung und liegt damit für einen Großteil der Personen vor der Aufnahme

in die jeweilige Studie: Für die Personen aus Studie 101 vor dem Zeitfenster der letzten Visite, für die 35 in Studie 102 übernommenen Personen in Studie 101, für Personen aus Studie 114 mit Behandlung vor der EAP-Aufnahme vor dieser Aufnahme; für die 12 Personen der Studie 102 ohne Teilnahme an Studie 101 ist ISE-Tag 1 nicht eindeutig festgelegt. Laut ISE-SAP wird die Überlebenszeit ab Behandlungsbeginn als Differenz zwischen dem Alter bei Tod bzw. beim letzten bekannten Lebenszeichen und dem Alter bei Behandlungsbeginn berechnet; ein verzögerter Eintritt in die Risikomenge („Delayed Entry") ab Studienaufnahme ist nicht vorgesehen. Die Zeit zwischen Behandlungsbeginn und letzter Visite (Studie 101) bzw. Studienaufnahme (Studie 102), in der die Personen definitionsgemäß überleben und weiterbehandelt werden mussten, geht damit als ereignisfreie Zeit unter Risiko in die Auswertung ein. Für die Auswertung ab Symptombeginn gilt dies zusätzlich für die Zeit vom Symptombeginn bis zum Behandlungsbeginn. Die Folgen sind eine systematische Überschätzung der Überlebenszeit und eine Unterschätzung der Ereignisraten, die umso stärker ausfallen, je länger diese Zeitspanne ist, und die sich insbesondere auf frühe Auswertungszeitpunkte auswirken, da die ereignisfreie Zeitspanne unmittelbar nach Behandlungsbeginn liegt. Für die Morbiditätsendpunkte bedeutet dies, dass Verläufe in dieser Zeitspanne nur retrospektiv und nur für die eingeschlossenen bzw. selektierten Personen vorliegen. Wie lange ISE-Tag 1 je Person vor der Studienaufnahme lag, ist nicht ausgewiesen; Kalenderdaten von Behandlungsbeginn und Studienaufnahme fehlen. Das Ausmaß der Verzerrung ist daher nicht abschätzbar. Maßnahmen zur Begrenzung des Immortal Time Bias sieht der ISE-SAP nur für den Vergleich mit unbehandelten Personen vor. Eine methodische Korrektur der Zeitachse, etwa durch einen verzögerten Eintritt in die Risikomenge, würde zudem nur die Zuordnung der Beobachtungszeit betreffen und die vorgelagerte Selektion der eingeschlossenen Personen nicht beheben.

Eine weitere Limitation betrifft die Abgrenzung des Anwendungsgebiets über das Alter bei Symptombeginn. Das zugelassene Anwendungsgebiet ist auf Personen mit einem Alter bei Symptombeginn bis zur Vollendung des 12. Lebensjahres beschränkt [13]; die anwendungsgebietsbezogene ISE-Population (N = 82) wie auch der ISE-Kernpool (N = 39) sind innerhalb des ISE-Gesamtpools bzw. des ISE-Kernpools gesamt ausschließlich über das Merkmal „Symptombeginn im Alter von ≤ 12 Jahren" abgegrenzt – die vorgelagerten Aufnahmebedingungen der Einzelstudien (Behandlung, Überleben bis zur maßgeblichen Visite der Studie 101 bzw. laufende Behandlung bei Aufnahme in Studie 102; siehe Schritt 1) bleiben davon unberührt –, ohne dass ausgewiesen wäre, ob darunter Personen mit Symptombeginn nach dem 12. Geburtstag sind (im ISE-Kernpool betrug das höchste Alter bei Symptombeginn 11,7 Jahre). Eine Definition des Symptombeginns liegt jedoch in keiner der vier Studien vor. In den Studien 101 und 107 wurde das Alter bei Symptombeginn retrospektiv aus den Krankenakten übernommen, in Studie 114 anamnestisch über das Aufnahmeformular erhoben, und für die 12 Personen der Studie 102 ohne Teilnahme an Studie 101 stammt es aus der retrospektiv erfassten Vorgeschichte. Welche Symptome als Erstmanifestation gewertet wurden, ob der Symptombeginn zeitnah dokumentiert oder erst Jahre später rückblickend angegeben wurde und ob er bei Personen, deren Diagnose genetisch vor dem Auftreten von Symptomen gestellt wurde, überhaupt eindeutig bestimmbar war, ist nicht berichtet. Dass alle Personen der anwendungsgebietsbezogenen ISE-Population und des ISE-Kernpools tatsächlich einen Symptombeginn im Alter von ≤ 12 Jahren hatten, lässt sich daher nicht sicherstellen; ebenso wenig ist für die Subgruppen mit Symptombeginn ≤ 2 Jahre und > 2 bis ≤ 12 Jahre gesichert, dass die Zuordnung zutrifft. Fehlzuordnungen sind in beide Richtungen nicht auszuschließen. Für die Abgrenzung der Zulassungspopulation ist vor allem relevant, dass Personen mit tatsächlich späterem Symptombeginn (etwa, wenn unspezifische Befunde rückblickend als Erstsymptom gewertet wurden) in die anwendungsgebietsbezogene Population eingegangen sein könnten; ein tatsächlich früherer Symptombeginn würde dagegen die Zuordnung zu den Alterssubgruppen, gegebenenfalls die Einordnung der Krankheitsschwere sowie die Berechnung der Überlebenszeiten betreffen. Da der Symptombeginn zugleich Startpunkt

der Überlebenszeitanalysen ab Symptombeginn ist, wirkt sich diese Unsicherheit auch unmittelbar auf die Berechnung der Überlebenszeit aus.

Darüber hinaus bestehen Unterschiede in Behandlung und Beobachtung. Zwischen der letzten Visite in Studie 101 (spätestens 15.12.2018) und der ersten Aufnahme in Studie 102 (10.07.2019) liegen mindestens knapp sieben Monate, deren Verlauf nur retrospektiv und teilweise erfasst ist. Formulierung, Dosierung und Behandlungsdauer unterscheiden sich zwischen den Studien erheblich, und ob gemäß Fachinformation titriert und eingenommen wurde, ist für keine Studie sichergestellt. ISE-Tag 1 und Baseline bezeichnen je nach Herkunft der Personen unterschiedliche Behandlungsphasen bzw. Zustände, die teilweise nach Behandlungsbeginn liegen; für die 12 Personen der Studie 102 ohne Teilnahme an Studie 101 liegt die Baseline unter laufender Behandlung. Für die Studien 107 und 114 liegen Angaben zu Behandlung und Exposition zudem nur für alle behandelten Personen vor, nicht für die Teilpopulation mit Symptombeginn im Alter von ≤ 12 Jahren; wie viele Personen aus diesen beiden Studien in die anwendungsgebietsbezogene ISE-Population (N = 82) eingingen, weist der pU nicht aus.

Zusammenfassend lassen sich die Ergebnisse der anwendungsgebietsbezogenen ISE-Population (N = 82) daher nicht als Ergebnisse einer vergleichbar ausgewählten, behandelten und beobachteten Kohorte interpretieren. Die Auswertungen der anwendungsgebietsbezogenen ISE-Population (N = 82) werden in der Nutzenbewertung nicht herangezogen.

Die Studien 101 und 102 werden von der EMA als Hauptstudien („Main Clinical Studies“) eingestuft [2]. Der ISE-Kernpool gesamt wurde erst mit Amendment 2 des ISE-SAP vom 03.05.2024 als zusätzliche Auswertungspopulation festgelegt. Im ISE-Kernpool mit Symptombeginn im Alter von ≤ 12 Jahren (N = 39) weist der pU für alle Personen Angaben zu allen drei Morbiditätsendpunkten aus (ob sie bei allen tatsächlich erhoben wurden, ist nicht sichergestellt; siehe Kapitel 2.3.2), alle wurden mindestens zwei Jahre behandelt, und die Heterogenität durch die Studien 107 und 114 entfällt. Die übrigen Limitationen gelten jedoch auch für den ISE-Kernpool (N = 39; Datenschnitt 14.03.2025): Selektionsbias und Immortal Time Bias (auch der Kernpool umfasst ausschließlich Personen, die bis zur maßgeblichen Visite der Studie 101 bzw. bis zur Aufnahme in Studie 102 überlebt hatten und behandelt wurden), die uneinheitliche Bestimmung von ISE-Tag 1 und Baseline und die Regeln für fehlende Angaben. Für die Personen der Studie 102 ohne vorherige Teilnahme an Studie 101 (insgesamt 12, davon 10 mit Symptombeginn im Alter von ≤ 12 Jahren) fehlen zudem Angaben zur Vorbehandlung und ein vor Behandlungsbeginn erhobener Ausgangswert. Die Ergebnisse des ISE-Kernpools unterliegen somit ebenfalls der beschriebenen Selektion und werden daher nicht herangezogen, jedoch deskriptiv im Anhang dargestellt.

2.2.3 Integrierte Sicherheitsauswertung (ISS)

Die ISS führt die Sicherheitsdaten behandelter Personen aus drei der vier Studien des ISE-Gesamtpools (N = 104) zusammen: Aus den retrospektiven, nicht interventionellen Krankenaktenauswertungen der Studien 101 und 107 sowie aus der prospektiven Phase-II-Studie 102. Nach Zusammenführung mehrfach erfasster Personen ergibt sich der ISS-Gesamtpool (N = 67). Die Daten aus Studie 114 werden separat in Modul 4 berichtet und sind nicht Teil des ISS-Gesamtpools; der pU begründet dies mit den Merkmalen des EAP und den unvollständigen Sicherheitsdaten (nur SUE und, soweit gemeldet, weitere UE). Wie die ISE ist die ISS keine eigenständige Studie, sondern eine nachträglich gebildete, studienübergreifende Zusammenführung individueller Personendaten.

Die nachfolgende Beschreibung folgt dem Aufbau der ISE mit einer Abweichung: Die Bildung des ISS-Gesamtpools (N = 67) und seiner Teilpopulationen wird vor Behandlung und Exposition beschrieben, da die Sicherheitsauswertungen je nach Variable auf unterschiedliche Teilpopulationen zurückgreifen und die Expositionsangaben nur vor diesem Hintergrund einzuordnen sind.

Die Beschreibung beschränkt sich auf die Punkte, in denen sich die ISS von der ISE unterscheidet oder die für die Sicherheitsauswertung zusätzlich von Bedeutung sind. Soweit die Angaben mit der ISE deckungsgleich sind, wird auf den entsprechenden Abschnitt der ISE verwiesen. Für die anwendungsgebietsbezogene Bewertung sind die Personen mit Symptombeginn im Alter von ≤ 12 Jahren maßgeblich; die Darstellung konzentriert sich daher auf diese Personen.

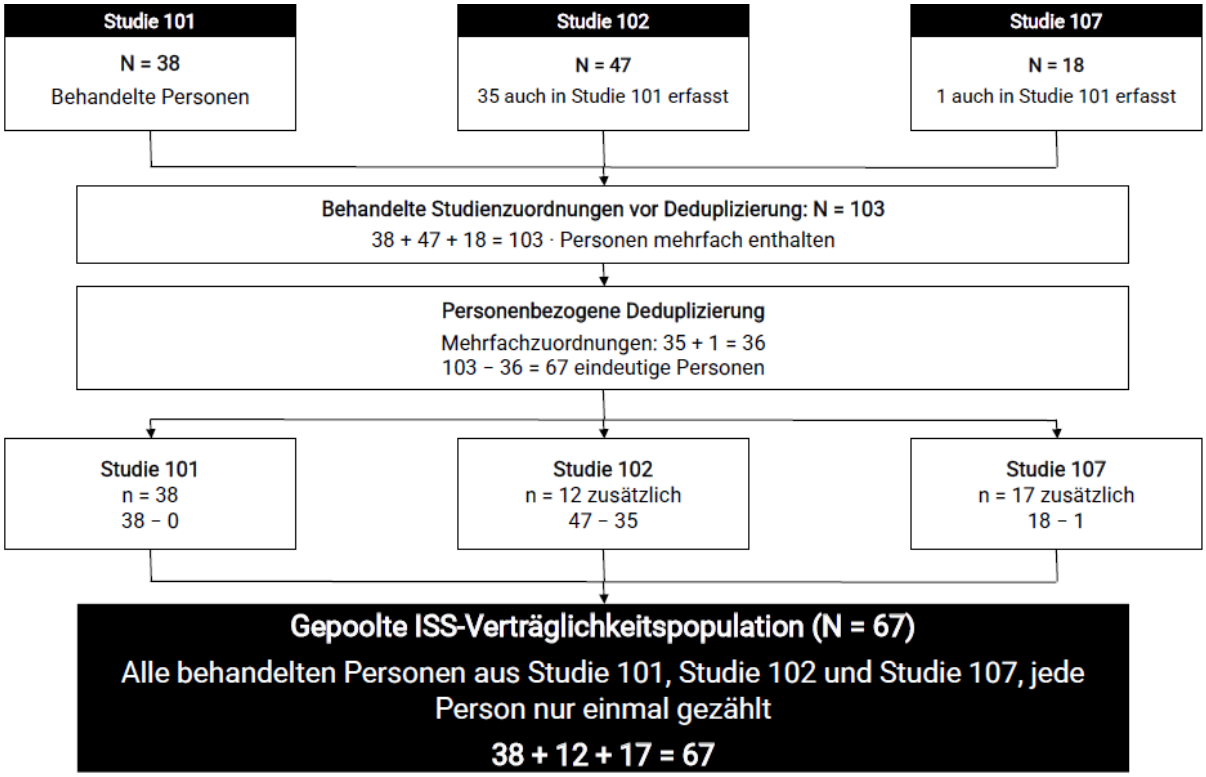

Abbildung 2: Patientenfluss der ISS (Eigene Darstellung) [10, 12]

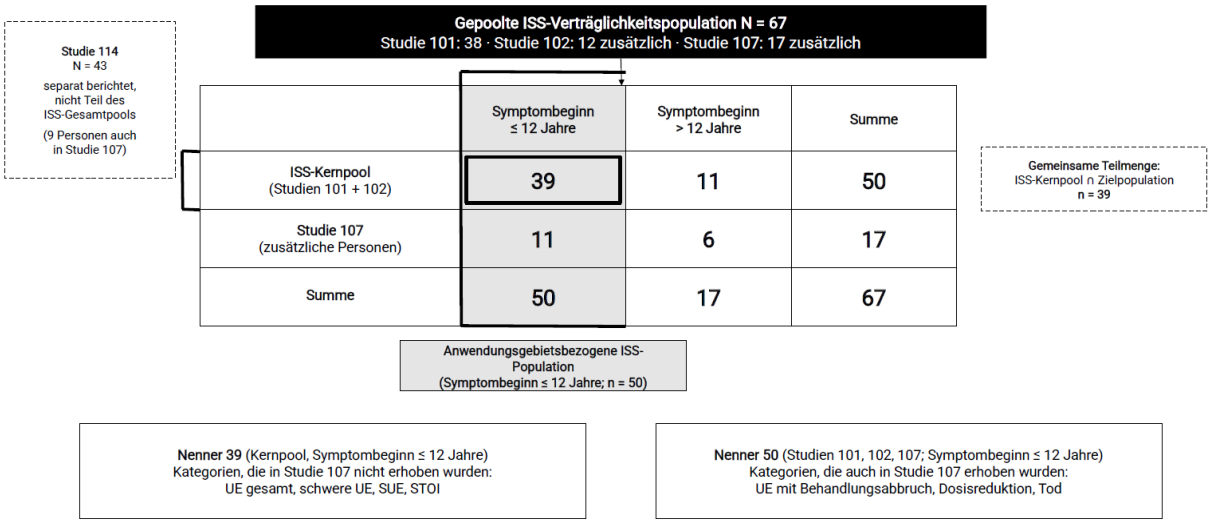

Abbildung 3: Teilpopulation der ISS und Nenner je UE-Kategorie (Eigene Darstellung) [10, 12]

Quellstudien und zugrunde liegende Studiendaten

Datengrundlage, Erhebungszeiträume und Datenschnitte der Studien 101, 102 und 107 entsprechen den Angaben zur ISE (siehe Kapitel 2.2.2); für Studie 102 gilt in der ISS der Datenschnitt vom 15.03.2024. Wesentlich ist der unterschiedliche Umfang der Sicherheitserhebung: Studie 102 erhob Sicherheitsdaten prospektiv nach Prüfplan, jedoch erst ab Studieneintritt und damit nach Behandlungsbeginn; Studie 101 erhob sie retrospektiv aus Krankenakten, wobei SUE nicht an den Sponsor gemeldet wurden, da die Studie nicht unter dem US-amerikanischen Prüfantrag („Investigational New Drug Application") des Sponsors durchgeführt wurde; Studie 107 erhob nur tödliche SUE sowie UE und SUE, die zu Behandlungsabbruch, -unterbrechung oder Dosisreduktion führten.

Die Sicherheitsangaben der Studie 114 stammen aus der Sicherheitsdatenbank des Sponsors: Die behandelnden Ärztinnen und Ärzte mussten SUE und Schwangerschaften melden, weitere UE wurden freiwillig gemeldet. Für die 9 auch in Studie 107 erfassten Personen wurden in Studie 114 nur Angaben berichtet, die in Studie 107 nicht erhoben worden waren. Modul 4 stellt die UE der Studie 114 für die 29 der 34 ausschließlich im EAP erfassten Personen mit Symptombeginn im Alter von ≤ 12 Jahren dar; die mediane Behandlungsdauer beträgt 9,9 Monate (Spannweite: 0,0–57,5), die mediane Dauer der Studienteilnahme (Einwilligung bis letzter Kontakt bzw. bei laufender Behandlung bis zur letzten Behandlung) 12,4 Monate (Spannweite: 0,0–58,2).

Auswahl und Bildung des ISS-Gesamtpools und der Teilpopulationen

Schritt 1 – Einschluss in die Einzelstudien: Es gelten die Angaben zur ISE. Für die Sicherheit ist die Auswahl in Studie 102 besonders relevant: Aufgenommen wurden nur Personen mit laufender Behandlung und stabilem Behandlungsregime; Personen mit Lebererkrankung in der Vorgeschichte oder Leberwerten ≥ 2 × ULN waren ohne Freigabe durch den Sponsor ausgeschlossen. Die 2 Personen, die die Behandlung in Studie 101 wegen erhöhter Leberwerte (GGT bzw. AST) abgebrochen hatten, konnten mangels laufender Behandlung nicht in Studie 102 übergehen; arzneimittelbezogene Leberstörungen sind daher potentiell unterrepräsentiert (Selektion auf Verträglichkeit). Da die Aufnahme in die Studien 101 und 102 Überleben und Behandlung mindestens bis zum 01.06.2018 (Studie 101) bzw. bis zur Aufnahme in Studie 102 (erste Aufnahme am 10.07.2019) voraussetzte, fehlen UE von Personen, die vorher verstarben oder die Behandlung abbrachen (Selektionsbias, Immortal Time Bias). Studie 107 erfasste zwar auch verstorbene Personen, erhob Sicherheitsereignisse aber nur eingeschränkt. Eine Darstellung aller behandelten Personen einschließlich früher Abbrüche liegt nicht vor.

Schritt 2 und 3 – ISS-Gesamtpool (N = 67) und Teilpopulationen: Die drei Studien umfassen 103 Studienteilnahmen behandelter Personen; nach Zusammenführung mehrfach erfasster Personen (35 in den Studien 101 und 102, eine in den Studien 101 und 107) ergeben sich 67 Personen (38 aus Studie 101, 12 zusätzliche aus Studie 102, 17 zusätzliche aus Studie 107). Davon hatten 50 einen Symptombeginn im Alter von ≤ 12 Jahren (anwendungsgebietsbezogene ISS-Population: 39 aus den Studien 101 und 102, 11 aus Studie 107) und 17 einen späteren (11 aus den Studien 101 und 102, 6 aus Studie 107). Da nur 50 der 67 Personen (74,6 %) dem Anwendungsgebiet entsprechen, wird der ISS-Gesamtpool nicht herangezogen. Für Variablen, die in Studie 107 nicht erhoben wurden, ist laut ISS-SAP der ISS-Kernpool gesamt (N = 50; Studien 101 und 102; Datenschnitt 15.03.2024) die Auswertungspopulation; er umfasst wie der ISE-Kernpool alle Personen der Studien 101 und 102 unabhängig vom Alter bei Symptombeginn. Maßgeblich ist der ISS-Kernpool (N = 39) mit Symptombeginn im Alter von ≤ 12 Jahren, der personengleich mit dem ISE-Kernpool ist und denselben vorgelagerten Aufnahmebedingungen (Behandlung, Überleben bis zur maßgeblichen Visite bzw. laufende Behandlung bei Studienaufnahme) unterliegt. Von den 9 in den Studien 107 und 114 erfassten Personen gehen die 6 in Studie 107 behandelten mit ihren Angaben aus Studie 107 in den ISS-Gesamtpool ein; die 3 in Studie 107 unbehandelten Personen sind nicht Teil des ISS-Gesamtpools.

Behandlung und Exposition

Formulierungen, Titration und Dosierungsangaben entsprechen der ISE; die Einschätzung der EMA zur Übertragbarkeit betrifft auch hier nur die Formulierung. Die ISS unterscheidet Behandlungsabschnitte mit nicht GMP-konformen Formulierungen und mit DC/DT; dieselbe Person wird nach einem Formulierungswechsel in mehreren Formulierungsgruppen gezählt, sodass diese Gruppen nicht addierbar und keine unabhängigen Behandlungsgruppen sind. Die nicht GMP-konforme Behandlung zwischen den Studien 101 und 102 wurde teilweise aus der Begleitmedikation der Studie 102 übernommen und als Studienmedikation gewertet. Für Studie 107 weisen die ISS-Unterlagen 12 Personen mit dC/dT und 3 mit dCMP/dTMP aus, die Unterlagen der Studie 107 13 bzw. 4.

Die ISS berichtet zwei Fälle mit Dosierungen oberhalb der in der Fachinformation empfohlenen Erhaltungsdosis von 400/400 mg/kg/Tag: 1 Person aus Studie 107, die anschließend in Studie 114 erfasst wurde, mit 600/600 mg/kg/Tag über 29,3 Monate in Studie 107 und zehn Tage in Studie 114 sowie 1 Person in Studie 102, die über bis zu 25 Monate 2–4 % mehr als die Erhaltungsdosis einnahm. Die ISS bezeichnet beide Fälle als versehentlich; bei der Person aus Studie 107 handelt es sich jedoch um eine anhaltende Dosierung oberhalb der Erhaltungsdosis, und ob vollständige Sicherheitsangaben dazu vorliegen, ist wegen der eingeschränkten Erhebung nicht beurteilbar.

Für den ISS-Kernpool (N = 39) betrug die Behandlungsdauer im Median 78,4 Monate (Spannweite: laut ISE-Unterlagen für dieselben 39 Personen 53,2–157,1; Monate zu rund 28 Tagen) alle 39 Personen wurden mindestens zwei Jahre, 31 (79,5 %) mindestens fünf Jahre behandelt. Ein Monat entspricht dabei 365,25/13 Tagen, also rund 28 Tagen. Die durchschnittlich erhaltene Tagesdosis betrug im Median 393,6 mg/kg/Tag (Spannweite: 169,6–400,0); sie bezieht sich laut ISS-Unterlagen auf jede Einzelkomponente.

Zeitliche Vereinheitlichung der Studiendaten

ISE-Tag 1, Baseline, Auswertungszeitpunkte, Datenschnitte und der Umgang mit unvollständigen Angaben folgen den Regeln der ISE (Baseline ist der letzte Wert vor der ersten Gabe der Studienmedikation „in der ersten Studie bzw. im EAP“ innerhalb von sechs Monaten vor Behandlungsbeginn; für den Beatmungs- und Sondenstatus gilt der letzte Wert vor Behandlungsbeginn ohne zeitliche Begrenzung; fehlt ein Wert vor Behandlungsbeginn, wird der erste Wert danach als Baseline verwendet und nicht mehr den Auswertungsfenstern nach Behandlungsbeginn zugeordnet). Abweichend gilt bei widersprüchlichen Angaben der Wert der zeitlich letzten Studie, sodass dieselbe Person in ISE und ISS mit unterschiedlichen Angaben eingehen kann. Bei zusammengeführten UE werden Ausgang und Enddatum aus der zeitlich letzten Studie übernommen.

Als behandlungsbegleitend gelten UE ab der ersten berücksichtigten Dosis bis 28 Tage nach der letzten Dosis (siehe Kapitel 2.3.4); bei laufender Behandlung gehen die Angaben bis zum Datenschnitt ein. Spätere Ereignisse sind nicht enthalten, so etwa drei der fünf Todesfälle in Studie 107. In der Einzelstudienauswertung der Studie 102 galten nur Ereignisse mit Beginn oder Zunahme des Schweregrads ab Studientag 1 als behandlungsbegleitend, sodass Einzelstudien- und ISS-Angaben nicht unmittelbar vergleichbar sind. Anamnestische Ereignisse ab der ersten Gabe wurden als UE übernommen; auch die Sicherheitsangaben der Studie 102 beruhen damit teilweise auf retrospektiver Erhebung.

Unvollständige Angaben zum Beginndatum eines UE wurden nach den Regeln des ISS-SAP ersetzt: Waren nur Monat und Jahr bekannt, wurde der erste Tag des Monats eingesetzt, entsprach der Monat jedoch dem Monat der ersten Dosis, deren Datum; war nur das Jahr bekannt, wurde der 1. Januar eingesetzt, im Jahr der ersten Dosis wiederum deren Datum. Ereignisse mit unvollständigem Beginndatum im Monat bzw. Jahr des Behandlungsbeginns gelten damit als behand-

lungsbegleitend. Fehlte das Behandlungsende, wurde es in Studie 101 auf den 31.12.2018 und in Studie 102 auf das Datum der letzten Beobachtung gesetzt; auf ISS-Ebene gilt das späteste Behandlungsende aus den Quellstudien. Exposition und Zeit unter Risiko können dadurch überschätzt und vor Behandlungsbeginn begonnene UE der Behandlungsphase zugeordnet werden. Wie häufig Angaben ersetzt wurden, ist nicht berichtet; Sensitivitätsanalysen liegen nicht vor.

Zusammenführung der Sicherheitsdaten

UE derselben Person aus verschiedenen Studien werden nur dann als ein Ereignis gezählt, wenn Beginn, MedDRA-Begriff, Intensität, Einstufung als schwerwiegend, CTCAE-Grad und Kausalitätsbewertung übereinstimmen; wie häufig dies erfolgte, ist nicht berichtet. UE wurden nach „Medical Dictionary for Regulatory Activities" (MedDRA, Version 26.0) kodiert und der Schweregrad mittels „Common Terminology Criteria for Adverse Events" (CTCAE, Version 5.0) eingestuft, jedoch nicht in allen Studien einheitlich (siehe Kapitel 2.3.4): In Studie 102 wurde der Schweregrad nach CTCAE oder nach einer an CTCAE angelehnten Skala des Prüfplans eingestuft; bei abweichenden Doppelbewertungen ging laut SAP-Addendum 1 vom 24.09.2021 die schwerere in die Auswertung ein. „Möglicherweise zusammenhängend" und „zusammenhängend" werden gemeinsam ausgewertet.

Für UE insgesamt, schwere UE und SUE sehen die ISS-Unterlagen den ISS-Kernpool gesamt als Auswertungspopulation vor; Studie 107 wird nur bei Behandlungsabbruch, Dosisreduktion und Tod einbezogen, UE mit Unterbrechung werden trotz Erhebung nicht gesondert mit Studie 107 ausgewertet. Nicht erhobene Ereignisse sind nicht mit Ereignisfreiheit gleichzusetzen. Modul 4 stellt die UE für 50 Personen dar; nach den ISS-Unterlagen ist dies die anwendungsgebietsbezogene ISS-Population (N = 50). Für den ISS-Kernpool mit Symptombeginn im Alter von ≤ 12 Jahren (N = 39) legt der pU in Modul 4 keine UE-Auswertungen vor; entsprechende Ergebnisse liegen in den Studienunterlagen zur ISS vor.

Fazit zur ISS

Die für die ISE beschriebenen Limitationen gelten für die ISS entsprechend, insbesondere die Selektion der Studienpopulationen und der Immortal Time Bias. Hinzu kommen die Selektion auf Verträglichkeit in Studie 102 (Ausschluss bei Lebererkrankung, kein Übergang nach Behandlungsabbruch), die auf tödliche und zu Behandlungskonsequenzen führende Ereignisse beschränkte Erhebung in Studie 107, die uneinheitliche Einstufung des Schweregrads sowie eine als Behandlungsdauer ausgewiesene Exposition, die unbekannte Behandlungsunterbrechungen und durch die letzte Beobachtung ersetzte fehlende Behandlungsenden einschließt, sodass die tatsächliche Exposition überschätzt werden kann; expositionsadjustierte UE-Raten liegen nicht vor. Studie 114 ist zudem nicht Teil des ISS-Gesamtpools, sodass ISE-Gesamtpool (N = 104) und ISS-Gesamtpool keine vergleichbare Basis für eine Abwägung von Nutzen und Schaden bilden.

Die Sicherheitsergebnisse der anwendungsgebietsbezogenen ISS-Population (N = 50) lassen sich daher nicht als Ergebnisse einer vergleichbar erhobenen und beobachteten Kohorte interpretieren. Die Auswertungen der anwendungsgebietsbezogenen ISS-Population (N = 50) werden in der Nutzenbewertung nicht herangezogen. Mit Ausnahme der eingeschränkten Erhebung in Studie 107 gelten die genannten Einschränkungen auch für den ISS-Kernpool (N = 39), in dessen Studien die Erfassung aller UE vorgesehen war. Die Ergebnisse des ISS-Kernpools mit Symptombeginn im Alter von ≤ 12 Jahren (N = 39; personengleich mit dem ISE-Kernpool; Datenschnitt der Studie 102 vom 15.03.2024) werden deskriptiv im Anhang dargestellt; die Einschätzung der Sicherheitsendpunkte in Kapitel 2.3.4 bezieht sich auf den ISS-Kernpool.

2.3 Endpunkte

In diesem Kapitel werden die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier, die zugehörigen Studienberichte, die Studienprotokolle, die SAP und die in diesen Quellen zitierte Literatur herangezogen. Alle Endpunkte der eingeschlossenen Studien (siehe Tabelle 3) wurden einer Prüfung unterzogen. Endpunkte, die in der nachfolgenden Tabelle nicht gelistet sind, wurden weder vom pU noch im Rahmen der Nutzenbewertung als patientenrelevant bzw. bewertungsrelevant eingestuft. Tabelle 5 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Die nachfolgende Darstellung beschränkt sich auf die Endpunkte, für die der pU in den studienübergreifenden Auswertungen ISE und ISS Ergebnisse vorgelegt hat. Endpunkte, die ausschließlich in Einzelstudien erhoben wurden und nicht Gegenstand der ISE oder ISS sind, werden nicht dargestellt. Die Einzelstudienunterlagen werden ausschließlich herangezogen, soweit sie für die Einordnung der Erhebung und Operationalisierung der in ISE bzw. ISS ausgewerteten Endpunkte erforderlich sind; eine Darstellung von Einzelstudienergebnissen erfolgt nicht. Tabelle 5 gibt wieder, ob die vom pU in Modul 4 vorgelegten Auswertungen der ISE und ISS in der Nutzenbewertung berücksichtigt werden; im Anhang dargestellte Auswertungspopulationen sind in den Fußnoten ausgewiesen. Der pU legt in Modul 4 Auswertungen für die anwendungsgebietsbezogene ISE-Population mit Symptombeginn im Alter von ≤ 12 Jahren (N = 82) und die anwendungsgebietsbezogene ISS-Population mit Symptombeginn im Alter von ≤ 12 Jahren (N = 50) vor, jeweils Teilpopulationen des ISE-Gesamtpools (N = 104) bzw. des ISS-Gesamtpools (N = 67); diese Auswertungen werden nicht herangezogen (siehe Kapitel 2.2.2 und 2.2.3). Die nachfolgende Endpunktbewertung bezieht sich ausschließlich auf den ISE-Kernpool mit Symptombeginn im Alter von ≤ 12 Jahren (Studien 101 und 102; N = 39), für die Sicherheit auf den personengleichen ISS-Kernpool (N = 39). Die Ergebnisse beider Kernpools werden im Anhang dargestellt. Für den ISE-Kernpool sind dies die in Modul 4 vorgelegten Auswertungen; für den ISS-Kernpool enthält Modul 4 keine Auswertung, die dargestellten Ergebnisse sind daher den ISS-Unterlagen des pU entnommen. Die Operationalisierung wird so beschrieben, wie sie für die ISE bzw. ISS festgelegt ist; sie gilt für die Kernpools unverändert. Die Bewertung der Operationalisierung und Validität bezieht sich ausschließlich auf den ISE-Kernpool bzw. ISS-Kernpool (Studien 101 und 102).

Über die vorgelegten Auswertungen zur Wirksamkeit in Modul 4 hinsichtlich „Motorische Meilensteine“, Künstliche Beatmung“ und „Künstlicher Ernährung“ hinaus wurden in den Einzelstudien weitere Endpunkte erhoben (siehe Tabelle 3). Für diese Endpunkte legt der pU in Modul 4 jedoch keine studienübergreifende Auswertung vor. Anhand der vorgelegten Unterlagen lässt sich nicht beurteilen, ob eine einheitliche Operationalisierung, ausreichende Standardisierung und eine für ein Pooling geeignete Datenbasis über mehrere Studien hinweg vorlagen. Ob eine ergebnisgesteuerte Auswahl der dargestellten Endpunkte vorliegt, kann damit nicht abschließend beurteilt werden.

Tabelle 5: Zusammenfassung der Endpunktbewertung der ISE und ISS

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
ISE (Studien 101, 102, 107 und 114)[1)]			
Gesamtüberleben[2)]	Mortalität	Ja	Nein[3)]
Motorische Meilensteine	Morbidität	Ja	Nein[4)]
Künstliche Beatmung		Ja	Nein[4)]
Künstliche Ernährung		Ja	Nein[4)]
ISS (Studien 101, 102 und 107)[5)]			
Unerwünschte Ereignisse	Sicherheit	Ja	Nein[6)]

[1)] Die Angaben zu „Gesamtüberleben", „Motorische Meilensteine", „Künstliche Beatmung" und „Künstliche Ernährung" beziehen sich auf die anwendungsgebietsbezogene ISE-Population (Personen mit Symptombeginn im Alter von ≤ 12 Jahren; N = 82).

[2)] Primärer Endpunkt der ISE.

[3)] Die Ergebnisse der anwendungsgebietsbezogenen ISE-Population (N = 82) werden nicht herangezogen. Anzahl und Anteil verstorbener Personen im ISE-Kernpool (N = 39) werden im Anhang dargestellt. Nähere Informationen hierzu siehe Kapitel 2.2.2 und 2.3.1.

[4)] Die vom pU vorgelegten Auswertungen der anwendungsgebietsbezogenen ISE-Population (N = 82) werden nicht herangezogen. Die Ergebnisse für den ISE-Kernpool (N = 39) werden im Anhang dargestellt. Nähere Informationen hierzu siehe Kapitel 2.3.2.

[5)] Die vom pU vorgelegten Angaben zu UE beziehen sich auf die anwendungsgebietsbezogene ISS-Population aus den Studien 101, 102 und 107 (Personen mit Symptombeginn im Alter von ≤ 12 Jahren; N = 50); Studie 114 ist nicht Bestandteil der ISS.

[6)] Die vom pU vorgelegten Auswertungen der anwendungsgebietsbezogenen ISS-Population (N = 50) werden nicht herangezogen. Die Ergebnisse für den ISS-Kernpool (N = 39) werden im Anhang dargestellt. Nähere Informationen hierzu siehe Kapitel 2.2.3 und 2.3.4.

Abkürzungen: ISE: Integrierte Wirksamkeitsanalyse; ISS: Integrierte Sicherheitsauswertung; pU: pharmazeutischer Unternehmer; UE: Unerwünschtes Ereignis.

2.3.1 Mortalität

Gesamtüberleben

Die vom pU vorgelegte Auswertung des Gesamtüberlebens in der anwendungsgebietsbezogenen ISE-Population (N = 82) wird wegen der in Kapitel 2.2.2 beschriebenen Limitationen der Datengrundlage – nicht wegen der fehlenden Eignung des Endpunkts – in der Nutzenbewertung nicht berücksichtigt. Die Endpunktbewertung bezieht sich auf den ISE-Kernpool (N = 39; Studien 101 und 102), dessen Ergebnisse im Anhang dargestellt werden.

Operationalisierung

Beschreibung

Der primäre Wirksamkeitsendpunkt der ISE-Auswertungen, also auch der anwendungsgebietsbezogenen ISE-Population (N = 82) und des ISE-Kernpools (N = 39), ist das Gesamtüberleben. Er umfasst drei Auswertungen; die ersten beiden sind als Zeit bis zum Tod jeglicher Ursache operationalisiert, die dritte als dichotomes Merkmal (ja / nein):

- Zeit bis zum Tod ab TK2d-Symptombeginn;
- Zeit bis zum Tod ab Behandlungsbeginn;
- Anzahl und Anteil verstorbener Personen im Nachbeobachtungszeitraum.

In den Ereigniszeitanalysen gilt der Tod jeglicher Ursache als Ereignis. Die ISE verwendet zwei unterschiedliche Startzeitpunkte, ab denen die Überlebenszeit berechnet wird:

- Ab Symptombeginn entspricht die Überlebenszeit der Differenz zwischen dem Alter bzw. Zeitpunkt bei Tod / letztem bekanntem Lebenszeichen und dem Alter bzw. Zeitpunkt bei TK2d-Symptombeginn.
- Ab Behandlungsbeginn entspricht die Überlebenszeit der Differenz zwischen dem Alter bzw. Zeitpunkt bei Tod / letztem bekanntem Lebenszeichen und dem Alter bzw. Zeitpunkt bei der ersten dokumentierten Behandlung (ISE-Tag 1).

Für beide Startzeitpunkte werden Personen ohne dokumentierten Tod zum Alter bzw. Zeitpunkt beim letzten bekannten Lebenszeichen zensiert.

Bewertung

Die Operationalisierung ist auf Auswertungsebene weitgehend nachvollziehbar: Ereignis (Tod jeglicher Ursache), Startzeitpunkte und Zensierung zum letzten bekannten Lebenszeichen sind einheitlich definiert. Für die in Studie 102 übernommenen Personen des ISE-Kernpools ist der Vitalstatus beim Übergang von Studie 101 in Studie 102 sichergestellt und wurde in Studie 102 prospektiv weiterverfolgt; in Studie 101 trat kein Todesfall auf.

Die Auswertungen der Zeit ab Symptombeginn und ab Behandlungsbeginn setzen voraus, dass Symptombeginn und Behandlungsbeginn valide bestimmt sind; dies ist insbesondere für den Symptombeginn nicht gesichert (siehe Kapitel 2.2.2). Diese Ereigniszeitanalysen und die historischen Überlebensvergleiche einschließlich der sequenziellen „Target Trial Emulation“ (TTE) werden in Kapitel 2.6 eingeordnet und im Anhang bewertet.

Patientenrelevanz

Der Endpunkt wird als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 der Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) angesehen.

Validität

Die Ausführungen zu Selektionsbias und Immortal Time Bias in Kapitel 2.2.2 sind zu berücksichtigen: In Studie 101 konnten nur Personen eingeschlossen werden, die bis zu einer Visite zwischen dem 01.06.2018 und dem 15.12.2018 überlebt hatten, in Studie 102 nur Personen unter laufender Behandlung mit Pyrimidin-Nukleos(t)iden. Wie viele Personen zuvor unter Behandlung verstarben oder nicht mehr nachbeobachtet wurden, ist nicht bekannt. Die Ergebnisse können daher nicht zur Nutzenbewertung herangezogen werden.

Innerhalb des ISE-Kernpools mit Symptombeginn ≤ 12 Jahren (N = 39; Datenschnitt 15.03.2024) ist die Erfassung des Vitalstatus weitgehend gesichert: 35 der 38 Personen aus Studie 101 gingen (unabhängig vom Alter zu Symptombeginn) in Studie 102 über und wurden dort prospektiv nachbeobachtet; für die 3 nicht in Studie 102 übernommenen Personen, davon nur 1 Person mit Symptombeginn im Alter von ≤ 12 Jahren, endet die Beobachtung mit der letzten Visite in Studie 101 und der Vitalstatus danach ist unklar (spätestens 15.12.2018; Zensierung, der weitere Vitalstatus ist unbekannt). Die Validität der Erhebung des Vitalstatus ist damit – abgesehen von dieser früh zensierten Person – nicht wesentlich eingeschränkt; maßgeblich bleiben Selektionsbias und Immortal Time Bias.

Die Ergebnisse des ISE-Kernpools werden im Anhang deskriptiv als Anzahl und Anteil verstorbener Personen dargestellt. Angaben zum Vitalstatus liegen nur zum Datenschnitt vom 15.03.2024 vor.

2.3.2 Morbidität

Funktionsbereich Motorik – Motorische Meilensteine

Die vom pU vorgelegte Auswertung zum Funktionsbereich Motorik in der anwendungsgebietsbezogenen ISE-Population (N = 82) wird wegen der in Kapitel 2.2.2 beschriebenen Limitationen der Datengrundlage – nicht wegen der fehlenden Eignung des Endpunkts – in der Nutzenbewertung nicht berücksichtigt. Die Endpunktbewertung bezieht sich auf den ISE-Kernpool (N = 39; Studien 101 und 102), dessen Ergebnisse im Anhang dargestellt werden.

Operationalisierung

Beschreibung

Für die ISE wurden neun motorische Meilensteine spezifiziert: 1) Kopf aufrecht halten ohne Unterstützung; 2) Aufrecht sitzen ohne Unterstützung; 3) Stehen mit Unterstützung; 4) Stehen ohne Unterstützung, 5) Gehen mit Unterstützung; 6) Gehen ohne Unterstützung; 7) Treppensteigen mit Unterstützung; 8) Treppensteigen ohne Unterstützung; 9) Laufen.

Je Meilenstein unterschied die ISE das erstmalige Erreichen, das Nicht-Erreichen, den Verlust und das Wiedererlangen. Der Netto-Score wurde über die neun Meilensteine getrennt für die Zeit vor Behandlungsbeginn (Symptombeginn bis erste Behandlung) und nach Behandlungsbeginn berechnet: Gewinn oder Wiedererlangen = +1; Verlust = −1; unveränderte oder nicht bestimmbare Veränderung = 0. Ein positiver Wert bedeutet mehr Gewinne bzw. Wiedererlangungen als Verluste. Der Netto-Score ist kein funktioneller Summenscore und kein Vergleich eines Ausgangsstatus mit einer späteren Messung.

Erreichen, Verlust und Wiedererlangen galten als Ereignis, wenn sie zu einem beliebigen Zeitpunkt der prä- bzw. posttherapeutischen Phase dokumentiert waren; maßgeblich war der zuletzt dokumentierte Status am Ende der jeweiligen Phase, nicht ein fester Auswertungszeitpunkt (Zeitbezüge siehe Kapitel 2.4). Zwischenzeitliche Erhebungen (planmäßige Visiten in Studie 102, Dokumentation in Studie 101) gingen nur als Beleg eines mit Datum oder Alter dokumentierten Ereignisses ein; zeitpunktbezogene Auswertungen oder Verläufe liegen nicht vor. Undatierte Ereignisse wurden anhand der im ISE-SAP festgelegten Reihenfolge der Meilensteine (von weniger zu stärker anspruchsvoll) aus datierten Ereignissen derselben Person abgeleitet: Ein nachweislich vor Behandlungsbeginn erreichter anspruchsvollerer Meilenstein galt auch für undatiert erreichte weniger anspruchsvolle Meilensteine, ein nachweislich vor Behandlungsbeginn verlorener weniger anspruchsvoller Meilenstein auch für undatiert verlorene anspruchsvollere Meilensteine (entsprechend für das Wiedererlangen). Bei Symptombeginn im Alter von mindestens 5 Jahren galt ein undatierter Meilenstein als vor Behandlungsbeginn erreicht. Fehlte nach einem Erreichen die Angabe zum Verlust bzw. nach einem Verlust die Angabe zum Wiedererlangen, galt der Meilenstein als nie verloren bzw. nie wiedererlangt.

Der Netto-Score wurde für alle Personen mit Angaben zu motorischen Meilensteinen ausgewertet. Verlust und Wiedererlangen wurden je Meilenstein auf ereignisspezifische Bezugsgruppen („at risk“) bezogen, die den Nenner der Anteile bilden:

- Für einen Verlust nach Behandlungsbeginn kam infrage, wer den Meilenstein vor Behandlungsbeginn erreicht und nicht verloren, ihn nach Behandlungsbeginn erstmals erreicht oder ihn vor bzw. nach Behandlungsbeginn wiedererlangt hatte;
- für ein Wiedererlangen nach Behandlungsbeginn, wer den Meilenstein vor Behandlungsbeginn verloren und nicht wiedererlangt oder ihn nach Behandlungsbeginn verloren hatte;
- vor Behandlungsbeginn kam für einen Verlust infrage, wer den Meilenstein erreicht hatte, und für ein Wiedererlangen, wer ihn erreicht und verloren hatte.

Erstmaliges Erreichen und Nicht-Erreichen wurden je Phase ebenfalls ausgewiesen und gehen mit Verlust und Wiedererlangen in den Netto-Score ein; die Ereignisanteile beziehen sich somit auf meilenstein- und ereignisspezifische Bezugsgruppen, nicht durchgehend auf alle Personen.

Bewertung

Die Operationalisierung ist auf Auswertungsebene nur teilweise nachvollziehbar: Meilensteine, Ereigniskategorien, gemeinsamer Zeitanker (ISE-Tag 1) und bezugsgruppenspezifische Auswertung sind einheitlich definiert. Nicht nachvollziehbar ist, ob die Fähigkeiten in den Studien 101 und 102 nach identischen Leistungskriterien und mit vergleichbarer Erhebungsdichte beurteilt wurden: Studie 101 erfasste Erreichen, Verlust und Wiedererlangen retrospektiv aus Krankenakten, Studie 102 im Rahmen der planmäßigen Funktionsuntersuchungen; Leistungskriterien für die Meilensteine wurden dort erst mit Amendment 3 vom 29.04.2021 aufgenommen, und laut Prüfplan waren Entwicklungsmeilensteine nur für Personen bis fünf Jahre vorgesehen. Da der 6-Monate-Zeiitraum der ISE-Baseline für Meilensteine nicht gilt, kann der Ausgangsstatus deutlich vor ISE-Tag 1 liegen; eine Begründung hierfür fehlt. Je nach Funktionsstatus bei Behandlungsbeginn sind Boden- und Deckeneffekte zu erwarten. Wer noch keinen Meilenstein erreicht hat, kann keinen verlieren, und wer alle erreicht hat, keinen hinzugewinnen. Für die 12 Personen der Studie 102 ohne Teilnahme an Studie 101 beschreibt die erste Erhebung einen Zustand unter laufender Behandlung, sodass für sie kein vor Behandlungsbeginn erhobener Ausgangsstatus vorliegt.

<u>Patientenrelevanz</u>

Der Endpunkt „Motorische Meilensteine“ wird als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen, weil er Erwerb, Erhalt, Verlust oder Wiedererwerb grundlegender körperlicher Funktionen und Aktivitäten abbildet. Die TK2d ist eine autosomal-rezessiv vererbte mitochondriale Erkrankung, die überwiegend als progrediente Myopathie mit axialer und proximaler Muskelschwäche in Erscheinung tritt und damit Muskelgruppen betrifft, die für Kopf- und Rumpfkontrolle, Sitzen, Stehen, Gehen und weitere motorische Aktivitäten erforderlich sind [1, 5, 7].

Für die Interpretation ist die Altersverteilung bei Symptom- und Behandlungsbeginn wesentlich; da der Symptombeginn in den Studienunterlagen weder definiert noch sein Erhebungsvorgehen beschrieben ist, ist die Validität dieser Angaben nicht bewertbar (siehe Kapitel 2.2.2). Einen Symptombeginn im Alter von ≤ 2 Jahren hatten 21 Personen (53,8 %), im Alter von > 2 bis ≤ 12 Jahren 18 Personen (46,2 %); die Population umfasst somit Kinder in noch nicht abgeschlossener motorischer Entwicklung ebenso wie Personen, deren bereits erworbene Fähigkeiten durch die progrediente Myopathie gefährdet waren. Bei Behandlungsbeginn waren dagegen nur 9 Personen (23,1 %) ≤ 2 Jahre alt, 21 (53,8 %) > 2 bis ≤ 12 Jahre und 9 (23,1 %) älter als 12 Jahre, davon 6 ≥ 18 Jahre. Eine Gegenüberstellung von Symptom- und Behandlungsbeginn je Person legt der pU nicht vor; ein präsymptomatischer Behandlungsbeginn nach genetischer Diagnose ist nicht auszuschließen (siehe Anhang, Abschnitt „Informationen zur Diagnose“).

Bei Symptombeginn im Alter von ≤ 2 Jahren ist insbesondere das Erreichen motorischer Meilensteine patientenrelevant: Die früh beginnende TK2d kann sich als muskuläre Hypotonie, motorische Entwicklungsverzögerung, motorische Regression und rasch progrediente generalisierte Muskelschwäche manifestieren, sodass der Erwerb einer zuvor nicht vorhandenen Fähigkeit einen unmittelbar wahrnehmbaren Gewinn an körperlicher Funktion und Mobilität beschreibt. Bei Symptombeginn im Alter von > 2 bis ≤ 12 Jahren ist insbesondere (aber nicht ausschließlich) der Erhalt vorhandener Fähigkeiten patientenrelevant: Der Verlust des selbstständigen Sitzens, Stehens, Gehens oder Treppensteigens bildet eine unmittelbar erlebbare Verschlechterung von Mobilität und Selbstständigkeit ab, das Wiedererlangen einer krankheitsbedingt verlorenen Fähigkeit entsprechend eine unmittelbare Verbesserung der körperlichen Funktion [1, 3].

Validität

Selektionsbias und Immortal Time Bias gelten wie beim Gesamtüberleben (siehe Kapitel 2.2.2); über die motorische Entwicklung nicht eingeschlossener Personen und deren Anzahl ist nichts bekannt. Die Ergebnisse können daher nicht zur Nutzenbewertung herangezogen werden.

Innerhalb des ISE-Kernpools (N = 39) ist die Validität eingeschränkt, weil nicht gesichert ist, dass Erreichen, Verlust und Wiedererlangen in den Studien 101 und 102 nach vergleichbaren Kriterien, regelmäßig und vollständig dokumentiert wurden: Leistungskriterien, Hilfsmittelbedingungen, Untersucherschulungen und Erhebungszeitpunkte sind nicht harmonisiert, und eine zentrale Beurteilung über beide Studien hinweg ist nicht nachvollziehbar. In Studie 101 hängen Vollständigkeit, Zeitpunkt und Beurteilungsgrundlage der Ereignisse von der Routinedokumentation ab; nicht dokumentierte Ereignisse konnten nicht erfasst werden. In Studie 102 ist für Personen über 5 Jahre und für Visiten vor Amendment 3, mit dem die Leistungskriterien ergänzt wurden, nicht sichergestellt, dass die Meilensteine systematisch und nach einheitlichen Kriterien erhoben wurden; die Leistungskriterien sollten zudem nach lokalen Verfahren erhoben werden. Die WHO-Referenz, mit der der pU in Modul 4 die Auswahl der Meilensteine belegt, beschreibt den Erwerb von sechs Meilensteinen bei gesunden Kindern, liefert keinen Maßstab für den Neuerwerb bei älteren Personen mit TK2d und stellt die einheitliche Beurteilung in den Studien 101 und 102 nicht sicher. Das Kriterium ≤ 12 Jahre bezieht sich auf den Symptombeginn, nicht auf das Alter bei Behandlungsbeginn oder Erhebung: Bei Behandlungsbeginn waren 30 der 39 Personen älter als 2 Jahre und 9 älter als 12 Jahre; wie viele zum jeweiligen Erhebungszeitpunkt älter als 5 Jahre waren, für welche Meilensteine verwertbare Angaben vorlagen und wie viele Meilensteine zu Behandlungsbeginn vorlagen, ist nicht nachvollziehbar. Da nur Ereignisse und der letzte Status je Phase ausgewertet wurden, bleibt zudem unbekannt, wann innerhalb der Beobachtung (siehe Kapitel 2.4) ein Meilenstein verloren oder wiedererlangt wurde und ob der Zustand bis zum Beobachtungsende anhielt.

Die Ersetzungsregeln (siehe Beschreibung) beeinflussen Ereigniszahlen und Bezugsgruppen: Nicht dokumentierte Verluste gehen nicht als Ereignis ein und können die Bezugsgruppen für einen Verlust verkleinern, bei einem dokumentierten Verlust galt das vorherige Erreichen als gegeben, und die zur Zuordnung undatierter Ereignisse verwendete Reihenfolge der Meilensteine ist keine klinische Schweregradskala und unterstellt eine im Einzelfall nicht zwingend zutreffende feste Abfolge. Wie häufig diese Regeln angewandt wurden und welchen Einfluss sie auf Bezugsgruppen und Ereigniszahlen hatten, ist nicht berichtet (siehe Kapitel 2.4).

Im Anhang werden für den ISE-Kernpool (N = 39) das Wiedererlangen und der Verlust von Meilensteinen nach Behandlungsbeginn dargestellt (Anzahl der jeweils infrage kommenden Personen, Anzahl wiedererlangter bzw. verlorener Meilensteine je Person sowie Anteile je Meilenstein). Nicht dargestellt werden die Auswertungen der Zeit vor Behandlungsbeginn, die ausschließlich auf retrospektiv erfassten Angaben beruhen und keinen belastbaren Vorher-Nachher-Vergleich erlauben (siehe Kapitel 2.4), sowie der Netto-Score und das erstmalige Erreichen von Meilensteinen nach Behandlungsbeginn, die der pU für den ISE-Kernpool zum Datenschnitt vom 14.03.2025 nicht ausweist; der für Personen mit Symptombeginn im Alter von ≤ 2 Jahren besonders relevante Neuerwerb von Fähigkeiten bleibt damit unberücksichtigt.

Funktionsbereich Atmung – Künstliche Beatmung

Die vom pU vorgelegte Auswertung zum Funktionsbereich Atmung in der anwendungsgebietsbezogenen ISE-Population (N = 82) wird wegen der in Kapitel 2.2.2 beschriebenen Limitationen der Datengrundlage – nicht wegen der fehlenden Eignung des Endpunkts – in der Nutzenbewertung nicht berücksichtigt. Die Endpunktbewertung bezieht sich auf den ISE-Kernpool (N = 39; Studien 101 und 102), dessen Ergebnisse im Anhang dargestellt werden.

Operationalisierung

Beschreibung

Der respiratorische Status wurde in der ISE über die Inanspruchnahme künstlicher Beatmung operationalisiert; Hustenassistenzgeräte galten nicht als künstliche Beatmung. Laut ISE-SAP wurde die Beatmungsform als nicht invasiv, invasiv ohne Tracheostoma oder invasiv mit Tracheostoma kategorisiert; in den ISE-Ergebnistabellen werden die beiden invasiven Formen zu einer Kategorie zusammengefasst. Laut ISE-SAP wurden Anzahl und Anteil der behandelten Personen in folgenden Kategorien dargestellt:

- Personen mit Beatmung zum Behandlungsbeginn: Wechsel der Beatmungsform nach Behandlungsbeginn (von nicht invasiver zu invasiver Beatmung bzw. umgekehrt), Beendigung der Beatmung oder Abnahme der täglichen Beatmungsstunden (kombiniert), Beendigung der Beatmung (gesondert), Fortbestehen der Beatmung bei Beobachtungsende sowie Abnahme bzw. Zunahme der täglichen Beatmungsstunden (insgesamt und um mindestens vier Stunden pro Tag).
- Personen ohne Beatmung zu Behandlungsbeginn: Beginn einer Beatmung nach Behandlungsbeginn einschließlich Beatmungsform sowie deren anschließende Beendigung.

Ergänzend wurden zwei vom Ausgangsstatus unabhängige Kategorien ausgewertet: Beatmung zu irgendeinem Zeitpunkt vor Behandlungsbeginn (anamnestische Angabe einschließlich einer bereits vor Behandlungsbeginn beendeten Beatmung; sie unterscheidet sich vom Ausgangsstatus nur bei einer vor Behandlungsbeginn beendeten Beatmung) sowie Beatmung zu beliebigem Zeitpunkt nach Behandlungsbeginn (fortbestehende und neu begonnene Beatmung zusammen; ausgewiesen werden Abnahme bzw. Zunahme der täglichen Beatmungsstunden, insgesamt und um mindestens vier Stunden pro Tag, sowie die Beendigung). Deskriptiv zusammengefasst wurden zudem die Dauer einer vor bzw. nach Behandlungsbeginn begonnenen Beatmung (in Tagen) und die täglichen Beatmungsstunden (jeweils letzte Angabe); für Personen mit Beatmung zu Behandlungsbeginn war eine Ereigniszeitanalyse der Zeit (in Monaten) von ISE-Tag 1 bis zur ersten Beendigung der Beatmung vorgesehen (Kaplan-Meier-Schätzung). Alle Kategorien sollten insgesamt und nach Alter bei Symptombeginn ausgewertet werden.

Die Zuordnung folgt den in Kapitel 2.4 beschriebenen Zeitbezügen: Beginn, Beendigung und Wechsel der Beatmungsform galten als Ereignis, wenn sie zu einem beliebigen Zeitpunkt nach Behandlungsbeginn dokumentiert waren (im Anhang „nach Behandlungsbeginn begonnen“, „eingestellt“ bzw. „umgestellt“ sowie „zu beliebigem Zeitpunkt nach Behandlungsbeginn“); das Fortbestehen der Beatmung bezeichnet den Status bei Beobachtungsende. Zu- und Abnahmen der täglichen Beatmungsstunden ergeben sich allein aus der Differenz zwischen der letzten Angabe nach und der letzten Angabe vor Behandlungsbeginn, ohne Vergleich zu festen Zeitpunkten oder Mittelung über den Verlauf; zwischenzeitlich dokumentierte Angaben (planmäßige Visiten in Studie 102, Routinedokumentation in Studie 101) gingen nur als Beleg eines Ereignisses ein.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar: ISE-Tag 1, Ausgangsstatus und einheitliche Kategorien harmonisieren die Auswertung formal. Nicht sichergestellt sind gleiche Kriterien für Beginn, Änderung und Beendigung der Beatmung sowie eine vergleichbare Erhebungsdichte: Studie 102 dokumentierte prospektiv zu festgelegten Visiten, Studie 101 retrospektiv nach Verfügbarkeit der Krankenakten. Die Beatmungskategorien der Studie 102 („nicht invasive Beatmung“,

„Tracheostomie/vollständige maschinelle Beatmung“) lassen sich nicht vollständig in die drei ISE-Kategorien überführen.

Die Schwelle einer Änderung um mindestens vier Stunden pro Tag entspricht der Responderdefinition des ursprünglichen SAP der Studie 101, ist für die TK2d nicht validiert und blieb nach Addendum 2 nur als deskriptive Kategorie erhalten.

Patientenrelevanz

Der Bedarf an künstlicher Beatmung kann eine krankheitsbedingte Einschränkung der selbstständigen Atmung abbilden; im vorliegenden Anwendungsgebiet ist die respiratorische Insuffizienz die führende Todesursache [3]. Patientenrelevant erscheint insbesondere die Vermeidung einer dauerhaften, bis zum Beobachtungsende fortbestehenden Beatmungsabhängigkeit. Die in der ISE ausgewerteten Kategorien bilden dies in unterschiedlichem Maße ab:

- Der Beginn einer Beatmung bei Personen ohne Beatmung zu Behandlungsbeginn bildet den Verlust der selbstständigen Atmung und einen neuen, in der Regel dauerhaften Unterstützungsbedarf ab und ist patientenrelevant, in besonderem Maße bei invasiver Beatmung mit Tracheostoma.
- Die Beendigung einer Beatmung (zu Behandlungsbeginn bestehend oder danach begonnen) bildet die Wiedererlangung der selbstständigen Atmung ab und ist patientenrelevant, sofern sie anhält; erfasst werden die Beendigung als Ereignis und das Fortbestehen bei Beobachtungsende, nicht die Dauerhaftigkeit einer Beendigung.
- Ein Wechsel zwischen nicht invasiver und invasiver Beatmung betrifft Eingriffe, Alltag und Pflegebedarf unmittelbar und ist patientenrelevant. Die ISE weist die Richtung des Wechsels aus (von nicht invasiv zu invasiv bzw. umgekehrt), legt jedoch keine Rangfolge der Beatmungsformen fest, fasst invasive Beatmung mit und ohne Tracheostoma zusammen und differenziert nicht invasive Verfahren (z. B. BiPAP, CPAP) nicht weiter, sodass das Ausmaß der Veränderung nur grob abgebildet ist.
- Eine Zu- oder Abnahme der täglichen Beatmungsstunden (insgesamt bzw. um mindestens vier Stunden) kann patientenrelevant sein, etwa beim Übergang von kontinuierlicher zu ausschließlich nächtlicher Beatmung; ihre Bedeutung hängt jedoch von Ausgangsdauer, Beatmungsform und tatsächlichem Unterstützungsbedarf ab, und eine für die TK2d validierte Relevanzschwelle liegt nicht vor (siehe Bewertung der Operationalisierung). Die Patientenrelevanz dieser Kategorie ist daher nicht beurteilbar.

Der Endpunkt „Künstliche Beatmung“ wird insgesamt als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen; dies gilt für Beginn und Beendigung einer Beatmung sowie für den Wechsel zwischen nicht invasiver und invasiver Beatmung. Diesbezügliche Ergebnisse werden für den ISE-Kernpool (N = 39) im Anhang dargestellt. Nicht beurteilbar ist die Patientenrelevanz der vorliegenden Ergebnisse zu Änderungen der täglichen Beatmungsstunden (siehe oben), infolgedessen werden die Ergebnisse nicht dargestellt.

Validität

Selektionsbias und Immortal Time Bias gelten wie beim Gesamtüberleben (siehe Kapitel 2.2.2 und Abschnitt „Gesamtüberleben“); über den respiratorischen Verlauf nicht eingeschlossener Personen und deren Anzahl ist nichts bekannt. Die Ergebnisse können daher nicht zur Nutzenbewertung herangezogen werden.

Innerhalb des ISE-Kernpools (N = 39) ist die Validität eingeschränkt, weil nicht sichergestellt ist, dass Vorliegen und Form einer künstlichen Beatmung in beiden Studien nach vergleichbaren Kriterien, in vergleichbarer Dichte und vollständig dokumentiert wurden (siehe Bewertung der Operationalisierung). Die Indikation zu Beginn, Änderung und Beendigung der Beatmung lag im ärztlichen Ermessen und war studienübergreifend nicht vereinheitlicht, sodass Statusänderungen neben dem Krankheitsverlauf auch lokale Versorgungspraxis, Dokumentationsdichte und individuelle Therapieentscheidungen widerspiegeln können. Fehlende oder unvollständige Start- und

Enddaten wurden nach den Regeln der Einzelstudien ersetzt (fehlendes Enddatum: Tag vor dem nächsten Beginndatum; dadurch entstandene Lücken oder Überlappungen als durchgehende Unterstützung gewertet), sodass tatsächliche Unterbrechungen oder Beendigungen unberücksichtigt bleiben können; wie häufig Ersetzungen erfolgten und wie sie Beginn, Ende oder Fortbestehen beeinflussten, ist nicht berichtet (siehe Kapitel 2.4). Der pU weist für alle 39 Personen Angaben zur Beatmung aus (keine Person als „nicht erhoben" oder ohne Daten); ob der Beatmungsstatus tatsächlich bei allen Personen erhoben oder fehlende Dokumentation – insbesondere in der retrospektiv aus Krankenakten erfassten Studie 101 – als „keine Beatmung" gewertet wurde, ist nicht sichergestellt. Beginn, Beendigung und Formwechsel beziehen sich zudem auf kleine, unterschiedlich definierte Teilgruppen; wann innerhalb der unterschiedlich langen Beobachtung (siehe Kapitel 2.4) ein Ereignis eintrat und wie lange ein Zustand anhielt, ist nicht erkennbar. Die gepoolte Auswertung erlaubt daher nur eine vorsichtige deskriptive Einordnung dokumentierter Beatmungsereignisse, keine belastbare Aussage zu einem einheitlich erfassten respiratorischen Verlauf.

Funktionsbereich Ernährung – Künstliche Ernährung

Die vom pU vorgelegte Auswertung zum Funktionsbereich Ernährung in der anwendungsgebietsbezogenen ISE-Population (N = 82) wird wegen der in Kapitel 2.2.2 beschriebenen Limitationen der Datengrundlage – nicht wegen der fehlenden Eignung des Endpunkts – in der Nutzenbewertung nicht berücksichtigt. Die Endpunktbewertung bezieht sich auf den ISE-Kernpool (N = 39; Studien 101 und 102), dessen Ergebnisse im Anhang dargestellt werden.

Operationalisierung

Beschreibung

Der Ernährungsstatus wurde in der ISE anhand der Nutzung einer nasogastralen oder gastrostomischen Sonde und des Umfangs der ergänzenden Sondenernährung operationalisiert. Laut ISE-SAP wurden Anzahl und Anteil der behandelten Personen in folgenden Kategorien dargestellt:

- Sonde vor Behandlungsbeginn, mit Grund der Einlage sowie Entfernung vor Behandlungsbeginn und deren Grund.
- Sonde zu einem beliebigen Zeitpunkt nach Behandlungsbeginn, mit Entfernung nach Behandlungsbeginn und deren Grund.
- Sonde zu Behandlungsbeginn, mit Entfernung nach Behandlungsbeginn und deren Grund sowie Fortbestehen bei Beobachtungsende.
- Keine Sonde zu Behandlungsbeginn, mit Einlage nach Behandlungsbeginn und deren Grund sowie anschließender Entfernung und deren Grund (im Unterschied zur zweiten Kategorie, die alle Personen mit Sonde nach Behandlungsbeginn unabhängig vom Ausgangsstatus umfasst).

Der ungefähre Anteil der täglichen Sondennahrung wurde vor Behandlungsbeginn, bei der letzten Visite nach Behandlungsbeginn für Personen mit Sonde zu Behandlungsbeginn sowie nach Behandlungsbeginn für Personen mit Sonde zu einem beliebigen Zeitpunkt dargestellt; die Gesamtdauer der Sondenernährung vor und nach Behandlungsbeginn (in Tagen) wurde deskriptiv zusammengefasst. Alle Kategorien sollten insgesamt und nach Alter bei Symptombeginn ausgewertet werden.

Die Zuordnung folgt den in Kapitel 2.4 beschriebenen Zeitbezügen: Einlage und Entfernung einer Sonde galten als Ereignis, wenn sie zu einem beliebigen Zeitpunkt nach Behandlungsbeginn dokumentiert waren (im Anhang „nach Behandlungsbeginn begonnen" bzw. „entfernt" sowie „zu beliebigem Zeitpunkt nach Behandlungsbeginn"), das Fortbestehen der Sonde bezeichnet den Status bei Beobachtungsende; der Anteil der Sondennahrung wurde als letzte Angabe vor Behandlungsbeginn und als Angabe bei der letzten Visite nach Behandlungsbeginn ausgewiesen, nicht als Verlauf oder Mittelwert; zwischenzeitlich dokumentierte Angaben gingen nur als Beleg eines Ereignisses ein.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar: Sondenstatus, Einlage, Entfernung, Fortbestehen, dokumentierte Gründe, Anteil der Sondennahrung und Unterstützungsdauer werden nach einer gemeinsamen Logik zusammengefasst. Dies vereinheitlicht Zuordnung und Darstellung, nicht aber Indikation, Dokumentationsdichte und Abstand des Ausgangsstatus zu ISE-Tag 1; die Erhebung war wie beim Beatmungsstatus nicht einheitlich (siehe Abschnitt „Funktionsbereich Atmung – Künstliche Beatmung“).

Der Anteil der Sondennahrung wird für unterschiedliche Personengruppen und Bezugspunkte berichtet und erlaubt keinen einheitlichen Vorher-Nachher-Vergleich; Gründe für die Entfernung einer Sonde sind nicht für alle Personen angegeben, und ohne Grund ist eine Einlage oder Entfernung nicht eindeutig als Verschlechterung oder Verbesserung zu interpretieren.

Patientenrelevanz

Der Funktionsbereich Ernährung betrifft die Fähigkeit, Nahrung ausreichend und möglichst oral aufzunehmen. Der Bedarf an einer nasogastralen oder gastrostomischen Sonde kann eine krankheitsbedingte Einschränkung der oralen Nahrungsaufnahme abbilden, schließt eine selbstständige Nahrungsaufnahme jedoch nicht aus; inwieweit diese durch weitere Beeinträchtigungen der TK2d (z. B. der Motorik) erschwert wird, ist nicht beurteilbar. Da Schluckstörungen und eine unzureichende orale Nahrungsaufnahme im vorliegenden Anwendungsgebiet den Krankheitsverlauf prägen können, erscheint insbesondere die Vermeidung einer dauerhaften Sondenernährung patientenrelevant. Die in der ISE ausgewerteten Kategorien bilden dies in unterschiedlichem Maße ab:

- Die Einlage einer Sonde bei Personen ohne Sonde zu Behandlungsbeginn bildet den Verlust einer ausreichenden oralen Nahrungsaufnahme und einen neuen, meist dauerhaften Aufwand und gegebenenfalls externen Unterstützungsbedarf ab und ist patientenrelevant; maßgeblich für die Einordnung ist der dokumentierte Grund der Einlage (Dysphagie; Ergänzung der oralen Nahrungsaufnahme; andere), da eine Sonde auch vorsorglich zur Sicherung der Ernährung eingelegt werden kann.
- Die Entfernung einer Sonde (zu Behandlungsbeginn bestehend oder danach eingelegt) bildet die Unabhängigkeit von künstlicher Ernährung ab und ist patientenrelevant, wenn sie auf eine verbesserte Schluckfunktion oder orale Nahrungsaufnahme zurückgeht; erfasst wird der Grund der Entfernung (Verbesserung bezüglich Dysphagie bzw. oraler Nahrungsaufnahme; sondenbezogenes UE; andere), nicht jedoch, ob die orale Ernährung anschließend dauerhaft ausreichte. Das Fortbestehen der Sonde bei Beobachtungsende beschreibt den Zustand, nicht dessen Verlauf.
- Der Anteil der Sondennahrung an der täglichen Nahrungsaufnahme (Stufen von 25 %) kann eine Zu- oder Abnahme der Abhängigkeit von künstlicher Ernährung anzeigen; er ist jedoch grob gestuft, wird für unterschiedliche Personengruppen und Bezugspunkte berichtet und ist ohne Angaben zum Ernährungszustand nur eingeschränkt interpretierbar, sodass die Patientenrelevanz einer Veränderung nicht beurteilbar ist. Die Gesamtdauer der Sondenernährung beschreibt den Zustand ohne eigenständige Aussage zur Veränderung.

Der Endpunkt „Künstliche Ernährung“ wird insgesamt als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen; dies gilt für Einlage und Entfernung einer Sonde unter Berücksichtigung der dokumentierten Gründe, nicht beurteilbar ist die Patientenrelevanz des Anteils der Sondennahrung. Im Anhang werden daher für den ISE-Kernpool (N = 39) die Sonde zu Behandlungsbeginn (mit Entfernung nach Behandlungsbeginn einschließlich Gründen und Fortbestehen bei Beobachtungsende), die Sonde zu beliebigem Zeitpunkt nach Behandlungsbeginn (mit Entfernung einschließlich Gründen) sowie die Einlage einer Sonde nach Behandlungsbeginn (mit Gründen und anschließender Entfernung einschließlich Gründen) dargestellt.

Validität

Selektionsbias und Immortal Time Bias gelten wie beim Gesamtüberleben (siehe Kapitel 2.2.2 und Abschnitt „Gesamtüberleben“); über den Ernährungsverlauf nicht eingeschlossener Personen und deren Anzahl ist nichts bekannt. Die Ergebnisse können daher nicht zur Nutzenbewertung herangezogen werden.

Innerhalb des ISE-Kernpools (N = 39) ist die Validität eingeschränkt, weil nicht gesichert ist, dass Sondenstatus und Sondenereignisse in beiden Studien vergleichbar und vollständig dokumentiert wurden: Die Indikationen für Einlage, Wechsel oder Entfernung und die Dokumentation ihrer Gründe sind nicht harmonisiert, die Ernährungs- und Schluckfunktion selbst wird nur indirekt abgebildet. Studie 101 übernahm die Angaben (soweit verfügbar Beginn und Ende der Sondennutzung, Sondentyp und Anteil der Sondennahrung) retrospektiv aus Krankenakten, wobei Entscheidung und Management ohne standardisierte Vorgabe im ärztlichen Ermessen lagen; Studie 102 dokumentierte prospektiv, aber individuell angepasst, und wie der dort vorgesehene Netto-Ernährungsscore in die ISE-Kategorien überführt wurde, ist nicht nachvollziehbar. Für fehlende oder unvollständige Start- und Enddaten galten dieselben Ersetzungsregeln wie für die Beatmung, einschließlich der Annahme einer durchgehenden Unterstützung; wie häufig sie angewandt wurden und wie sie Einlage, Entfernung oder Fortbestehen beeinflussten, ist nicht berichtet (siehe Kapitel 2.4).

Der pU weist für alle 39 Personen Angaben zum Sondenstatus aus (keine Person als „nicht erhoben“ oder ohne Daten); ob der Sondenstatus tatsächlich bei allen Personen erhoben oder fehlende Dokumentation – insbesondere in der retrospektiv aus Krankenakten erfassten Studie 101 – als „keine Sonde“ gewertet wurde, ist nicht sichergestellt. Einlage und Entfernung beziehen sich auf kleine, unterschiedliche Teilgruppen, Zeitpunkt und Dauer eines Ereignisses innerhalb der Beobachtung sind nicht erkennbar, und einheitlich erhobene Angaben zu Schluckfähigkeit, oraler Nahrungsaufnahme, Ernährungszustand und Gründen der Versorgungsentscheidung liegen nicht vor. Die gepoolte Auswertung erlaubt daher nur eine vorsichtige Beschreibung dokumentierter Sondenereignisse, keine belastbare Aussage zu einem einheitlich erfassten Ernährungsverlauf.

2.3.3 Lebensqualität

Für die Lebensqualität legt der pU keine studienübergreifende Auswertung vor (zu den in den Einzelstudien erhobenen Instrumenten siehe Tabelle 3).

2.3.4 Sicherheit

Unerwünschte Ereignisse

Die Auswertung der anwendungsgebietsbezogenen ISS-Population (N = 50) wird aus den in Kapitel 2.2.3 beschriebenen Gründen nicht berücksichtigt. Für den ISS-Kernpool (N = 39; Studien 101 und 102) legt der pU in Modul 4 keine UE-Auswertungen vor; die Endpunktbewertung bezieht sich auf diesen Kernpool, dessen den ISS-Unterlagen entnommene Ergebnisse im Anhang dargestellt werden.

Operationalisierung

Beschreibung

Die ISS wertet behandlungsbegleitende unerwünschte Ereignisse (UE) deskriptiv und ohne inferenzstatistische Tests aus. Ein UE ist jedes ungünstige medizinische Ereignis, einschließlich eines auffälligen Laborbefunds, eines Symptoms oder einer neu aufgetretenen bzw. verschlechterten Erkrankung. Als behandlungsbegleitend gilt ein UE mit Beginn am oder nach der ersten Dosis einer berücksichtigten Nukleos(t)id-Behandlung in der ersten Studie oder im globalen EAP 113 und bis 28 Tage nach der letzten Dosis in der letzten Studie; Behandlungspausen bleiben unberücksichtigt. Ereignisse vor der ersten Dosis werden als medizinische Vorgeschichte geführt

(erhoben nur in den Studien 101 und 102); als Vorgeschichte dokumentierte Ereignisse mit Beginn ab der ersten Dosis werden als UE ausgewertet. UE und Vorgeschichte werden nach MedDRA (Version 24.0 oder höher, in den Ergebnissen Version 26.0) als Systemorganklasse (SOC) und „Preferred Term" (PT) kodiert. Stimmen Beginn, PT, CTCAE-Grad, Intensität, Einstufung als schwerwiegend und Kausalitätsbewertung desselben Ereignisses in mehreren Studien überein, werden die Einträge zu einem UE zusammengeführt; Ausgang und Enddatum stammen dann aus der späteren Studie. Andernfalls werden die Einträge als getrennte UE gezählt. Bei sonstigen widersprüchlichen Angaben zwischen Studien gilt die Angabe der jüngsten Studie, sofern sie dort nicht fehlt.

Der Schweregrad wird nach CTCAE (Version 5.0) bestimmt; schwere UE sind UE mit einem CTCAE-Grad ≥ 3. Die ISS-Studienunterlagen unterscheiden dabei zwei Angaben: den CTCAE-Grad (dort als „toxicity grade" bezeichnet) und die Intensität des Ereignisses (leicht, mittel, schwer, lebensbedrohlich, tödlich; dort als „severity" bezeichnet). Beide Angaben werden über eine feste Zuordnung ineinander überführt (von Grad 1 = leicht bis Grad 5 = tödlich): Fehlt eine der beiden Angaben, wird sie aus der anderen abgeleitet; widersprechen sie sich, wird für beide die schwerere Einstufung übernommen; ist der CTCAE-Grad als unbekannt dokumentiert, gilt auch die Intensität als unbekannt.

Ein schwerwiegendes unerwünschte Ereignis (SUE) ist ein UE, das zum Tod, zu einer lebensbedrohlichen Situation, zu einer stationären Aufnahme oder Verlängerung einer Hospitalisierung, zu einer anhaltenden oder erheblichen Beeinträchtigung normaler Lebensfunktionen oder zu einer kongenitalen Anomalie führt. Auch ein anderes medizinisch bedeutsames Ereignis kann als SUE eingestuft werden, wenn es die Person gefährden und eine medizinische oder chirurgische Intervention zur Vermeidung einer dieser Folgen erfordern kann.

Gezählt werden Personen mit mindestens einem Ereignis je Kategorie; bei wiederholten Ereignissen geht das schwerste bzw. das mit dem engsten Zusammenhang zur Behandlung ein. Laut ISS-SAP vorgesehen sind folgende Kategorien:

- UE gesamt und SUE nach SOC und PT, jeweils auch nach höchstem CTCAE-Grad;
- UE und SUE mit Therapieabbruch oder Dosisreduktion sowie SUE mit Todesfolge;
- UESI: nicht infektiöse Diarrhö, arzneimittelbedingte Lebererkrankungen (jeweils MedDRA-SMQ), Bauchschmerz (PT Bauchschmerz und Oberbauchschmerz) und Erbrechen (PT), einschließlich Zeit bis zum ersten Auftreten, Dauer und Ausgang;
- UE zu Wachstum, Neuroentwicklung, Verhalten und endokrinen Ereignissen bei Personen, die bei Behandlungsbeginn höchstens 2 Jahre alt waren.

Darüber hinaus wurden in Modul 4 erkrankungsbezogene Ereignisse definiert. Hierzu zählen bspw. die PT „Muskuläre Schwäche" und „Gastrointestinale Motilitätsstörung". Die Einstufung der PT als erkrankungsbezogen beruht auf den charakteristischen Krankheitsmanifestationen der TK2d und wurde von der medizinischen Fachabteilung des pU vorgenommen.

Bewertung

Die Operationalisierung ist auf Ebene der ISS weitgehend nachvollziehbar: Behandlungsbegleitender Zeitraum, MedDRA-Kodierung und Regeln zur Zusammenführung wiederholter Ereignisse sind einheitlich festgelegt. Die ISS standardisiert jedoch nur die Zusammenführung vorhandener Angaben, nicht deren Erhebung: Studie 101 beruht auf einer retrospektiven Krankenaktenauswertung ohne festgelegte Visiten, Studie 102 erfasste UE prospektiv und visitenbezogen. Ob ein Ereignis als UE oder als Vorgeschichte gilt, hängt allein vom dokumentierten bzw. ersetzten Beginndatum ab.

Ein Teil der UE beruht auf retrospektiv erfasster Vorgeschichte: In Studie 101 wurden UE und Vorgeschichte gemeinsam als medizinische Ereignisse erfasst, und die Vorgeschichte der Studie 102 enthält UE, die unter Behandlung in Studie 101 auftraten. Abweichende Angaben zum Schweregrad (CTCAE-Grad bzw. Intensität) gleicht die ISS nachträglich an. Wie häufig dies erforderlich war und wie es die Kategorien „Schwere UE" und „SUE" beeinflusste, ist nicht nachvollziehbar. Für den

ISS-Kernpool (N = 39) liegen weder Angaben zu schweren UE (CTCAE-Grad ≥ 3) noch eine Darstellung nach SOC noch eine Auswertung ohne erkrankungsbezogene Ereignisse vor; letztere legt der pU nur für die anwendungsgebietsbezogene ISS-Population (N = 50) vor. In den Auswertungen können daher Manifestationen der TK2d enthalten sein. So beschreiben die ISS-Unterlagen erhöhte Transaminasen sowohl als häufige Manifestation der TK2d als auch als Nebenwirkung der Behandlung. Im Anhang werden die Ergebnisse des ISS-Kernpools (N = 39) deskriptiv dargestellt.

Patientenrelevanz

Der Endpunkt wird als patientenrelevanter Endpunkt entsprechend § 2 Satz 3 AM-NutzenV angesehen.

Validität

Selektionsbias und Immortal Time Bias gelten wie beim Gesamtüberleben (siehe Kapitel 2.2.2), in Studie 102 zusätzlich die Selektion auf Verträglichkeit (siehe Kapitel 2.2.3). UE von Personen, die zuvor verstarben oder die Behandlung abbrachen, fehlen; ihre Anzahl ist nicht bekannt. Die Ergebnisse können daher nicht zur Nutzenbewertung herangezogen werden.

Innerhalb des ISS-Kernpools (N = 39) ist die Validität eingeschränkt, weil Ereigniserhebung, Erfassungsdichte und Datentiefe zwischen den Studien 101 und 102 unterschiedlich waren und ein großer Teil der behandlungsbegleitenden Zeit nur retrospektiv dokumentiert ist. Unvollständige Beginndaten im Monat bzw. Jahr der ersten Dosis wurden auf den Tag der ersten Dosis gesetzt, sodass das Ereignis als behandlungsbegleitend gilt. UE, die bei der letzten Visite der Studie 101 noch andauerten und in Studie 102 als medizinische Vorgeschichte dokumentiert wurden, wurden zu einem Ereignis zusammengeführt (anhand von PT, Beginn, Kausalität und Intensität, laut ISS-SAP zusätzlich Einstufung als schwerwiegend und CTCAE-Grad); verwendet wird der vollständigere Eintrag aus Studie 102. Wie häufig Daten ersetzt oder Einträge zusammengeführt wurden und welchen Einfluss dies auf die Sicherheitsauswertungen hatte, ist nicht berichtet (siehe Kapitel 2.4).

Die Ergebnisse des ISS-Kernpools (Studien 101 und 102; N = 39) werden im Anhang dargestellt: Anzahl und Anteil der Personen mit mindestens einem UE, einem SUE bzw. einem UE mit Abbruch der Studienmedikation sowie UE mit einer Inzidenz ≥ 10 % und SUE mit einer Inzidenz ≥ 5 % nach PT. Zu den fehlenden Auswertungen (schwere UE, Darstellung nach SOC, Auswertung ohne erkrankungsbezogene Ereignisse) siehe Bewertung der Operationalisierung. Eine Auswertung unter Nichtberücksichtigung erkrankungsbezogener Ereignisse bzw. von Ereignissen der Grunderkrankung legt der pU für den ISS-Kernpool nicht vor (nur für die anwendungsgebietsbezogene ISS-Population, N = 50; siehe oben). Es ist nicht auszuschließen, dass Ereignisse der Grunderkrankung mit in die Erfassung der UE eingingen.

2.3.5 Erhebungszeitpunkte

Für den ISE-Kernpool lassen sich keine einheitlichen endpunktspezifischen Studienvisiten angeben. Den gepoolten Daten liegen unterschiedliche Erhebungsstrukturen der Studien 101 und 102 zugrunde. Die nachträgliche Zuordnung zu Auswertungszeitpunkten relativ zu ISE-Tag 1 ersetzt keinen einheitlichen prospektiven Erhebungsplan (siehe Kapitel 2.2.2, Abschnitt „Zeitliche Vereinheitlichung der Studiendaten“).

Für die retrospektive Studie 101 sind keine systematischen Erhebungsregeln berichtet. Die Erhebung folgte der klinischen Routine und war damit patientenindividuell. Systematische Erhebungszeitpunkte sind ausschließlich für Personen, die in Studie 102 eingeschlossen wurden, für den Zeitraum der Studie beschrieben.

Für Studie 102 ist ein prospektiver Erhebungsplan festgelegt. Die motorische Funktion sollte zu Studienbeginn, in Monat 1 sowie bei den weiteren planmäßigen Verlaufsvisiten erhoben werden. Künstliche Beatmung und künstliche Ernährung wurden ebenfalls bei den planmäßigen Visiten

dokumentiert (siehe Abschnitte „Funktionsbereich Atmung“ und „Funktionsbereich Ernährung“). UE wurden zu Studienbeginn und bei den planmäßigen Visiten erhoben; UE waren zudem ereignisbezogen zu dokumentieren. Für Personen mit vorzeitigem Therapieabbruch konnten Überlebensdaten mindestens jährlich erhoben werden, solange Studie 102 lief.

Auch der ISS-Kernpool beruht auf Daten mit unterschiedlichen Erhebungsstrukturen aus den Studien 101 und 102. Ein einheitlicher Visitenplan für die Sicherheitsauswertung ist daher nicht gegeben (siehe Kapitel 2.2.3 und 2.3.4).

2.4 Statistische Methoden

Grundlage der statistischen Auswertungen sind die studienübergreifenden Analysepläne der ISE (ISE-SAP vom 08.10.2021 mit Amendment 1 vom 08.06.2023 und Amendment 2 vom 03.05.2024) und der ISS (ISS-SAP Version 1.0 vom 20.09.2021, finale Version 2.0 vom 31.01.2022, Version 3.0 (Amendment 2) vom 08.06.2023 und Version 4.0 (Amendment 3) vom 10.11.2023). Der indirekte Vergleich des Gesamtüberlebens mit unbehandelten Personen wird in Kapitel 2.6eingeordnet und im Anhang beschrieben bzw. bewertet.

Bezug der Analysepläne der ISE und ISS auf den Kernpool der Studien 101 und 102

Die Festlegungen des ISE-SAP und des ISS-SAP beziehen sich auf die jeweiligen Gesamtpools der ISE und der ISS, die neben den Studien 101 und 102 weitere Datenquellen einschließen. Diese Gesamtpools und ihre anwendungsgebietsbezogenen Populationen werden für die Nutzenbewertung nicht herangezogen (siehe Kapitel 2.2.2 und 2.2.3).

Betrachtet wird ausschließlich der Kernpool der Studien 101 und 102. Er ist im ISE-SAP (Amendment 2) als zusätzlich ausgewertete Teilpopulation („clinical trial subjects“) vorgesehen und umfasst 50 behandelte Personen, davon 39 dem Anwendungsgebiet entsprechend mit Symptombeginn im Alter von ≤ 12 Jahren. Eine eigenständige Auswertung für sie ist in den Analyseplänen der ISE und ISS nicht vorgesehen. Die Ergebnisse dieser 39 Personen werden aufgrund der Vielzahl an Limitationen (siehe Kapitel 2.2.2 und 2.2.3) deskriptiv im Anhang dargestellt. Die folgende Beschreibung dient der Einordnung, nach welchen Regeln diese Teilpopulation ausgewertet wurde.

Zeitbezüge der Auswertungen

Für die im Anhang dargestellten Morbiditätsendpunkte („Motorische Meilensteine“, „Künstliche Beatmung“, „Künstliche Ernährung“) gelten einheitlich die folgenden Zeitbezüge, die auch in Kapitel 2.3 und im Anhang verwendet werden:

- Behandlungsbeginn (ISE-Tag 1): Die erste dokumentierte Gabe einer berücksichtigten Formulierung (dC/dT, dCMP/dTMP oder DC/DT) liegt bei allen Personen des Kernpools vor der Aufnahme in Studie 101 bzw. 102, der Behandlungsbeginn ist also weder der Studieneintritt noch Studientag 1 der Studie 102 (siehe Kapitel 2.2.2, Abschnitt „Zeitliche Vereinheitlichung der Studiendaten“). Für die 29 Personen aus Studie 101 liegt er in der dort retrospektiv dokumentierten, außerhalb der Studie begonnenen Behandlung, für die 10 Personen der Studie 102 ohne Teilnahme an Studie 101 in der nicht näher beschriebenen Vorbehandlung.
- Ausgangsstatus (im ISE-SAP: „treatment baseline“; in den Tabellen: Status „zu Behandlungsbeginn“, „at treatment start“): der zuletzt vor ISE-Tag 1 dokumentierte Status, der für den Tag des Behandlungsbeginns übernommen wird – ohne zeitliche Begrenzung, er kann also lange vor dem Behandlungsbeginn dokumentiert worden sein. „Zu Behandlungsbeginn“ bezeichnet damit keinen eigenen Erhebungszeitpunkt, sondern den auf ISE-Tag 1 bezogenen Ausgangsstatus. Fehlte er in Studie 101, wurde ersatzweise der erste Status nach Behandlungsbeginn verwendet; für die 10 Personen der Studie 102 ohne vorherige Teilnahme an Studie 101 beruht er vermutlich auf anamnestischen Angaben bei Studien-

eintritt unter laufender Behandlung und schließt die außerhalb der Studien begonnene Vorbehandlung ein (siehe Kapitel 2.2.2).

- Zeitraum „nach Behandlungsbeginn“: Die gesamte individuelle Beobachtung von ISE-Tag 1 bis zur letzten dokumentierten Angabe vor dem Datenschnitt der Studie 102 vom 14.03.2025 (für die eine nicht in Studie 102 übernommene Person bis zur letzten Visite in Studie 101). Ereignisse zählen, wenn sie zu irgendeinem Zeitpunkt in diesem Zeitraum dokumentiert sind. Der Zeitraum ist zwischen den Personen unterschiedlich lang (Behandlungsdauer zum Datenschnitt vom 15.03.2024 im Median 78,4 Monate (Spannweite: 53,2–157,1), jeweils Monate zu rund 28 Tagen; zum Datenschnitt vom 14.03.2025 liegen hierzu keine Angaben vor) und schließt die Zeit vor der Studienaufnahme ein, die nur retrospektiv erfasst ist.
- Beobachtungsende („bei Beobachtungsende“): Der zuletzt dokumentierte Status am Ende dieses Zeitraums. Zeitpunkt und Dauer eines Ereignisses sowie zwischenzeitliche Verläufe werden nicht ausgewiesen; zeitpunktbezogene Auswertungen liegen nicht vor.
- Sicherheit (ISS): Bezugspunkt ist die erste Gabe einer beliebigen Formulierung der Studienmedikation (Tag 1), bei Personen aus Studie 101 also der Behandlungsbeginn in Studie 101. Für die 10 Personen der Studie 102 ohne vorherige Teilnahme an Studie 101 (Behandlungsbeginn außerhalb der Studien) ist nicht angegeben, ob Tag 1 die erste Gabe vor der Studienaufnahme oder die erste Gabe in Studie 102 bezeichnet; im ersten Fall fiele die nicht systematisch erfasste Vorbehandlungszeit in die behandlungsbegleitende Zeit. Als behandlungsbegleitend gelten alle Ereignisse ab Tag 1 bis 28 Tage nach der letzten Dosis; Ereignisse, die in Studie 102 als Vorgeschichte dokumentiert wurden, aber am oder nach Tag 1 begannen, werden nachträglich als UE gewertet. Personen, die an beiden Studien teilnahmen, werden über eine studienübergreifende Kennung nur einmal gezählt.

Die Zeit vor Behandlungsbeginn (prätherapeutische Phase von Symptombeginn bis ISE-Tag 1) ist ausschließlich retrospektiv dokumentiert; Auswertungen hierzu werden im Anhang nicht dargestellt (siehe Abschnitt „Statistische Auswertung der im Anhang dargestellten Ergebnisse“).

Datenschnitte

Studie 101 ist abgeschlossen (Ende der Datenerhebung 28.03.2019 mit letzter möglicher Visite am 15.12.2018). Für die laufende Studie 102 liegen Datenschnitte vom 15.03.2024 und vom 14.03.2025 vor; der spätere wurde von der EMA nur für die Wirksamkeitsauswertungen der Studie 102 angefordert. Für den Kernpool liegen zu diesem Datenschnitt in den ISE-Unterlagen nur die Morbiditätsendpunkte vor; Angaben zum Gesamtüberleben des Kernpools sind nur zum 15.03.2024 ausgewiesen, eine Begründung hierfür ist nicht angegeben. Im Anhang werden daher die Patientencharakteristika, die Exposition, das Gesamtüberleben und die UE zum Datenschnitt vom 15.03.2024 und die Morbiditätsendpunkte zum Datenschnitt vom 14.03.2025 dargestellt. Ergebnisse zum finalen Datenschnitt der Studie 102 (geplant August 2026) liegen nicht vor.

Statistische Auswertung der im Anhang dargestellten Ergebnisse

Alle Ergebnisse werden ausschließlich deskriptiv dargestellt. Die Ergebnisse werden nicht zu festen Auswertungszeitpunkten, sondern über die gesamte, zwischen den Personen unterschiedlich lange Beobachtungszeit nach Behandlungsbeginn zusammengefasst (Behandlungsdauer zum Datenschnitt vom 15.03.2024 im Median 78,4 Monate (Spannweite: 53,2–157,1); zum Datenschnitt vom 14.03.2025 liegen hierzu keine Angaben vor, siehe Abschnitt „Zeitbezüge der Auswertungen“). Zeitpunkt eines Ereignisses und Dauerhaftigkeit eines Zustands – etwa eines wiedererlangten Meilensteins oder einer beendeten Beatmung – sind daraus nicht erkennbar. Die Darstellung erfolgt für den ISE-/ISS-Kernpool insgesamt.

- *Motorische Meilensteine:* Dargestellt werden Wiedererlangen und Verlust nach Behandlungsbeginn. Für das Wiedererlangen kommt infrage, wer vor Behandlungsbeginn mindestens einen Meilenstein verloren und nicht wiedererlangt oder ihn nach Behandlungsbeginn

verloren hat; für den Verlust, wer vor Behandlungsbeginn mindestens einen Meilenstein erreicht und nicht verloren, ihn nach Behandlungsbeginn erreicht oder ihn wiedererlangt hat. Die Anteile beziehen sich auf diese ereignisspezifischen Bezugsgruppen (N_1) bzw. je Meilenstein auf die Personen, die ihn wiedererlangen oder verlieren konnten (N_x; 14 bis 38 Personen); eine Person kann bei mehreren Meilensteinen gezählt werden, sodass die Anteile je Meilenstein nicht addierbar sind. Ob wiedererlangte Fähigkeiten bis zum Beobachtungsende erhalten blieben, wird nicht ausgewiesen. Auswertungen zu Verlust und Wiedererlangen vor Behandlungsbeginn liegen für den Kernpool zum Datenschnitt vom 14.03.2025 ebenfalls vor, werden jedoch nicht dargestellt; der Netto-Score (Saldo aus erreichten bzw. wiedererlangten und verlorenen Meilensteinen, siehe Kapitel 2.3.2), das erstmalige Erreichen von Meilensteinen nach Behandlungsbeginn und Angaben zur Anzahl der zu Behandlungsbeginn vorliegenden Meilensteine liegen für den Kernpool zu diesem Datenschnitt nicht vor.

- *Künstliche Beatmung* und *Künstliche Ernährung:* Dargestellt werden Ergebnisse für ineinander geschachtelte Teilgruppen (mit bzw. ohne Beatmung oder Sonde zu Behandlungsbeginn sowie mit Beatmung oder Sonde zu beliebigem Zeitpunkt nach Behandlungsbeginn) mit der jeweils übergeordneten Teilgruppe als Nenner. Einzelne Teilgruppen umfassen weniger als zehn Personen. Der Ausgangsstatus entscheidet über die Zuordnung zu den Teilgruppen, in denen Beginn bzw. Beendigung ausgewertet werden (Zeitbezüge siehe oben). Dauer der Unterstützung, tägliche Beatmungsstunden, Anteil der Sondennahrung und die Kategorien „vor Behandlungsbeginn“ werden nicht dargestellt (siehe Kapitel 2.3.2).
- *Unerwünschte Ereignisse:* Anzahl der Personen mit mindestens einem UE, SUE bzw. UE mit Abbruch der Studienmedikation, bezogen auf N = 39, unabhängig von der individuellen Beobachtungs- und Expositionsdauer; expositionsadjustierte Raten liegen nicht vor. Mehrere Ereignisse derselben Person gehen einmal ein, bei unterschiedlicher Einstufung mit dem schwersten Ereignis. Eine Darstellung nach SOC, UE mit CTCAE-Grad ≥ 3 und eine Auswertung ohne erkrankungsbezogene Ereignisse liegen nicht vor.

Fehlende Werte und Daten-Imputation

Fehlende Werte wurden grundsätzlich nicht ersetzt. Unvollständige Datumsangaben wurden nach den studienübergreifenden Imputationsregeln ersetzt, die je Variable die Regeln der Einzelstudien übernehmen und für ISE und ISS ergänzen. Folgende Regeln betreffen die im Anhang dargestellten Ergebnisse:

- *Motorische Meilensteine:* Unvollständige Datumsangaben zu „Erreichen“, „Verlust“ und „Wiedererlangen“ wurden in beiden Studien durch den 15. des Monats bzw. den 30. Juni ersetzt. Fehlte nach einem Erreichen die Angabe zum Verlust, galt der Meilenstein als nie verloren; fehlte nach einem Verlust die Angabe zum Wiedererlangen, galt er als nie wiedererlangt. Undatierte Ereignisse wurden nach einer festgelegten Reihenfolge der Meilensteine der Zeit vor oder nach Behandlungsbeginn zugeordnet (Zuordnung aus Studie 101 übernommen, in Studie 102 nicht vorgenommen und auf ISE-Ebene bei fehlender Phase nachgeholt). Diese Reihenfolge ist keine klinische Schweregradskala.
- *Künstliche Beatmung* und *Künstliche Ernährung:* Unvollständige Start- und Enddaten wurden nach den Regeln der Einzelstudien ersetzt; fehlte ein Enddatum vollständig, wurde der Tag vor dem nächsten Beginndatum eingesetzt; Lücken wurden als durchgehende Unterstützung gewertet; auf ISE-Ebene erfolgte keine weitere Ersetzung.
- *Unerwünschte Ereignisse:* Lag ein unvollständiges Beginndatum (nur Monat und Jahr bzw. nur Jahr bekannt) im Monat bzw. Jahr der ersten Dosis, wurde das Datum der ersten Dosis eingesetzt, andernfalls der erste Tag des Monats bzw. der 1. Januar. Auf ISS-Ebene wurden die Ersetzungen der Einzelstudien übernommen, Einträge derselben Person studien-

übergreifend zusammengeführt und erst bei danach weiterhin unvollständigen Angaben die Start- bzw. Enddatumsregeln erneut angewandt.

- *Exposition:* Fehlte das Behandlungsende, wurde es in Studie 101 auf den 31.12.2018 und in Studie 102 auf das Datum der letzten Beobachtung gesetzt; die Behandlungsdauer kann dadurch überschätzt sein. Unvollständige Datumsangaben zur genetischen Testung wurden in beiden Studien ersetzt; dieses Datum wird in ISE und ISS nicht ausgewertet.
- *Patientencharakteristika:* Fehlende Angaben sind als eigene Kategorie ausgewiesen (z. B. Mobilitätshilfe bei 21, Muskelbiopsiebefunde bei 15, Atmungskettenkomplexe bei 29 Personen); die Anteile beziehen sich auf N = 39 und nicht auf die Personen mit vorliegender Angabe. Unvollständige Datumsangaben zum Symptombeginn wurden in beiden Studien durch den 15. des Monats bzw. den 30. Juni ersetzt; lagen Angaben aus beiden Studien vor, galt vorrangig ein vollständiges Datum, andernfalls das Datum der späteren Studie.

Wie häufig diese Regeln angewandt wurden, ist nicht berichtet; Sensitivitätsanalysen liegen nicht vor. Die Ersetzungen wirken nicht in eine einheitliche Richtung: Sie verschieben Ereignisse in die Behandlungsphase (UE mit unvollständigem Beginndatum im Monat bzw. Jahr der ersten Dosis), verlängern Unterstützungszeiten (Lücken als durchgehende Unterstützung) und werten fehlende Angaben zu Verlust bzw. Wiedererlangen als „nicht eingetreten".

Einschätzung der Auswertungen

Für den Kernpool der Studien 101 und 102 bestehen folgende Einschränkungen, die für alle im Anhang dargestellten Ergebnisse gelten:

- *Selektionsbias und Immortal Time Bias* (siehe Kapitel 2.2.2): Die Aufnahme in Studie 101 setzte Überleben und eine letzte dokumentierte Visite zwischen dem 01.06.2018 und dem 15.12.2018 voraus, die Aufnahme in Studie 102 eine laufende Behandlung. Die Zeit zwischen Behandlungsbeginn und letzter dokumentierter Visite (Studie 101) bzw. Studienaufnahme (Studie 102) mussten alle Personen überleben (Immortal Time); Personen mit frühem Tod oder Behandlungsabbruch sind potentiell unterrepräsentiert.
- *Ausgangswert:* Der Ausgangswert beruht für Studie 101 auf Krankenakten und für die Personen der Studie 102 ohne vorherige Teilnahme an Studie 101 auf anamnestischen Angaben, die bei Studieneintritt unter laufender Behandlung erhoben wurden.
- *Erhebung:* Die Zeit vor Behandlungsbeginn ist – wenn überhaupt – ausschließlich retrospektiv dokumentiert, die Zeit danach in Studie 101 retrospektiv und im Zuge der Studie 102 prospektiv, jedoch für Studie 102 erst nach dem tatsächlichen Behandlungsbeginn. Die Vollständigkeit der Erhebung in Studie 101 lässt sich nicht prüfen. Unvollständigkeit und Unterschiede in der Erhebungsdichte können die Validität der Ergebnisse beeinträchtigen (siehe Kapitel 2.3.2 und 2.3.4).
- *Sicherheit:* Die behandlungsbegleitende Phase beginnt mit der ersten Gabe in Studie 101, auch mit nicht GMP-konformen Arzneimittel-Formulierungen und möglichen Dosisabweichungen gegenüber den Empfehlungen der Fachinformation, und ist bis zum Eintritt in Studie 102 nur retrospektiv dokumentiert. Erkrankungsbezogene Ereignisse werden nicht abgegrenzt, und die Aufnahme in Studie 102 setzte eine bis dahin ausreichende Verträglichkeit voraus (siehe Selektionsbias).
- *Behandlung:* Vor Studie 102 erfolgte die Behandlung mit nicht GMP-konformen Arzneimittel-Formulierungen in individueller Dosierung.
- *Fehlende Werte:* Die Ersetzungs- und Zuordnungsregeln wirken nicht in eine einheitliche Richtung; da ihre Häufigkeit nicht berichtet ist und Sensitivitätsanalysen fehlen, lässt sich ihr Einfluss auf die dargestellten Anteile nicht abschätzen.
- *Deskriptive Auswertung ohne Vergleich:* Berichtet werden ausschließlich Anzahl und Anteil; Konfidenzintervalle, Tests oder Adjustierungen liegen nicht vor, und ohne Kontrollgruppe

lässt sich der natürliche Verlauf – bei Kindern einschließlich der physiologischen motorischen Entwicklung – nicht von einem Behandlungseffekt trennen.

- *Bezugsgruppen und Nenner:* Die Anteile beziehen sich auf endpunkt- und ereignisspezifische Teilgruppen, die über den Ausgangsstatus definiert sind (Bezugsgruppen N_1 bzw. N_x bei den Meilensteinen, Teilgruppen mit bzw. ohne Beatmung oder Sonde); Ergebnisse verschiedener Kategorien sind daher weder untereinander noch mit auf N = 39 bezogenen Angaben vergleichbar.
- *Symptombeginn und Anwendungsgebiet:* Die Zuordnung zum Anwendungsgebiet beruht auf dem nicht definierten, retrospektiv erhobenen Alter bei Symptombeginn (siehe Kapitel 2.2.2); für die 10 Personen, die keine Indexfälle waren, ist ein Behandlungsbeginn vor dem Symptombeginn nicht auszuschließen (siehe Anhang, Abschnitt „Informationen zur Diagnose“).
- *Fehlende Auswertungen:* Netto-Score, erstmaliges Erreichen von Meilensteinen nach Behandlungsbeginn, Anzahl der zu Behandlungsbeginn vorliegenden Meilensteine, schwere UE (CTCAE-Grad ≥ 3), UE nach SOC und eine Auswertung ohne erkrankungsbezogene Ereignisse liegen für den Kernpool nicht vor.
- *Präspezifikation:* Der Kernpool wurde erst mit Amendment 2 des ISE-SAP vom 03.05.2024 als Auswertungspopulation festgelegt; die Analysepläne der ISE und ISS entstanden nach Abschluss der Datenerhebung der Studie 101 und in Kenntnis eines großen Teils der Daten der Studie 102.
- *Boden- und Deckeneffekte:* Bei den motorischen Meilensteinen hängt die Möglichkeit eines Gewinns oder Verlusts vom Funktionsstatus bei Behandlungsbeginn ab; wer noch keinen Meilenstein erreicht hatte, konnte keinen verlieren, wer alle erreicht hatte, keinen hinzugewinnen.

Die Ergebnisse des Kernpools sind damit nicht hinreichend valide, um sie für die Bewertung heranzuziehen; sie werden ausschließlich deskriptiv im Anhang dargestellt und erlauben keine Aussage zu einem Behandlungseffekt.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Da die Studien 101 und 102 unkontrolliert waren, ist von einem hohen Verzerrungspotential auf Studien- und Endpunktebene auszugehen, mit der zusätzlichen Einschränkung, dass die Vollständigkeit der Ergebnisse der retrospektiven Studie 101 nicht geprüft werden kann.

2.6 Indirekter Vergleich

Der vom pU vorgelegte indirekte Vergleich des Gesamtüberlebens mit einer historischen, unbehandelten Kontrollgruppe beruht auf der anwendungsgebietsbezogenen ISE-Population. Da die anwendungsgebietsbezogene ISE-Population (N = 82) für die Nutzenbewertung nicht herangezogen wird, werden auch der indirekte Vergleich für diese Population und die ergänzenden Analysen einschließlich der TTE nicht berücksichtigt. Der indirekte Vergleich wird im Anhang (Kapitel 7.5) beschrieben, und die Limitationen, die über die in Kapitel 2.2.2 beschriebenen Limitationen der behandelten Population (z. B. Selektionsbias) hinausgehen, werden dort ausführlich dargelegt.

3 Ergebnisse der eingeschlossenen Studien

3.1 Studiencharakteristika und Studienmedikation

Die Studiencharakteristika und die Studienmedikation des ISE-/ISS-Kernpools (N = 39) sind im Anhang dargestellt (zur Einordnung siehe Kapitel 2.2.2 und 2.2.3).

3.2 Mortalität

Die Ergebnisse des ISE-Kernpools (N = 39) werden aufgrund der in den Kapiteln 2.2.2 und 2.3.1 beschriebenen Limitationen nicht herangezogen; Anzahl und Anteil verstorbener Personen sind dem Anhang zu entnehmen.

3.3 Morbidität

Die Ergebnisse des ISE-Kernpools (N = 39) werden aufgrund der in den Kapiteln 2.2.2 und 2.3.2 beschriebenen Limitationen nicht herangezogen; sie sind dem Anhang zu entnehmen.

3.4 Lebensqualität

Für die Lebensqualität legt der pU keine studienübergreifende Auswertung der ISE vor. Ergebnisse zur gesundheitsbezogenen Lebensqualität liegen damit auch für den ISE-Kernpool nicht vor.

3.5 Sicherheit

Die Ergebnisse des ISS-Kernpools (N = 39) werden aufgrund der in den Kapiteln 2.2.3 und 2.3.4 beschriebenen Limitationen nicht herangezogen; sie sind dem Anhang zu entnehmen.

4 Diskussion der Methodik und Ergebnisse

Die Bewertung bezieht sich auf die Behandlung von Kindern und Erwachsenen mit genetisch bestätigter Thymidinkinase-2-Defizienz (TK2d) mit einem Alter bei Symptombeginn bis zur Vollendung des 12. Lebensjahres.

4.1 Design und Methodik der studienübergreifenden Auswertungen (ISE und ISS)

Studiendesign

Für die Bewertung von Doxecitin/Doxribtimin legt der pU keine Einzelstudie, sondern zwei studienübergreifende Auswertungen vor: eine integrierte Wirksamkeitsanalyse (ISE) der Studien 101, 102, 107 und 114 (anwendungsgebietsbezogene ISE-Population: N = 82) und eine integrierte Sicherheitsauswertung (ISS) der Studien 101, 102 und 107 (anwendungsgebietsbezogene ISS-Population: N = 50). Alle Datenquellen sind einarmig; keine wurde darauf ausgelegt, den Effekt der Behandlung gegenüber einer Vergleichstherapie zu bestimmen. Ein randomisierter Vergleich wurde nicht angestrebt; das CHMP hält einen solchen aufgrund der Seltenheit der Erkrankung und der für Überleben und motorische Entwicklung erforderlichen Beobachtungsdauer für nicht durchführbar [2]. Die Limitationen beider Auswertungen sind in den Kapiteln 2.2.2 und 2.2.3 ausführlich dargelegt; nachfolgend werden die für die Bewertung maßgeblichen Punkte zusammengefasst. In Anbetracht des langen Zeitraums seit Beginn der Studien (Beginn der Datenerhebung in Studie 101 im Oktober 2018, erste Aufnahme in Studie 102 im Juli 2019) sowie der wesentlichen Limitationen und geringen Aussagekraft der vorliegenden Daten wäre eine prospektive Studie – wenn nicht gar eine vergleichende Studie – mit erster Behandlung nach Studienaufnahme wünschenswert gewesen.

Die ISE ist weder eine prospektiv geplante Kohorte noch eine Metaanalyse, sondern eine nachträgliche Zusammenlegung von Individualdaten aus vier einarmigen Studien unterschiedlichen Designs – zwei retrospektiven Krankenaktenauswertungen (Studien 101 und 107), einer retrospektiven Auswertung von EAP-Formularen (Studie 114) und einer prospektiven Fortsetzungsstudie (Studie 102) – ohne Stratifizierung nach Studie und ohne Prüfung der Ähnlichkeit der Studienpopulationen. Maßgeblich sind folgende Limitationen:

- **Selektionsbias:** Die retrospektiven Studien beruhen auf Krankenakten bzw. EAP-Formularen; weder die Auswahl der Zentren noch ein Patientenfluss (gesichtete Akten, nicht eingeschlossene Personen, Gründe) sind beschrieben, sodass unbekannt bleibt, welcher Anteil der insgesamt behandelten Personen erfasst wurde. Hinzu kommt, dass der Einschluss in den Studien 101, 102 und 114 an Bedingungen geknüpft war, die erst nach Behandlungsbeginn erfüllt sein konnten: eine dokumentierte Visite zwischen dem 01.06.2018 und dem 15.12.2018 (Studie 101), eine laufende und vertragene Behandlung bei Aufnahme (Studie 102) bzw. die Aufnahme in das EAP (Studie 114; für bereits vorbehandelte Personen also das Überleben bis zur EAP-Aufnahme). Personen, die vorher verstarben, die Behandlung abbrachen oder nicht weiter nachbeobachtet wurden, konnten damit nicht eingeschlossen werden; dass solche Verläufe unter Behandlung vorkommen, zeigen die fünf Todesfälle unter den 18 behandelten Personen der Studie 107, darunter zwei nach 23 bzw. 44 Behandlungstagen. Die eingeschlossenen Personen dürften daher einen günstigeren Verlauf aufweisen als die Gesamtheit der behandelten Personen; das Ausmaß dieser Selektion lässt sich nicht quantifizieren und durch weitere Auswertungen der eingeschlossenen Personen nicht ausgleichen.
- **Immortal Time Bias:** Als Startpunkt der Auswertungen (ISE-Tag 1) gilt der Behandlungsbeginn, nicht der Studieneintritt. Für den Großteil der Personen liegt der Behandlungsbeginn vor der Studienaufnahme, in Studie 101 teils um Jahre; weil die Aufnahme Überleben und fortgesetzte Behandlung voraussetzte, konnte in dieser Zeitspanne definitionsgemäß kein Tod und kein Behandlungsabbruch beobachtet werden. Dennoch geht sie als ereignisfreie

Zeit unter Risiko in die Auswertung ein; ein verzögerter Eintritt in die Risikomenge ab Studienaufnahme ist im ISE-SAP nicht vorgesehen und Maßnahmen zur Begrenzung dieses Bias betreffen nur den Vergleich mit unbehandelten Personen. Überlebenszeiten dürften dadurch überschätzt und Ereignisraten unterschätzt sein, am stärksten in den ersten Behandlungsjahren, in denen die Sterblichkeit bei frühem Symptombeginn am höchsten ist. Da die Zeit zwischen Behandlungsbeginn und Studienaufnahme je Person nicht ausgewiesen ist, lässt sich das Ausmaß nicht abschätzen.

- **Ausgangsstatus:** Ob ein Meilenstein wiedererlangt oder verloren wurde und ob eine Person zu Behandlungsbeginn beatmet oder sondenernährt war, wird am letzten vor ISE-Tag 1 dokumentierten Status festgemacht. Ein einheitlich und vor Behandlungsbeginn erhobener Ausgangsstatus liegt jedoch für keinen Morbiditätsendpunkt vor: In Studie 101 stammt er retrospektiv aus Krankenakten und wurde, wenn er fehlte, durch den ersten Status nach Behandlungsbeginn ersetzt, also durch einen möglicherweise bereits behandlungsbeeinflussten Wert; für die 12 Personen der Studie 102 ohne Teilnahme an Studie 101 stammt er vermutlich aus anamnestischen Angaben, die bei Studieneintritt unter laufender Vorbehandlung unbekannter Art und Dauer erhoben wurden; für Personen ausschließlich aus Studie 114 fehlt er. Je nachdem, aus welcher Studie eine Person stammt, beziehen sich Veränderungen damit auf unterschiedlich erhobene Ausgangswerte und sind untereinander nicht vergleichbar. Die Richtung einer möglichen Verzerrung ist nicht eindeutig: Ein bereits unter Behandlung erhobener Ausgangsstatus könnte frühe Veränderungen nach Behandlungsbeginn ebenso verdecken wie vorbestehende Verluste.
- **Abgrenzung des Anwendungsgebiets:** Das Anwendungsgebiet ist über zwei Merkmale definiert, eine genetisch bestätigte TK2d und ein Alter bei Symptombeginn bis zur Vollendung des 12. Lebensjahres [13]. Die anwendungsgebietsbezogenen Populationen und der Kernpool wurden aus den Gesamtpools, deren Zusammensetzung bereits durch die Aufnahmebedingungen der Einzelstudien (Behandlung, Überleben bis zur maßgeblichen Visite bzw. laufende Behandlung bei Studienaufnahme) bestimmt ist, allein über das Alter bei Symptombeginn abgegrenzt, das in keiner Studie definiert und überwiegend retrospektiv aus Krankenakten oder anamnestisch erhoben wurde. Welche Symptome als Erstmanifestation galten und ob der Beginn zeitnah oder erst Jahre später rückblickend dokumentiert wurde, ist nicht berichtet. Fehlzuordnungen zum Anwendungsgebiet und zu den Alterssubgruppen (≤ 2 bzw. > 2 bis ≤ 12 Jahre) sind daher nicht auszuschließen – für 1 Person der Studie 107 weicht die Zuordnung zum Anwendungsgebiet zwischen Studienbericht und ISE bereits ab (siehe Kapitel 2.2.2) –; sie würden sich zugleich auf die Überlebenszeitanalysen ab Symptombeginn auswirken. Die genetische Bestätigung war zwar Einschlusskriterium aller Studien; Angaben dazu, wie viele Personen sie tatsächlich aufwiesen und welche Mutationen vorlagen, wurden jedoch nicht vorgelegt (siehe Kapitel 4.2).
- **Behandlung und Beobachtung:** Die Studien erfassten Behandlung und Verlauf zu unterschiedlichen Zeiten, ab unterschiedlichen Bezugspunkten und in unterschiedlicher Dichte, prospektiv nach Visitenplan nur in Studie 102, sonst nach Verfügbarkeit der Routinedokumentation. Behandelt wurde überwiegend mit nicht GMP-konformen Formulierungen (dC/dT, teils dCMP/dTMP) in individuell festgelegter Dosierung; ob gemäß Fachinformation titriert und dreimal täglich mit Nahrung eingenommen wurde, ist für keine Studie sichergestellt, und die Einnahme mit Nahrung war in Studie 102 erst ab 2021 vorgegeben. Die Datenschnitte der Studien liegen gut sechs Jahre (ISE) bzw. gut fünf Jahre (ISS) auseinander, die Beobachtungsdauer nach Behandlungsbeginn ist weder je Studie noch je Population ausgewiesen und die Zeit zwischen der letzten Visite in Studie 101 und der Aufnahme in Studie 102 (mindestens knapp sieben Monate) ist nur retrospektiv und teilweise erfasst. Die Regeln des ISE-SAP vereinheitlichen Startzeitpunkt, Ausgangsstatus, Auswertungszeitpunkte und Datenschnitte rechnerisch, nicht jedoch die zugrunde liegende Datenerhebung; Unterschiede zwischen den Studien könnten daher als Unterschiede im Verlauf erscheinen.

- **Datenlücken:** In der anwendungsgebietsbezogenen ISE-Population (N = 82) wurden künstliche Beatmung und künstliche Ernährung bei 29 Personen (35,4 %) nicht erhoben; Angaben liegen für 53 bzw. 52 Personen vor. Bei den motorischen Meilensteinen hatten je nach Meilenstein nur 22 bis 49 Personen den Meilenstein vor Behandlungsbeginn erreicht und für einen Verlust nach Behandlungsbeginn kamen je Meilenstein 21 bis 46 Personen infrage. Die Lücken betreffen ausschließlich Personen außerhalb des Kernpools und entstehen hinsichtlich der Studienherkunft systematisch, nicht zufällig: Studie 107 erhob den Verlauf nur nach Verfügbarkeit der Akten, Studie 114 außer dem Vitalstatus keinen Wirksamkeitsendpunkt; für den Kernpool (N = 39) weist der pU bei allen drei Morbiditätsendpunkten keine fehlenden Angaben aus. Die gepoolten Morbiditätsergebnisse dürften daher überwiegend den Kernpool abbilden, ohne dass dies ausgewiesen wäre.
- **Präspezifikation:** Die Analysepläne der ISE und ISS wurden 2021 erstellt, nach Abschluss der Datenerhebung der Studie 101 und mehr als zwei Jahre nach Beginn der Studie 102, also in Kenntnis eines großen Teils der Daten. Auch Regeln mit Einfluss auf die Ergebnisse, etwa der Ersatz eines fehlenden Ausgangsstatus durch den ersten Wert nach Behandlungsbeginn, wurden erst zu diesem Zeitpunkt festgelegt und der Kernpool der Studien 101 und 102 wurde erst mit Amendment 2 des ISE-SAP vom 03.05.2024 als zusätzliche Auswertungspopulation ergänzt. Eine datengetriebene Festlegung von Populationen und Auswertungsregeln lässt sich daher nicht ausschließen.

Für die ISS, die nur die Studien 101, 102 und 107 umfasst, gelten diese Limitationen entsprechend. Hinzukommen:

- **Selektion auf Verträglichkeit:** In Studie 102 wurden nur Personen mit laufender und stabiler Behandlung aufgenommen; Personen mit Lebererkrankung oder Leberwerten ≥ 2 × ULN waren ohne Freigabe durch den Sponsor ausgeschlossen, und die 2 Personen, die die Behandlung in Studie 101 wegen erhöhter Leberwerte abgebrochen hatten, konnten nicht übergehen. UE, die zu einem Behandlungsabbruch führen, und insbesondere arzneimittelbezogene Leberstörungen dürften in der prospektiv erfassten Phase daher unterrepräsentiert sein; die 2 Personen mit leberwertbedingtem Abbruch sind im ISE-/ISS-Kernpool gesamt (N = 50) enthalten und entfallen im Kernpool (N = 39) wegen eines Symptombeginns nach dem 12. Lebensjahr (siehe Kapitel 2.2.2); dass dort kein UE zum Behandlungsabbruch führte, ist daher nicht Folge dieser Selektion, schließt vor Studieneinschluss aufgetretene und nicht erfasste Abbrüche aber nicht aus.
- **Uneinheitliche Erhebung:** Der Umfang der Sicherheitserhebung unterscheidet sich grundlegend. Studie 102 erfasste UE prospektiv nach Prüfplan, jedoch erst ab Studieneintritt und damit nach teils jahrelanger Vorbehandlung; Studie 101 übernahm sie retrospektiv aus Krankenakten ohne festgelegte Visiten, SUE wurden nicht an den Sponsor gemeldet; Studie 107 erhob nur tödliche SUE sowie UE mit Behandlungsabbruch, -unterbrechung oder Dosisreduktion. Ein großer Teil der behandlungsbegleitenden Zeit ist damit nur retrospektiv dokumentiert, sodass UE vor Eintritt in Studie 102 unvollständig erfasst sein könnten; nicht erhobene Ereignisse sind nicht mit Ereignisfreiheit gleichzusetzen. Da Studie 114 nicht Teil der ISS ist, beruhen ISE (N = 104) und ISS (N = 67) zudem auf unterschiedlichen Populationen und bilden keine gemeinsame Basis für eine Abwägung von Nutzen und Schaden.
- **Zusammenführung und Schweregrad:** Für die 35 Personen, die aus Studie 101 in Studie 102 übergingen, können dieselben Ereignisse in beiden Studien dokumentiert sein. Ein solches Ereignis zählt in der ISS nur dann als ein UE, wenn Beginn, MedDRA-Begriff, Schweregrad, Intensität, Einstufung als schwerwiegend und Kausalitätsbewertung in beiden Einträgen übereinstimmen; andernfalls geht es als zwei UE ein. Der Schweregrad wurde in den Studien nicht einheitlich erhoben (CTCAE-Grad bzw. Intensität, in Studie 102 teils nach einer protokolleigenen Skala) und in der ISS nachträglich vereinheitlicht, wobei bei abweichenden Angaben die schwerere Einstufung übernommen wurde; bei sonstigen Widersprüchen

zwischen den Studien gilt die Angabe der zeitlich letzten Studie. Wie häufig Ereignisse zusammengeführt oder Einstufungen angepasst wurden, ist nicht berichtet; Doppelzählungen und eine Verschiebung zu schwereren Graden sind daher nicht auszuschließen.

- **Zeit unter Risiko:** Als behandlungsbegleitend gelten UE von der ersten Gabe bis 28 Tage nach der letzten Dosis. Behandlungsunterbrechungen bleiben dabei unberücksichtigt, fehlende Behandlungsenden wurden in Studie 101 auf den 31.12.2018 und in Studie 102 auf das Datum der letzten Beobachtung gesetzt und unvollständige Angaben zum Beginndatum eines UE im Monat bzw. Jahr der ersten Dosis wurden auf deren Tag gesetzt. Die Exposition könnte dadurch überschätzt sein und einzelne vor Behandlungsbeginn begonnene Ereignisse könnten der Behandlungsphase zugeordnet sein (konservative Zuordnung, die die UE-Häufigkeiten eher über- als unterschätzt). Expositionsadjustierte Raten liegen nicht vor, sodass sich die UE-Anteile bei Behandlungsdauern zwischen 53 und 157 Monaten nicht ins Verhältnis zur Beobachtungszeit setzen lassen.
- **Auswertungspopulation:** Modul 4 stellt die UE für die anwendungsgebietsbezogene ISS-Population (N = 50) dar, also einschließlich der 11 Personen aus Studie 107, bei denen UE insgesamt, schwere UE und SUE nicht vollständig erhoben wurden; die dort berichteten Anteile könnten dadurch unterschätzt sein. Für den ISS-Kernpool (N = 39) legt der pU in Modul 4 keine UE-Auswertung vor; die im Anhang dargestellten Ergebnisse stammen aus den ISS-Unterlagen und enthalten weder schwere UE (CTCAE-Grad ≥ 3) noch eine Darstellung nach SOC noch eine Auswertung ohne erkrankungsbezogene Ereignisse. Manifestationen der TK2d, etwa eine respiratorische Insuffizienz oder erhöhte Transaminasen, lassen sich in den UE-Raten daher nicht von Nebenwirkungen abgrenzen.

Die gepoolten Auswertungen lassen sich daher nicht als Ergebnisse einer vergleichbar ausgewählten, behandelten und beobachteten Kohorte interpretieren und werden nicht herangezogen. Die Darstellung im Anhang beschränkt sich auf den ISE-/ISS-Kernpool (N = 39) der von der EMA als zulassungsbegründete Hauptstudien eingestuften, vom pU gesponserten Studien 101 und 102. In ihm entfällt die Heterogenität durch die Studien 107 und 114, für alle Personen werden Angaben zu allen drei Morbiditätsendpunkten ausgewiesen, der Vitalstatus ist bis auf eine früh zensierte Person gesichert, und Studie 102 bietet eine prospektive Nachbeobachtung. Die übrigen Limitationen bestehen jedoch fort, weil sie in den beiden Studien selbst angelegt sind: Die Aufnahmebedingungen beider Studien (siehe Abschnitt „Selektionsbias“) setzten Überleben und fortgesetzte, vertragene Behandlung voraus, die Behandlung vor Studie 102 erfolgte außerhalb der Studien mit nicht GMP-konformen Formulierungen nach ärztlichem Ermessen, und die prospektive Erhebung der Studie 102 setzt erst nach dieser Selektion unter laufender Behandlung ein, sodass ihre Baseline keinen unbehandelten Ausgangszustand abbildet. Der Kernpool ist zudem keine prospektiv definierte Kohorte, sondern wurde erst 2024 als zusätzliche Auswertungspopulation festgelegt; die dem Anwendungsgebiet entsprechenden 39 Personen ergeben sich als Teilpopulation mit dem Alter bei Symptombeginn von ≤ 12 Jahren, wobei zur Belastbarkeit der Angaben zum Alter bei Symptombeginn Unklarheiten bestehen (siehe Abschnitt „Studienpopulation“).

Ohne Vergleichsgruppe und ohne unbehandelten Ausgangszustand lassen sich Veränderungen nach Behandlungsbeginn nicht vom natürlichen Verlauf – bei Kindern einschließlich der physiologischen motorischen Entwicklung – abgrenzen; Selektionsbias und Immortal Time Bias könnten den Verlauf zudem günstiger erscheinen lassen, als er in der Gesamtheit der behandelten Personen wäre. Die Ergebnisse des Kernpools werden daher ausschließlich deskriptiv im Anhang dargestellt; die für den Kernpool im Einzelnen fortbestehenden Einschränkungen werden im Abschnitt „Bewertung der Validität und Aussagekraft“ zusammengefasst.

Studienpopulation

Die Studienpopulation ist auf mehreren Ebenen das Ergebnis einer Auswahl: Der ISE-Gesamtpool (N = 104) umfasst alle behandelten Personen der vier Studien, ohne dass für die retrospektiven Studien beschrieben wäre, welche Zentren und Krankenakten gesichtet und aus welchen Gründen Personen nicht eingeschlossen wurden. Die anwendungsgebietsbezogene ISE-Population (N = 82) wird aus dem Gesamtpool allein über das Alter bei Symptombeginn von ≤ 12 Jahren abgegrenzt – ein Merkmal, das in keiner Studie definiert und überwiegend retrospektiv oder anamnestisch erhoben wurde – und enthält aus den Studien 107 und 114 Personen, zu denen außer dem Vitalstatus kaum Verlaufsdaten vorliegen. Der ISE-/ISS-Kernpool (N = 39) beschränkt die Population auf die Studien 101 und 102 und damit auf Personen, die die Einschlussbedingungen der Studie 101 und/oder der Studie 102 erfüllten, also bis zur letzten Visite zwischen dem 01.06.2018 und dem 15.12.2018 bzw. bis zur Aufnahme in Studie 102 überlebt hatten und behandelt wurden (siehe Abschnitt „Studiendesign"). Die folgenden Angaben beziehen sich auf diese 39 Personen, die eine selektierte Population beschreiben dürften und die Zielpopulation der Zulassung nicht ohne Weiteres abbilden.

Alle 39 Personen hatten laut Studienunterlagen einen Symptombeginn im Alter von ≤ 12 Jahren (im Median 1,89 Jahre (Spannweite: 0,31–11,7); 21 Personen im Alter von ≤ 2 Jahren, 18 Personen im Alter von > 2 bis ≤ 12 Jahren) und entsprechen damit formal dem Anwendungsgebiet, sofern diese Angaben valide sind. Die Behandlung begann jedoch im Median erst mit 4,67 Jahren (Spannweite: 0,73–35,5): Nur 9 Personen wurden im Alter von ≤ 2 Jahren erstmals behandelt, 9 erst nach dem 12. Lebensjahr und bei 13 Personen (33,3 %) lagen mehr als fünf Jahre zwischen Symptom- und Behandlungsbeginn. Der Kernpool dürfte damit überwiegend eine Behandlung nach längerem Krankheitsverlauf abbilden. Für Kinder mit frühem Symptombeginn und frühem Behandlungsbeginn – die am schwersten betroffene Gruppe, in der die EMA einen Effekt auf das Überleben für am bedeutsamsten und am ehesten beurteilbar hält [2] – liegen Daten von höchstens 9 Personen vor (siehe Kapitel 4.2). Ob sich Beobachtungen bei spätem Behandlungsbeginn auf eine frühe Behandlung übertragen lassen, bleibt offen.

Wie belastbar die Zuordnung zum Anwendungsgebiet ist, lässt sich nicht beurteilen: Eine Definition des Symptombeginns fehlt in beiden Studien. Für die 29 Personen aus Studie 101 wurde das Alter bei Symptombeginn retrospektiv aus Krankenakten übernommen – Voraussetzung für den Einschluss waren lediglich Aufzeichnungen ab Symptombeginn –, für die 10 Personen ohne Teilnahme an Studie 101 stammt es aus der in Studie 102 retrospektiv erhobenen Vorgeschichte. Ob der Symptombeginn zeitnah dokumentiert oder erst Jahre später rückblickend angegeben wurde, welche Symptome als Erstmanifestation galten und wie er bei familiär früh diagnostizierten Personen bestimmt wurde, ist nicht berichtet. Fehlzuordnungen sind daher nicht auszuschließen: Die Grenze des Anwendungsgebiets von 12 Jahren betrifft bei einem Maximum von 11,7 Jahren einzelne Personen, die Subruppengrenze von 2 Jahren dagegen die Hälfte der Population. Dass alle 39 Personen tatsächlich einen Symptombeginn im Alter von ≤ 12 Jahren hatten, lässt sich mit den vorliegenden Angaben nicht sichern.

Studienmedikation

Die Personen des Kernpools wurden im Median 78,4 Monate (Spannweite: 53,2–157,1; Monate zu rund 28 Tagen) behandelt. Die kürzeste Expositionsdauer von 53,2 Monaten könnte auf die Selektion hindeuten. Es ist anzunehmen, dass Personen, die die Behandlung früh abbrachen oder verstarben, im Kernpool unterrepräsentiert sind, da die Aufnahme in Studie 101 Überleben und dokumentierte Behandlung bis mindestens zum 01.06.2018 und die Aufnahme in Studie 102 eine laufende Behandlung voraussetzte (Aufnahmen ab 10.07.2019). Die Person mit der kürzesten Exposition (Behandlungsbeginn rechnerisch Anfang 2020) muss zu den 10 erst in Studie 102 aufgenommenen Personen gehören; wann die Behandlung der übrigen Personen frühestens begann, ist nicht berichtet. Die lange Exposition dürfte damit zumindest teilweise auf die

Aufnahmekriterien zurückzuführen sein und ließe sich nicht ohne Weiteres als Hinweis auf eine gute Verträglichkeit oder Wirksamkeit der Behandlung deuten. Die durchschnittlich erhaltene Tagesdosis betrug im Median 393,6 mg/kg/Tag (Spannweite: 169,6–400,0) und lag damit nahe der maximalen Erhaltungsdosis laut Fachinformation von je 400 mg/kg/Tag Doxecitin und Doxribtimin. Die untere Grenze der Spannweite deutet jedoch darauf hin, dass ein Teil der Personen über die Behandlungsdauer deutlich geringer dosiert wurde; Gründe hierfür (z. B. Titration, Unverträglichkeit oder Unterbrechungen) sind nicht berichtet. Die Angabe bezieht sich laut ISS-Unterlagen auf jede Einzelkomponente der Wirkstoffkombination (siehe Kapitel 2.2.3).

Der überwiegende Teil der Beobachtungszeit entfällt auf eine Behandlung mit nicht GMP-konformen Formulierungen (dC/dT, teils dCMP/dTMP) in individuell festgelegter Dosierung; erst mit Studie 102 wurde auf DC/DT umgestellt. Genaue Angaben dazu, wie viele Personen wie lange welche Formulierung erhielten, liegen nicht vor. Titration, Einnahme mit Nahrung und Dosisschema entsprachen weder in Studie 101 noch vollständig in Studie 102 der Fachinformation; die Einnahme mit Nahrung war erst ab 2021 vorgegeben und die Regeln zur Wiederaufnahme nach leberwertbedingter Unterbrechung weichen von der Fachinformation ab. Die Einschätzung der EMA, Daten der nicht GMP-konformen Formulierungen ließen sich auf das zugelassene Arzneimittel übertragen [2], betrifft nur die Formulierung, nicht Dosierung und Behandlungsverlauf; Studien zur Dosisfindung oder Dosis-Wirkungs-Beziehung wurden nicht durchgeführt [2]. Die vorliegenden Daten dürften die zugelassene Anwendung daher nur eingeschränkt abbilden. Supportive Maßnahmen wurden nicht standardisiert eingesetzt und nicht erfasst; ein möglicher Einfluss auf die Endpunkte lässt sich nicht von dem der Behandlung trennen.

Bewertung der Validität und Aussagekraft der studienübergreifenden Auswertungen (ISE und ISS) und des Verzerrungspotentials

Die Aussagekraft einarmiger Daten hängt davon ab, ob die Population vorab definiert und vollständig erfasst ist, ob ein Ausgangszustand vor Behandlungsbeginn einheitlich erhoben wurde, ob die Endpunkte in allen Phasen gleich dokumentiert wurden und ob die Auswertung vor Kenntnis der Daten festgelegt war. Für die Gesamtpools der ISE und ISS sind diese Voraussetzungen nicht gegeben, und Selektionsbias, Immortal Time Bias, uneinheitlicher Ausgangsstatus, Behandlung mit überwiegend nicht GMP-konformen Formulierungen in individueller Dosierung und späte Präspezifikation gelten unverändert auch für den Kernpool der Studien 101 und 102 (siehe Abschnitt „Studiendesign“). Darüber hinaus bestehen für den Kernpool folgende Einschränkungen:

- **Datenerhebung:** Die Zeit vor Behandlungsbeginn ist ausschließlich retrospektiv dokumentiert, die Zeit danach in Studie 101 retrospektiv und in Studie 102 prospektiv. Unterschiede in Erhebungsdichte und Vollständigkeit können Veränderungen nach Behandlungsbeginn häufiger oder seltener erscheinen lassen als vor Behandlungsbeginn.
- **Endpunkte:** Motorische Meilensteine wurden nicht mit einem validierten Instrument erhoben, sondern aus Akten und Visiten abgeleitet; undatierte Ereignisse wurden nach festen Regeln einer Phase zugeordnet. Beginn und Beendigung von Beatmung und Sondenernährung lagen im ärztlichen Ermessen. UE schließen erkrankungsbezogene Ereignisse ein, schwere UE sind für den ISS-Kernpool nicht ausgewertet.
- **Fehlende Werte:** Wie häufig Ersetzungsregeln angewandt wurden, ist nicht berichtet; Sensitivitätsanalysen liegen nicht vor.
- **Fallzahl:** Die Nenner sind klein und unterscheiden sich je Endpunkt, sodass einzelne Personen die Anteile deutlich verändern.

Die Ergebnisse des Kernpools der gemäß EMA pivotalen Studien 101 und 102 sind daher nicht hinreichend valide, um daraus einen Effekt von Doxecitin/Doxribtimin abzuleiten; die Darstellung im Anhang dient allein der Beschreibung der Daten. Das Verzerrungspotential ist auf Studien- und Endpunktebene hoch (siehe Kapitel 2.5).

Indirekter Vergleich

Den Vergleich mit unbehandelten Personen einer historischen Kontrollgruppe legt der pU als Beleg für einen Überlebensvorteil vor. Da die anwendungsgebietsbezogene ISE-Population (N = 82) für die Nutzenbewertung nicht herangezogen wird, werden auch der indirekte Vergleich für diese Population und die ergänzenden Analysen einschließlich der TTE nicht berücksichtigt. Der indirekte Vergleich wird im Anhang beschrieben und die Limitationen, die über die in Kapitel 2.2.2 beschriebenen Limitationen der behandelten Population (z. B. Selektionsbias) hinausgehen, werden dort ausführlich dargelegt. Auch die EMA hat die Überlebensdaten nicht in die Fachinformation aufgenommen [2].

4.2 Zulassungsstatus und Zulassungspopulation

Zulassungsstatus und Anwendungsgebiet

Doxecitin/Doxribtimin ist zugelassen zur Behandlung von pädiatrischen und erwachsenen Patientinnen und Patienten mit genetisch bestätigter TK2d und einem Alter bei Symptombeginn bis zur Vollendung des 12. Lebensjahres [13]. Die Einschränkung auf eine genetisch bestätigte Erkrankung wurde im Verfahren auf Anforderung des CHMP aufgenommen, um das Anwendungsgebiet an die untersuchte Population anzugleichen [2].

Die Zulassung erfolgte unter „außergewöhnlichen Umständen“. Das CHMP stuft die klinischen Daten als nicht umfassend ein und sieht es jedoch als hinreichend dargelegt an, dass umfassende Daten unter normalen Anwendungsbedingungen nicht beigebracht werden können. In seiner Begründung benennt das CHMP zugleich die Grenzen der Evidenz: Ein großer Teil der Wirksamkeitsdaten wurde retrospektiv erhoben, bei den behandelten Personen sei eine positive und bei den unbehandelten eine negative Selektion möglich, die Dosis-Wirkungs-Beziehung ist nicht quantifiziert und die Größe des Behandlungseffekts lasse sich nicht verlässlich schätzen [2].

Der ISE-Kernpool deckt das Anwendungsgebiet hinsichtlich des Alters bei Symptombeginn formal ab (alle 39 Personen laut Studienunterlagen mit Symptombeginn im Alter von ≤ 12 Jahren), bildet aber nur die durch die Studienaufnahme auf Behandlung und Überleben selektierten Personen ab (siehe Kapitel 4.1). Für die am schwersten betroffene Teilpopulation liegen jedoch kaum Daten vor: Je früher die TK2d beginnt, desto rascher verläuft sie. In einer retrospektiven Verlaufsstudie (N = 92) betrug bspw. das mediane Überleben nach Symptombeginn bei infantilem Beginn (≤ 1 Jahr) 1,0 Jahr (95%-KI: [0,58; 2,33]), bei Beginn im Kindesalter dagegen mindestens 13 Jahre; Todesursache war eine respiratorische Insuffizienz [3]. Auch die EMA bezeichnet Kinder mit Symptombeginn im Alter von ≤ 2 Jahren als die am schwersten betroffene Teilgruppe, in der ein Behandlungseffekt am bedeutsamsten und am ehesten zu beurteilen wäre [2]. Zwar hatten im ISE-Kernpool gemäß den vorliegenden Angaben 21 Personen (53,8 %) einen Symptombeginn im Alter von ≤ 2 Jahren, die Behandlung begann jedoch nur bei 9 Personen in diesem Alter, die Erhebung des Symptombeginns ist nicht bewertbar und da die Aufnahme in die Studien ein Überleben bis mindestens zum 01.06.2018 voraussetzte, dürften gerade Kinder mit rasch progredientem, früh tödlichem Verlauf unterrepräsentiert sein (siehe Kapitel 4.1). Für die Patientengruppe mit dem schwersten Verlauf und dem höchsten Therapiebedarf erlauben die vorliegenden Daten daher vermutlich kaum Aussagen. Der späteste Behandlungsbeginn im ISE-Kernpool lag bei 35,5 Jahren; für einen Behandlungsbeginn im Erwachsenenalter (6 Personen im Alter von ≥ 18 Jahren) ist die Evidenz ebenfalls sehr begrenzt.

Eine bestätigte Mutation im TK2-Gen war Einschlusskriterium der Studien 101 und 102 und Teil der Erhebungen; auf dieses Kriterium geht die Beschränkung der Zulassung auf Personen mit genetisch bestätigter TK2d zurück, während die Altersgrenze (Symptombeginn im Alter von ≤ 12 Jahren) auf die Zulassungsentscheidung zurückgeht [2]. Eindeutige Angaben dazu, wie viele Personen des ISE-Kernpools tatsächlich eine genetische Bestätigung der TK2d aufwiesen und welche Mutationen

vorlagen, wurden nicht vorgelegt. Die vorliegenden Angaben legen nahe, dass der überwiegende Teil oder alle eingeschlossenen Personen eine genetisch bestätigte TK2d aufwiesen; infolge der fehlenden Angaben lässt sich dies jedoch nicht abschließend beurteilen, sodass Abweichungen von der Zulassungspopulation nicht gänzlich ausgeschlossen werden können.

Aufgrund des Selektionsbias sind Abweichungen zur Zulassungspopulation wahrscheinlich. Zudem bestehen wesentliche Unklarheiten zur Bestimmung des Alters bei Symptombeginn und zum Nachweis der genetischen Bestätigung. Die Übertragbarkeit der Ergebnisse auf die Zulassungspopulation erscheint daher in Teilen eingeschränkt, kann aufgrund der vorliegenden Informationslage jedoch nicht abschließend bewertet werden.

Übertragbarkeit auf deutschen Versorgungskontext

Ob Diagnose und Patientencharakteristika auf den deutschen Versorgungskontext übertragbar sind, lässt sich anhand der vorliegenden Angaben nicht beurteilen.

Die TK2d wurde erst 2001 als eigenständige Erkrankung beschrieben [7]. Die Personen des Kernpools wurden zwischen 1983 und 2019 geboren, sodass die Diagnosen über einen langen Zeitraum und mit sich wandelnden diagnostischen Möglichkeiten gestellt wurden. Wann und auf welchem Weg die Diagnose jeweils gesichert wurde, ist den vorliegenden Unterlagen nicht zu entnehmen; das in den Studien 101 und 102 erhobene Datum der genetischen Testung und die nachgewiesenen Mutationen werden in ISE und ISS nicht ausgewertet. Die Zeit von Symptombeginn bis zur Diagnose reichte von −60 bis 360 Monaten, zur Zeit vom Symptombeginn bis zum Behandlungsbeginn liegen nur kategorisierte Angaben vor (siehe Kapitel 4.1 und Anhang). Dabei müssen auch die Unsicherheiten zur Bestimmung des Alters bei Symptombeginn berücksichtigt werden. Das Alter bei Diagnose lag im Median bei 35,2 Monaten (Spannweite: 1–423; siehe Anhang); wie es bestimmt wurde und ob es sich auf die genetische Bestätigung der TK2d oder auf eine klinische Diagnose bezieht, ist nicht berichtet. Die über den genetischen Nachweis hinausgehende diagnostische Charakterisierung ist lückenhaft (siehe Anhang, Abschnitt „Informationen zur Diagnose“). Es bleibt daher unklar, inwieweit die Personen des Kernpools der Population entsprechen, die heute in Deutschland mit genetisch bestätigter TK2d diagnostiziert wird. Eine heute frühere Diagnosestellung könnte zudem mit einem anderen Krankheitsstadium bei Behandlungsbeginn einhergehen.

Weitere Einschränkungen ergeben sich daraus, dass ein Drittel der Personen außerhalb Europas und Nordamerikas lebte, dass die supportive Versorgung (Beatmung, Ernährung, Physiotherapie) mangels Angaben zu Begleittherapien nicht beurteilt werden kann und sich über den langen Beobachtungszeitraum verändert haben dürfte sowie dass die Behandlung überwiegend nicht der Anwendung gemäß Fachinformation entsprach. Die Übertragbarkeit der Ergebnisse auf den deutschen Versorgungskontext ist damit unklar.

4.3 Zusammenfassende Beurteilung zum Verzerrungspotential und zur Ergebnissicherheit

Das Verzerrungspotential wird auf Studien- und Endpunktebene als hoch eingestuft. Maßgeblich sind die fehlende Kontrollgruppe, die Selektion auf Überleben und fortgesetzte Behandlung, die überwiegend retrospektive und nicht standardisierte Datenerhebung sowie die erst in Kenntnis der Daten festgelegten Auswertungsregeln der Analysepläne der ISE und ISS. Diese Limitationen betreffen nicht nur die gepoolten Auswertungen, sondern auch den Kernpool.

Die Ergebnissicherheit ist so gering, dass die vorliegenden Daten keine Aussage zu einem Effekt von Doxecitin/Doxribtimin zulassen. Die Ergebnisse des ISE-/ISS-Kernpools (N = 39) werden im Anhang dargestellt.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Doxecitin/Doxribtimin ist zugelassen zur Behandlung von Kindern und Erwachsenen mit genetisch bestätigter Thymidinkinase-2-Defizienz (TK2d) und einem Alter bei Symptombeginn bis zur Vollendung des 12. Lebensjahres [13]. Zulassungsbegründend sind die Studien MT-1621-101 (Studie 101), ein retrospektiver Chart Review bereits behandelter Personen, und TK0102 (Studie 102), eine einarmige Fortsetzungsstudie; beide sind unkontrolliert.

Der pU legt die Ergebnisse in den studienübergreifenden Auswertungen ISE und ISS sowie in historischen Vergleichen mit unbehandelten Personen vor. Diese Auswertungen werden für die Nutzenbewertung nicht herangezogen (siehe Kapitel 2.2.2, 2.2.3 und 2.6). Dargestellt werden lediglich die Ergebnisse des zulassungsbegründenden ISE-/ISS-Kernpools der pivotalen Studien 101 und 102 (N = 39; Symptombeginn im Alter von ≤ 12 Jahren) im Anhang. Auch dieser umfasst nur Personen, die bis zur maßgeblichen Visite der Studie 101 bzw. bis zur Aufnahme in Studie 102 überlebt hatten und behandelt wurden, sodass Selektionsbias und Immortal Time Bias nicht ausgeschlossen werden können. Die Ergebnisse zu den Endpunkten „Gesamtüberleben“, „Motorische Meilensteine“, „Künstliche Beatmung“, „Künstliche Ernährung“ und „Unerwünschte Ereignisse“ werden ausschließlich deskriptiv dargestellt; Daten zur Lebensqualität liegen nicht vor.

Aus den vorliegenden Daten lässt sich keine Aussage zu einem Effekt von Doxecitin/Doxribtimin auf patientenrelevante Endpunkte ableiten.

Referenzen

1. **Berardo A, Dominguez-Gonzalez C, Engelstad K, Hirano M.** Advances in thymidine kinase 2 deficiency: clinical aspects, translational progress, and emerging therapies. J Neuromuscul Dis 2022;9(2):225-235. https://dx.doi.org/10.3233/JND-210786.

2. **European Medicines Agency (EMA).** Kygevvi: European public assessment report EMEA/H/C/005119/0000 [online]. 29.01.2026. Amsterdam (NED): EMA. [Zugriff: 10.09.2026]. URL: https://www.ema.europa.eu/en/documents/assessment-report/kygevvi-epar-public-assessment-report_en.pdf.

3. **Garone C, Taylor RW, Nascimento A, Poulton J, Fratter C, Dominguez-Gonzalez C, et al.** Retrospective natural history of thymidine kinase 2 deficiency. J Med Genet 2018;55(8):515-521. https://dx.doi.org/10.1136/jmedgenet-2017-105012.

4. **Lesko N, Naess K, Wibom R, Solaroli N, Nennesmo I, von Döbeln U, et al.** Two novel mutations in thymidine kinase-2 cause early onset fatal encephalomyopathy and severe mtDNA depletion. Neuromuscular Disorders 2010;20(3):198-203. https://dx.doi.org/10.1016/j.nmd.2009.11.013.

5. **Mancuso M, Salviati L, Sacconi S, Otaegui D, Camano P, Marina A, et al.** Mitochondrial DNA depletion: mutations in thymidine kinase gene with myopathy and SMA. Neurology 2002;59(8):1197-1202. https://dx.doi.org/10.1212/01.wnl.0000028689.93049.9a.

6. **Rizk JG.** When and why to use overlap weighting: clarifying its role, assumptions, and estimand in real-world studies. J Clin Epidemiol 2025;187:111942. https://dx.doi.org/10.1016/j.jclinepi.2025.111942.

7. **Saada A, Shaag A, Mandel H, Nevo Y, Eriksson S, Elpeleg O.** Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy. Nat Genet 2001;29(3):342-344. https://dx.doi.org/10.1038/ng751.

8. **Thomas LE, Li F, Pencina MJ.** Overlap weighting: a propensity score method that mimics attributes of a randomized clinical trial. JAMA 2020;323(23):2417-2418. https://dx.doi.org/10.1001/jama.2020.7819.

9. **UCB Pharma.** Data collection for patients in the company-supported expanded access program; Final report TK0114 [unveröffentlicht]. 12.09.2024.

10. **UCB Pharma.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4 A: Doxecitin und Doxribtimin (KYGEVVI), Behandlung von Kindern und Erwachsenen mit genetisch bestätigter Thymidinkinase-2-Defizienz (TK2d) mit einem Alter bei Symptombeginn bis zur Vollendung des 12. Lebensjahres; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 14.08.2026.

11. **UCB Pharma.** Doxecitine and doxribtimine: Integrated summary of efficacy; Includes studies MT-1621-101, TK0102, MT-1621-107, TK0110, TK0114, TK0112 [unveröffentlicht]. 25.09.2024.

12. **UCB Pharma.** Doxecitine and doxribtimine: Integrated summary of safety; Includes studies MT-1621-101, MT-1621-103, MT-1621-105, MT-1621-106, MT-1621-107, TK0102, TK0114 [unveröffentlicht]. 30.09.2024.

13. **UCB Pharma.** KYGEVVI 2 g/2 g, Pulver zur Herstellung einer Lösung zum Einnehmen [online]. Stand: 03.2026. Frankfurt/Main. [Zugriff: 10.09.2026]. URL: https://www.fachinfo.de/fi/detail/025765/kygevvi-2-g-2-g-pulver-zur-herstellung-einer-loesung-zum-einnehmen.

14. **UCB Pharma.** A phase 2 open-label study of continuation treatment with combination pyrimidine nucleosides in patients with thymidine kinase 2 (TK2) deficiency; interim clinical study report TK0102 [unveröffentlicht]. 01.07.2024

15. **UCB Pharma.** A retrospective study of subjects with thymidine kinase 2 deficiency treated with the combination of pyrimidine nucleos(t)ides as well as untreated subjects to collect vital status data and supporting information; Clinical study report MT-1621-107 [unveröffentlicht]. 08.06.2023.

16. **UCB Pharma.** A retrospective study of the combination of pyrimidine nucleos(t)ides in patients with thymidine kinase 2 (TK2) deficiency; Clinical study report MT-1621-101 [unveröffentlicht]. 31.03.2021.

17. **Wang J, Kim E, Dai H, Stefans V, Vogel H, Al Jasmi F, et al.** Clinical and molecular spectrum of thymidine kinase 2-related mtDNA maintenance defect. Mol Genet Metab 2018;124(2):124-130. https://dx.doi.org/10.1016/j.ymgme.2018.04.012.

7 Anhang

7.1 Studiencharakteristika und Studienmedikation

Vorbemerkung zur Einordnung der dargestellten Ergebnisse

Die nachfolgend dargestellten Charakteristika und Ergebnisse des ISE-/ISS-Kernpools (N = 39; Studien 101 und 102) werden ausschließlich deskriptiv wiedergegeben. Sie sind vor dem Hintergrund der in den Kapiteln 2.2.2, 2.2.3, 2.3 und 2.4 beschriebenen Limitationen zu lesen, die für alle Abschnitte dieses Anhangs gleichermaßen gelten und in den einzelnen Abschnitten nicht wiederholt werden:

- Selektionsbias und Immortal Time Bias: Beide Studien schlossen nur Personen ein, die nach Behandlungsbeginn bis zu einer Visite zwischen dem 01.06.2018 und dem 15.12.2018 (Studie 101) bzw. bis zur Aufnahme unter laufender und vertragener Behandlung (Studie 102) überlebt hatten. Personen mit frühem Tod oder Behandlungsabbruch fehlen daher und die Zeit bis zur Studienaufnahme geht als ereignisfreie Zeit in die Auswertung ein; der Kernpool dürfte eine prognostisch günstigere Auswahl der behandelten Personen abbilden.
- Datenerhebung und Ausgangsstatus: Die Zeit vor Behandlungsbeginn ist ausschließlich retrospektiv aus Krankenakten dokumentiert, die Zeit danach in Studie 101 retrospektiv und erst in Studie 102 prospektiv. Für die 10 Personen ohne Teilnahme an Studie 101 beruht der Ausgangsstatus auf anamnestischen Angaben, die bei Studieneintritt unter laufender Vorbehandlung erhoben wurden; ein unbehandelter Ausgangszustand liegt für keinen Endpunkt vor.
- Erhebung und fehlende Angaben: Beginn und Beendigung von Beatmung und Sondenernährung sowie die Beurteilung motorischer Meilensteine lagen im ärztlichen Ermessen und folgten keinen studienübergreifend einheitlichen Kriterien. Undatierte und fehlende Angaben wurden nach festen Regeln zugeordnet (fehlende Angaben zu Verlust bzw. Wiedererlangen als „nicht eingetreten“, Dokumentationslücken als durchgehende Unterstützung); wie häufig diese Regeln angewandt wurden, ist nicht berichtet.
- Behandlung und Beobachtungszeit: Der überwiegende Teil der Beobachtungszeit entfällt auf nicht GMP-konforme Formulierungen in individuell festgelegter Dosierung. Die Ergebnisse werden nicht zu festen Zeitpunkten, sondern über die gesamte, zwischen den Personen unterschiedlich lange Beobachtung nach Behandlungsbeginn zusammengefasst (Behandlungsdauer zum Datenschnitt vom 15.03.2024 im Median 78,4 Monate (Spannweite: 53,2–157,1); Monate zu rund 28 Tagen); Zeitpunkt und Dauer eines Ereignisses sind daraus nicht erkennbar. Patientencharakteristika, Exposition, Gesamtüberleben und UE beziehen sich auf den Datenschnitt vom 15.03.2024, die Morbiditätsendpunkte auf den Datenschnitt vom 14.03.2025, zu dem keine Angaben zur Behandlungsdauer vorliegen.
- Fehlende Vergleichsgruppe und kleine Teilgruppen: Ohne Kontrollgruppe und ohne unbehandelten Ausgangszustand lassen sich Veränderungen nach Behandlungsbeginn nicht vom natürlichen Verlauf – bei Kindern einschließlich der physiologischen motorischen Entwicklung – abgrenzen. Die Anteile beziehen sich auf kleine, über den Ausgangsstatus definierte Teilgruppen.
- Berichterstattung: Der pU weist für den Kernpool bei keinem Morbiditätsendpunkt fehlende Angaben aus; ob die Endpunkte tatsächlich bei allen 39 Personen erhoben wurden oder fehlende Dokumentation – insbesondere in Studie 101 – als „kein Ereignis“ gewertet wurde, ist nicht sichergestellt. Netto-Score, erstmaliges Erreichen und Anzahl der zu Behandlungsbeginn vorliegenden Meilensteine, schwere UE (CTCAE-Grad ≥ 3), UE nach SOC und eine Auswertung ohne erkrankungsbezogene Ereignisse liegen für den Kernpool nicht vor; Zeitpunkt und Dauer der Ereignisse sowie der Erhalt eines wiedererlangten Zustands bis zum Beobachtungsende werden nicht ausgewiesen.

Die Einordnungen in den einzelnen Abschnitten beschränken sich daher auf endpunktspezifische Aspekte. Keines der dargestellten Ergebnisse erlaubt eine Aussage zu einem Behandlungseffekt von Doxecitin/Doxribtimin.

Demographische und krankheitsspezifische Charakteristika

Der Symptombeginn lag im Median bei 1,89 Jahren (Spannweite: 0,31–11,7), bei 53,8 % im Alter von ≤ 2 Jahren; zur Unsicherheit bei der Bestimmung des Symptombeginns siehe Kapitel 2.2.2. Die Behandlung begann im Median erst mit 4,67 Jahren (Spannweite: 0,73–35,5); 9 Personen (23,1 %) wurden erstmals im Alter von ≤ 2 Jahren behandelt, 9 (23,1 %) erst im Alter von > 12 Jahren (6 davon im Erwachsenenalter) und bei 13 Personen (33,3 %) lagen mehr als fünf Jahre zwischen Symptom- und Behandlungsbeginn. Zu Behandlungsbeginn waren 53,8 % gehfähig, 46,2 % beatmet (überwiegend nicht invasiv) und 28,2 % sondenernährt; 66,7 % waren männlich, 43,6 % lebten in Europa.

Dass trotz überwiegend frühem Symptombeginn mehr als die Hälfte der Personen bei Behandlungsbeginn noch gehfähig war und die Behandlung bei einem Drittel erst mehr als fünf Jahre nach Symptombeginn, bei 6 Personen erst im Erwachsenenalter, begann, deutet angesichts des natürlichen Verlaufs der früh beginnenden TK2d – medianes Überleben nach infantilem Symptombeginn etwa ein Jahr [3] – darauf hin, dass die Aufnahmebedingungen (siehe Vorbemerkung) Personen mit langsamerem Verlauf begünstigt haben könnten. Der Kernpool scheint damit prognostisch günstiger zu sein als die Patientinnen und Patienten mit früh beginnender, rasch progredienter Erkrankung, für die in der Versorgung über eine Behandlung entschieden wird. Die vom pU als „Behandlungsverzögerung“ bezeichnete Zeit zwischen Symptom- und Behandlungsbeginn setzt zudem voraus, dass die Behandlung nach dem Symptombeginn begann; da das Alter bei Symptombeginn in keiner Studie definiert ist und bei familiär bekannter Erkrankung Diagnose und Behandlungsbeginn vor dem Symptombeginn gelegen haben können (siehe Abschnitt „Informationen zur Diagnose“), ist dies nicht gesichert.

Die Angaben zum Anteil der Sondennahrung sind zwischen ISS- und ISE-Bericht nicht konsistent: Laut ISS-Bericht summieren sich die Kategorien auf 40 Personen bei N = 39; vermutlich wird eine Person, deren Sonde vor Behandlungsbeginn entfernt wurde, sowohl unter „0 %“ als auch unter „nicht anwendbar“ gezählt. Der ISE-Bericht weist unter „0 %“ keine Person aus (siehe Fußnote 8 zu Tabelle 6).

Tabelle 6: Charakterisierung der Studienpopulation; ISE-/ISS-Kernpool, Datenschnitt: 15.03.2024

Studie 101 + 102 **Charakterisierung der Studienpopulation**	**ISE-/ISS-Kernpool** **N = 39**
Demographische Variablen	
Alter bei Symptombeginn[1)] (Jahre) n (%) MW (SD) Median (min; max)	 39 (100) 2,34 (2,03) 1,89 (0,31; 11,7)
Alter bei Symptombeginn[1)] nach Kategorien, n (%) ≤ 2 Jahre > 2 bis ≤ 12 Jahre ≤ 12 Jahre > 12 Jahre	 21 (53,8) 18 (46,2) 39 (100) 0 (0)
Alter bei Behandlungsbeginn (jegliche Behandlung) (Jahre)[2)] n (%) MW (SD) Median (min; max)	 39 (100) 8,57 (8,99) 4,67 (0,73; 35,5)
Alter bei Behandlungsbeginn nach Kategorien[2)] ≤ 2 Jahre > 2 bis ≤ 12 Jahre > 12 bis < 18 Jahre ≥ 18 Jahre ≤ 18 Jahre > 18 Jahre	 9 (23,1) 21 (53,8) 3 (7,7) 6 (15,4) 33 (84,6) 6 (15,4)
Geburtsjahr[3)] n (%) MW (SD) Median (min; max)	 39 (100) 2008 (9,46) 2011 (1983; 2019)
Geburtsjahr nach Kategorien, n (%)[3)] < 2000 ≥ 2000 nicht berichtet	 6 (15,4) 33 (84,6) 0 (0)
Geschlecht, n (%) männlich weiblich	 26 (66,7) 13 (33,3)
Ethnische Herkunft, n (%) weiß schwarz asiatisch Ureinwohner Amerikas und Alaskas	 36 (92,3) 1 (2,6) 1 (2,6) 1 (2,6)
Ethnizität, n (%) hispanisch oder lateinamerikanisch nicht hispanisch oder lateinamerikanisch	 14 (35,9) 25 (64,1)
Geographische Region des Wohnorts, n (%) Nordamerika Europa Rest der Welt	 9 (23,1) 17 (43,6) 13 (33,3)

Studie 101 + 102 Charakterisierung der Studienpopulation	ISE-/ISS-Kernpool N = 39
Zeit zwischen Symptom und Behandlungsbeginn[3)4)]	
≤ 6 Monate	6 (15,4)
> 6 Monate bis ≤ 1 Jahr	6 (15,4)
> 1 Jahr bis ≤ 5 Jahre	14 (35,9)
> 5 Jahre bis ≤ 10 Jahre	6 (15,4)
> 10 Jahre	7 (17,9)
Krankheitsspezifische Variablen	
Gehfähigkeit[5)], *n (%)*	
ja	21 (53,8)
nein	18 (46,2)
Mobilitätshilfe, n (%)	
Bettlägerig	1 (2,6)
Steh-/Gehhilfe[6)]	2 (5,1)
Fahrbare Mobilitätshilfen[7)]	15 (38,5)
Fehlend	21 (53,8)
Künstliche Beatmung erforderlich, n (%)	
ja	18 (46,2)
nein	21 (53,8)
Art der künstlichen Beatmung, n (%)	
invasiv (Tracheotomie)	4 (10,3)
nicht-invasiv (z. B. BiPAP, CPAP)	14 (35,9)
nicht anwendbar	21 (53,8)
Umfang der künstlichen Beatmung (Std./Tag)	
n	17
MW (SD)	13,6 (6,95)
Median (min; max)	10,0 (8; 24)
Künstliche Ernährung erforderlich, n (%)	
ja	11 (28,2)
nein	28 (71,8)
Nahrungsaufnahme über künstliche Ernährung (prozentualer Anteil/Tag)	
0 %	1 (2,6)[8)]
25 %	0 (0)
50 %	0 (0)
75 %	3 (7,7)
100 %	5 (12,8)
fehlend	3 (7,7)
nicht anwendbar	28 (71,8)

[1)] Eine Definition des Symptombeginns und das Erhebungsvorgehen sind in den Studienunterlagen nicht beschrieben; das Alter bei Symptombeginn wurde retrospektiv bzw. anamnestisch bestimmt. Zur Unsicherheit bei der Bestimmung und zu möglichen Fehlzuordnungen siehe Kapitel 2.2.2 und 2.3.2.

[2)] Das Alter bei der ersten Behandlung entspricht dem Alter zum Zeitpunkt der ersten Behandlung mit jeglicher Formulierung der Studienmedikation.

[3)] Angabe stammt aus dem ISE-Bericht. Angabe im ISS-Bericht liegt nicht vor.

[4)] Vom pU als „Behandlungsverzögerung" bezeichnet; hier als Zeit zwischen Symptom- und Behandlungsbeginn ausgewiesen. Ein Behandlungsbeginn nach dem Symptombeginn ist damit nicht sichergestellt: Das Alter bei Symptombeginn ist nicht valide bestimmbar (siehe Kapitel 2.2.2) und bei familiären Fällen kann die Diagnose vor dem Symptombeginn gestellt worden sein (Zeit vom Symptombeginn bis zur Diagnose bis −60 Monate); ob ein Behandlungsbeginn vor dem Symptombeginn vorkam, ist aus den Kategorien nicht ersichtlich.

[5)] Als gehfähig galten Personen, die den Meilenstein „Gehen ohne Unterstützung“ zu Behandlungsbeginn erreicht hatten. 9 Kinder waren bei Behandlungsbeginn ≤ 2 Jahre alt.

[6)] Steh- und Gehhilfen umfassen Gehstöcke, Stehständer, Kinderwagen und Rollatoren. 9 Kinder waren bei Behandlungsbeginn ≤ 2 Jahre alt.

[7)] Fahrbare Mobilitätshilfen umfassen Rollstühle, motorisierte Rollstühle, Elektromobile, Kinderwagen und Skateboards. 9 Kinder waren bei Behandlungsbeginn ≤ 2 Jahre alt.

[8)] Angabe gemäß ISS-Unterlagen. In der Summe ergeben die Kategorien 40 Personen bei N = 39: Den 11 Personen mit künstlicher Ernährung zu Behandlungsbeginn stehen 12 Personen in den Kategorien „0 %“ bis „fehlend“ gegenüber. Die Kategorien beziehen sich vermutlich auf die 12 Personen mit einer Sonde vor Behandlungsbeginn, einschl. 1 Person, deren Sonde vor Behandlungsbeginn entfernt wurde und die unter „0 %“ geführt wird; „nicht anwendbar“ bezieht sich auf die Personen ohne Sonde zu Behandlungsbeginn. Im ISE-Bericht ist für die Kategorie „0 %“ keine Person angegeben. Die Bezugsgruppen sind in den Unterlagen nicht erläutert.

Abkürzungen: BiPAP: Bilevel Positive Airway Pressure; CPAP: Continuous Positive Airway Pressure; ISE: Integrierte Wirksamkeitsanalyse; ISS: Integrierte Sicherheitsauswertung; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung.

Informationen zur Diagnose

Eine bestätigte Mutation im TK2-Gen war Einschlusskriterium beider Studien und wurde erhoben. Angaben dazu, wie viele Personen sie tatsächlich aufwiesen und welche Mutationen vorlagen, wurden jedoch nicht vorgelegt. Berichtet ist zwar das Alter bei Diagnose (im Median 35,2 Monate (Spannweite: 1–423); vom pU als „Zeit von der Geburt bis zur Diagnose“ ausgewiesen), aber ohne Angabe, ob es sich auf die genetische Bestätigung oder eine klinische Diagnose bezieht. Die Diagnose wurde im Median 8,9 Monate nach Symptombeginn gestellt, bei einzelnen Personen jedoch bis zu 60 Monate davor. Die über den genetischen Nachweis hinausgehende Charakterisierung ist lückenhaft: Muskelbiopsiebefunde fehlen bei 15 Personen (38,5 %), eine mtDNA-Deletion wurde bei 21 Personen (53,8 %) nicht untersucht, Angaben zum betroffenen Atmungskettenkomplex fehlen bei 29 Personen und die TK2-Aktivität liegt nur für 3 Personen vor; einzelne Angaben weichen zwischen ISE- und ISS-Unterlagen voneinander ab (siehe Fußnoten zu Tabelle 7). Angaben zu den Mutationen könnten ergänzende Hinweise zur Krankheitsschwere liefern, da einzelne Varianten mit einem milderen oder schwereren Verlauf assoziiert zu sein scheinen [4, 17], auch wenn der Zusammenhang zwischen Genotyp und Phänotyp insgesamt schwach ist [2].

10 Personen (25,6 %) waren keine Indexfälle, in ihrer Familie war die Erkrankung also bereits zuvor diagnostiziert worden. 22 Personen (56,4 %) hatten kein weiteres an TK2d erkranktes Familienmitglied, bei 1 Person fehlt die Angabe, bei den übrigen waren Bruder (9), Schwester (7) oder andere Angehörige (5) betroffen (Mehrfachnennungen möglich). Erfasst wurden nur Angehörige, die ebenfalls in das klinische Entwicklungsprogramm eingeschlossen waren, sodass die familiäre Häufung unterschätzt sein dürfte. Geschwister im Kernpool sind zudem nicht unabhängig voneinander; gemeinsame genetische und familiäre Faktoren könne ihren Verlauf ähnlich beeinflussen, was bei einer Population von 39 Personen von Relevanz sein kann.

Für die Bewertung folgt daraus zweierlei: Weder lässt sich sicherstellen, dass alle 39 Personen die Kriterien des Anwendungsgebiets – genetisch bestätigte TK2d und Symptombeginn im Alter von ≤ 12 Jahren – erfüllten, noch beurteilen, inwieweit sie die heute mit genetisch bestätigter TK2d diagnostizierte Population abbilden (siehe Kapitel 4.2). Bei familiär bekannter Erkrankung ist außerdem eine Diagnose vor dem eigenen Symptombeginn und damit ein präsymptomatischer Behandlungsbeginn zumindest denkbar; ob dies vorkam, ist nicht berichtet. In diesem Fall wäre nicht nur die Angabe zur Zeit zwischen Symptom- und Behandlungsbeginn hinfällig, auch der unter Behandlung beobachtete Symptombeginn entspräche nicht dem natürlichen Krankheitsbeginn, auf den sich das Anwendungsgebiet bezieht.

Tabelle 7: Charakterisierung der Studienpopulation mit Informationen zur Diagnose; ISE-/ISS-Kernpool, Datenschnitt: 15.03.2024

Studie 101 + 102 **Informationen zur Diagnose**	**ISE-/ISS-Kernpool** **N = 39**
Genetische Bestätigung der TK2d, n (%)	k. A.
Mutationen im TK2-Gen, n (%)	k. A.
Zeit bis zur Diagnose seit Symptombeginn (Monate)	
n	39 (100)
MW (SD)	46,9 (81,4)
Median (min; max)	8,9 (-60; 360)
Alter bei Diagnose (Monate)[1]	
n	39 (100)
MW (SD)	75,0 (88,3)
Median (min; max)	35,2 (1; 423)
Befunde von Muskelbiopsien, n (%)	
Cox-Defizienz	6 (15,4)
Ragged-Red Fibers	5 (12,8)
Cox-Defizienz und Ragged-Red Fibers	11 (28,2)
Nekrotisierende Myopathie mit dystrophischen Merkmalen	1 (2,6)
Andere	1 (2,6)
Fehlend	15 (38,5)
mtDNA (%)	
n	15 (38,5)
MW (SD)	38,9 (42,0)
Median (min; max)	18,0 (8; 158)
TK2-Aktivität (%)	
n	3 (7,69)
MW (SD)	27,7 (14,0)
Median (min; max)	32,0 (12; 39)
mtDNA-Deletion, n (%)	
ja	6 (15,4)
nein	12 (30,8)
nicht getestet	21 (53,8)
Familienmitglieder mit TK2d[2)]*, n (%)*	
Bruder	9 (23,1)
Schwester	7 (17,9)
Andere	5 (12,8)
Keine	22 (56,4)
Fehlend	1 (2,6)
Indexpatientin oder -patient[3)]*, n (%)*	
ja	29 (74,4)
nein	10 (25,6)
Atmungskettenenzyme verringert, n (%)	
ja	10 (25,6)
nein	9 (23,1)
unbekannt	15 (38,5)
fehlend	5 (12,8)

Studie 101 + 102 Informationen zur Diagnose	ISE-/ISS-Kernpool N = 39
Atmungskettenkomplex, n (%)	
Komplex I	1 (2,6)
Komplex IV	1 (2,6)
Komplex I + III, Komplex II + III	2 (5,1)
Komplex I, Komplex III, Komplex IV	1 (2,6)
Komplex I, Komplex III, Komplex IV, Komplex II + III	1 (2,6)
Komplex I, Komplex III, Komplex IV, Komplex I + III, Komplex II + III	3 (7,7)
Komplex IV, Komplex I + III, Komplex II + III	1 (2,6)
Fehlend	29 (74,4)

[1)] Vom pU als „Zeit von der Geburt bis zur Diagnose" ausgewiesen; entspricht dem Alter bei Diagnose. Wie und wann die Diagnose gestellt wurde – klinisch oder genetisch, vor oder nach dem Symptombeginn – und ob es sich um den Zeitpunkt der genetischen Bestätigung der TK2d handelt, ist nicht berichtet.

[2)] Die Prozentanteile können sich zu mehr als 100 % aufsummieren, da 1 Person mehr als ein mit TK2d diagnostiziertes Familienmitglied haben kann. Das Verwandtschaftsverhältnis wurde nur für Personen mit mind. einem Familienmitglied erhoben, das eine TK2d-Diagnose erhielt und in das klinische Entwicklungsprogramm eingeschlossen wurde. „Keine" umfasst Personen ohne Familienmitglieder, die mit TK2d diagnostiziert wurden. „Fehlend" umfasst Personen, für die keine Informationen verfügbar sind. „Fehlend" umfasst ebenso Personen, deren Familienmitglieder mit TK2d diagnostiziert, aber nicht Teil des klinischen Entwicklungsprogramms sind.

[3)] Definiert als erste mit TK2d diagnostizierte Person in einer Familie.

Abkürzungen: ISE: Integrierte Wirksamkeitsanalyse; ISS: Integrierte Sicherheitsauswertung; k. A.: keine Angabe; mtDNA: mitochondriale DNA; MW: Mittelwert; pU: pharmazeutischer Unternehmer; SD: Standardabweichung; TK2: Thymidinkinase 2; TK2d: Thymidinkinase-2-Defizienz

Vortherapien und Begleitmedikation

Für den ISE-/ISS-Kernpool (N = 39) mit Symptombeginn im Alter von ≤ 12 Jahren liegen keine Daten zu Vortherapien und Begleitmedikationen vor.

Behandlungsexposition

Angaben zur Exposition liegen für den Kernpool nur zum Datenschnitt vom 15.03.2024 vor; für den Datenschnitt vom 14.03.2025, auf dem die Morbiditätsergebnisse beruhen, ist die Behandlungsdauer nicht ausgewiesen. Die Exposition betrug im Median 78,4 Monate (Spannweite: 53,2–157,1; Monate zu rund 28 Tagen), bei 31 Personen (79,5 %) mindestens 5 Jahre; die durchschnittlich erhaltene Tagesdosis lag im Median bei 393,6 mg/kg/Tag je Komponente (Spannweite: 169,6–400,0). Wie viele Personen wie lange welche Formulierung (dC/dT, dCMP/dTMP oder DC/DT) erhielten, ist für den Kernpool nicht ausgewiesen.

Die lange und nach unten begrenzte Exposition (kürzeste Behandlungsdauer 53,2 Monate) dürfte vor allem die Aufnahmebedingungen widerspiegeln (siehe Abschnitt „Vorbemerkung zur Einordnung der dargestellten Ergebnisse") und lässt sich daher nicht als Hinweis auf Verträglichkeit oder Wirksamkeit deuten. Der überwiegende Teil der Beobachtungszeit entfällt zudem auf nicht GMP-konforme Formulierungen in individuell festgelegter Dosierung und die untere Grenze der Dosisspannweite deutet darauf hin, dass ein Teil der Personen deutlich geringer als mit der Erhaltungsdosis laut Fachinformation dosiert wurde; Gründe hierfür (Titration, Unverträglichkeit, Unterbrechungen) sind nicht berichtet (siehe Kapitel 4.1).

Tabelle 8: Angaben zur Exposition mit der Studienmedikation; ISE-/ISS-Kernpool, Datenschnitt: 15.03.2024

Studie 101 + 102 **Exposition mit der Studienmedikation**	**ISE-/ISS-Kernpool** **N = 39**
Dauer der Exposition (Monate)[1)]	
MW (SD)	87,1 (27,5)
Median (min; max)	78,4 (53,2; 157,1)
Dauer der Studienbehandlung (Monate)	
MW (SD)	86,2 (27,0)
Median (min; max)	78,4 (53,2; 157,1)
Gesamtzahl der Personenjahre mit Exposition (Person Years) (Jahre)	
MW (SD)	6,63 (2,08)
Median (min; max)	6,03 (4,09; 12,08)
Behandlungsdauer nach Kategorien, n (%)	
Mindestens 1 Dosis	39 (100)
2 Jahre bis < 5 Jahre	8 (20,5)
5 Jahre oder mehr	31 (79,5)
Durchschnittlich erhaltene Tagesdosis (mg/kg/Tag)	
MW (SD)	359,9 (60,4)
Median (min; max)	393,6 (169,6; 400,0)
Durchschnittlich erhaltene Tagesdosis (mg/Tag)	
MW (SD)	10.049,8 (5.171,8)
Median (min; max)	8.898,8 (3.099,1; 31.508,7)

[1)] Monate: Einheiten zu 365,25/13 ≈ 28,1 Tagen (rund 28 Tage, keine Kalendermonate); Personenjahre: Kalenderjahre; 78,4 Monate entsprechen 6,03 Jahren.

Abkürzungen: ISE: Integrierte Wirksamkeitsanalyse; ISS: Integrierte Sicherheitsauswertung; MW: Mittelwert; SD: Standardabweichung

7.2 Mortalität

Bis zum Datenschnitt vom 15.03.2024 wurde für keine der 39 Personen des ISE-Kernpools ein Todesfall berichtet (38 Personen bis zum Datenschnitt beobachtet; für die eine nicht in Studie 102 übernommene Person endet die Beobachtung mit der letzten Visite in Studie 101, spätestens am 15.12.2018). Das Alter beim letzten bekannten Lebenszeichen lag im Median bei 12,35 Jahren (Spannweite: 5,1–41); Angaben zum Datenschnitt vom 14.03.2025 liegen auf Ebene der ISE nicht vor; laut Studienbericht 102 waren die 38 dort erfassten Personen des Kernpools zu diesem Datenschnitt weiterhin in der Studie, ein Todesfall ist nicht berichtet (siehe Kapitel 2.2.2).

Das Ausbleiben von Todesfällen ist vor dem Hintergrund der Aufnahmebedingungen zu sehen (siehe Abschnitt „Vorbemerkung zur Einordnung der dargestellten Ergebnisse“): Alle Personen mussten vom Behandlungsbeginn bis zur letzten Visite in Studie 101 bzw. bis zur Aufnahme in Studie 102 überlebt haben, und die kürzeste Expositionsdauer beträgt 53,2 Monate; Todesfälle in den ersten Behandlungsjahren, in denen die Sterblichkeit bei frühem Symptombeginn am höchsten ist, konnten damit designbedingt kaum erfasst werden. Bei infantilem Symptombeginn liegt das mediane Überleben im natürlichen Verlauf bei etwa einem Jahr [3]. Wie viele der 39 Personen einen infantilen Symptombeginn (≤ 1 Jahr) hatten, ist nicht ausgewiesen; der ISE-Bericht unterscheidet nur einen Symptombeginn im Alter von ≤ 2 Jahren (21 Personen, 53,8 %) und von > 2 bis ≤ 12 Jahren (18 Personen); das niedrigste Alter bei Symptombeginn betrug 0,31 Jahre. Die bereits dargelegten Unklarheiten zum Alter bei Symptombeginn sind zu berücksichtigen.

Das Ausbleiben von Todesfällen scheint daher vor allem die Aufnahmebedingungen widerzuspiegeln und lässt sich nicht als Überlebensvorteil interpretieren; auch das CHMP hat eine entsprechende Angabe nicht in die Fachinformation aufgenommen [2].

7.3 Morbidität

Funktionsbereich Motorik – Motorische Meilensteine

Von 31 Personen, die für ein Wiedererlangen infrage kamen, erlangten 26 (83,9 %) mindestens einen Meilenstein wieder; je Meilenstein zwischen 37,5 % (Gehen bzw. Treppensteigen ohne Unterstützung) und 88,2 % (Kopf aufrecht halten). Von 38 Personen, die für einen Verlust infrage kamen, verloren 10 (26,3 %) mindestens einen Meilenstein, je Meilenstein zwischen 0 und 13,8 % (Stehen ohne Unterstützung).

Die Bezugsgruppen sind klein, undatierte und fehlende Angaben wurden regelbasiert zugeordnet, ob wiedererlangte Fähigkeiten bis zum Beobachtungsende erhalten blieben, ist nicht dargestellt, und Netto-Score, erstmaliges Erreichen sowie die Anzahl der zu Behandlungsbeginn vorliegenden Meilensteine liegen für den Kernpool nicht vor (siehe Abschnitt „Vorbemerkung zur Einordnung der dargestellten Ergebnisse“ und Kapitel 2.3.2). Vor diesem Hintergrund deuten die Ergebnisse allenfalls darauf hin, dass unter Behandlung vor allem einfache Fähigkeiten wie Kopfkontrolle und Sitzen wiedererlangt wurden, freies Gehen, freies Treppensteigen und Laufen dagegen nur bei etwa einem Drittel bis zwei Fünfteln (37,5–41,2 %) der jeweils infrage kommenden Personen, und dass zugleich gut jede vierte Person mindestens einen Meilenstein verlor. Ob und in welchem Umfang diese Verläufe auf die Behandlung zurückgehen, lässt sich nicht beurteilen. Auswertungen zu Verlust und Wiedererlangen vor Behandlungsbeginn liegen für den ISE-Kernpool zwar ebenfalls vor (siehe Kapitel 2.4), ein belastbarer Vorher-Nachher-Vergleich ist jedoch wegen der ausschließlich retrospektiven Erfassung vor Behandlungsbeginn, der fehlenden Angaben zu den zu Behandlungsbeginn vorliegenden Meilensteinen und des fehlenden Netto-Scores nicht möglich.

Der Status vor Behandlungsbeginn ist ausschließlich retrospektiv aus Krankenakten erfasst, die Zeit danach teils retrospektiv (Studie 101), teils prospektiv (Studie 102); eine dichtere Dokumentation unter Studienbedingungen könnte Veränderungen nach Behandlungsbeginn häufiger sichtbar machen. Da die Häufigkeit der regelbasierten Zuordnungen undatierter und fehlender Angaben nicht berichtet ist, könnten diese die Ergebnisse in beide Richtungen beeinflussen. Für die 10 Personen ohne vorherige Teilnahme an Studie 101 beruht der Ausgangsstatus vermutlich auf anamnestischen Angaben, die bei Studieneintritt unter laufender Behandlung erhoben wurden. Bei Kindern mit frühem Symptombeginn fällt die Behandlung zudem in eine Phase physiologischer motorischer Entwicklung, sodass Veränderungen motorischer Fähigkeiten nur schwer vom natürlichen Verlauf abzugrenzen sind.

Tabelle 9: Ergebnisse zu „Anzahl wiedererlangter motorischer Meilensteine nach Behandlungsbeginn“ und „Wiedererlangen des motorischen Meilensteins nach Behandlungsbeginn“; ISE-Kernpool, Datenschnitt: 14.03.2025

Studien 101 + 102[1)] **Kategorie** Item	**ISE-Kernpool** **N = 39** *n (%)*
Nicht für das Wiedererlangen infrage kommend[2)3)]	8 (20,5)
Für das Wiedererlangen infrage kommend (N_1)[3)4)]	31 (79,5)
Anzahl wiedererlangter motorischer Meilensteine nach Behandlungsbeginn[5)]	*n (%)*
≥ 1	26 (83,9)
1[6)]	3 (9,7)
2[6)]	13 (41,9)
3[6)]	3 (9,7)
≥ 4[6)]	7 (22,6)
Art der wiedererlangten motorischer Meilensteine nach Behandlungsbeginn[7)8)]	*n/N_x (%)*
Kopf aufrecht halten ohne Unterstützung	15/17 (88,2)
Aufrecht sitzen ohne Unterstützung	10/14 (71,4)
Stehen mit Unterstützung	8/15 (53,3)
Stehen ohne Unterstützung	7/15 (46,7)
Gehen mit Unterstützung	9/16 (56,3)
Gehen ohne Unterstützung	6/16 (37,5)
Treppensteigen mit Unterstützung	9/19 (47,4)
Treppensteigen ohne Unterstützung	6/16 (37,5)
Laufen	7/17 (41,2)

[1)] Behandlungsbeginn = ISE-Tag 1, d. h. die erste dokumentierte Gabe einer berücksichtigten Formulierung (dC/dT, dCMP/dTMP oder DC/DT). Sie liegt bei allen Personen des ISE-Kernpools vor der Aufnahme in Studie 101 bzw. 102; der Behandlungsbeginn ist daher nicht mit dem Studieneintritt oder Studientag 1 der Studie 102 gleichzusetzen (siehe Kapitel 2.2.2). „Vor Behandlungsbeginn“: Zeitraum vom Symptombeginn bis ISE-Tag 1 (ausschließlich retrospektiv erfasst). „Nach Behandlungsbeginn“: Zeitraum von ISE-Tag 1 bis zum Ende der individuellen Beobachtung, d. h. bis zur letzten dokumentierten Angabe vor dem Datenschnitt vom 14.03.2025 (individuell unterschiedlich lang; Behandlungsdauer zum Datenschnitt vom 15.03.2024 im Median 78,4 Monate (Spannweite: 53,2–157,1), Monate zu 28 Tagen; zum Datenschnitt vom 14.03.2025 nicht berichtet). Ein Wiedererlangen zählt, wenn es zu irgendeinem Zeitpunkt des jeweiligen Zeitraums dokumentiert ist. Die Bezugsgruppen (infrage kommende Personen insgesamt bzw. je Meilenstein) ergeben sich aus den vor und nach Behandlungsbeginn dokumentierten Ereignissen (siehe Fußnoten zu „nicht infrage kommend“ und „infrage kommend“ sowie Kapitel 2.3.2). Zeitpunkt des Wiedererlangens und Erhalt eines wiedererlangten Meilensteins bis zum Beobachtungsende werden nicht ausgewiesen; undatierte Ereignisse und fehlende Angaben wurden nach festen Regeln zugeordnet (siehe Kapitel 2.4).

[2)] Nicht infrage kommend: Personen, die vor Behandlungsbeginn keinen MM verloren hatten, der nicht wiedererlangt wurde, und nach Behandlungsbeginn keinen MM verloren.

[3)] Die Prozentanteile werden unter Verwendung der Anzahl N der behandelten Personen berechnet.

[4)] Infrage kommend: Personen, die vor Behandlungsbeginn mind. einen MM verloren und nicht wiedererlangt hatten oder nach Behandlungsbeginn einen MM verloren.

[5)] Die Prozentanteile werden unter Verwendung der Anzahl N_1 der behandelten Personen berechnet, die MM potentiell wiedererlangen können.

[6)] Die spezifischen Anzahlen der wiedererlangten MM schließen sich gegenseitig aus.

[7)] Eine behandelte Person konnte bei mehr als einem MM berücksichtigt werden.

[8)] Die Prozentsätze basieren auf der Anzahl N_x der behandelten Personen, die jeweils einen MM potentiell wiedererlangen können, wobei maximal N_1 (= Anzahl der behandelten Personen, die den MM potentiell wiedererlangen können) berücksichtigt wird.

Abkürzungen: dC: Desoxycytidin; DC: Doxecitin; dCMP: Desoxycytidinmonophosphat; dT: Desoxythymidin; DT: Doxribtimin; dTMP: Desoxythymidinmonophosphat; ISE: Integrierte Wirksamkeitsanalyse; MM: Motorischer Meilenstein.

Tabelle 10: Ergebnisse zu „Anzahl verlorener motorischer Meilensteine nach Behandlungsbeginn“ und „Verlust des motorischen Meilensteins nach Behandlungsbeginn“; ISE-Kernpool, Datenschnitt: 14.03.2025

Studien 101 + 102[1)] **Kategorie** Item	**ISE-Kernpool** **N = 39** *n (%)*
Für den Verlust nicht infrage kommend[2)3)]	1 (2,6)
Für den Verlust infrage kommend (N_1)[3)4)]	38 (97,4)
Anzahl verlorener motorischer Meilensteine nach Behandlungsbeginn[5)]	*n (%)*
≥ 1	10 (26,3)
1[6)]	6 (15,8)
2[6)]	3 (7,9)
3[6)]	0 (0)
≥ 4[6)]	1 (2,6)
Verlorene motorische Meilensteine nach Behandlungsbeginn[7)8)]	*n/N_x (%)*
Kopf aufrecht halten ohne Unterstützung	1/38 (2,6)
Aufrecht sitzen ohne Unterstützung	1/36 (2,8)
Stehen mit Unterstützung	3/31 (9,7)
Stehen ohne Unterstützung	4/29 (13,8)
Gehen mit Unterstützung	3/30 (10,0)
Gehen ohne Unterstützung	1/27 (3,7)
Treppensteigen mit Unterstützung	2/26 (7,7)
Treppensteigen ohne Unterstützung	0/20 (0)
Laufen	2/20 (10,0)

[1)] Behandlungsbeginn = ISE-Tag 1, d. h. die erste dokumentierte Gabe einer berücksichtigten Formulierung (dC/dT, dCMP/dTMP oder DC/DT). Sie liegt bei allen Personen des ISE-Kernpools vor der Aufnahme in Studie 101 bzw. 102; der Behandlungsbeginn ist daher nicht mit dem Studieneintritt oder Studientag 1 der Studie 102 gleichzusetzen (siehe Kapitel 2.2.2). „Vor Behandlungsbeginn“: Zeitraum vom Symptombeginn bis ISE-Tag 1 (ausschließlich retrospektiv erfasst). „Nach Behandlungsbeginn“: Zeitraum von ISE-Tag 1 bis zum Ende der individuellen Beobachtung, d. h. bis zur letzten dokumentierten Angabe vor dem Datenschnitt vom 14.03.2025 (individuell unterschiedlich lang; Behandlungsdauer zum Datenschnitt vom 15.03.2024 im Median 78,4 Monate (Spannweite: 53,2–157,1), Monate zu 28 Tagen; zum Datenschnitt vom 14.03.2025 nicht berichtet). Ein Verlust zählt, wenn er zu irgendeinem Zeitpunkt des jeweiligen Zeitraums dokumentiert ist. Die Bezugsgruppen (infrage kommende Personen insgesamt bzw. je Meilenstein) ergeben sich aus den vor und nach Behandlungsbeginn dokumentierten Ereignissen (siehe die Fußnoten zu „nicht infrage kommend“ und „infrage kommend“ sowie Kapitel 2.3.2). Zeitpunkt des Verlusts und ein etwaiges späteres Wiedererlangen bis zum Beobachtungsende werden nicht ausgewiesen; undatierte Ereignisse und fehlende Angaben wurden nach festen Regeln zugeordnet (siehe Kapitel 2.4).

[2)] Nicht infrage kommend: Personen, die vor Behandlungsbeginn keinen MM erreicht oder wiedererlangt hatten, der nicht verloren wurde, und nach Behandlungsbeginn keinen MM erreichten oder wiedererlangten.

[3)] Die Prozentanteile werden unter Verwendung der Anzahl N der behandelten Personen berechnet.

[4)] Infrage kommend: Personen, die vor Behandlungsbeginn mind. einen MM erreicht oder wiedererlangt und nicht verloren hatten oder nach Behandlungsbeginn einen MM erreichten oder wiedererlangten.

[5)] Die Prozentanteile werden unter Verwendung der Anzahl N_1 der behandelten Personen berechnet, die MM potentiell verlieren können.

[6)] Die spezifischen Anzahlen der verlorenen MM schließen sich gegenseitig aus.

[7)] Eine behandelte Person konnte in mehr als einem MM berücksichtigt werden.

[8)] Die Prozentsätze basieren auf der Anzahl N_x der behandelten Personen, die jeweils einen MM potentiell verlieren können, wobei maximal N_1 (= Anzahl der behandelten Personen, die den MM potentiell verlieren können) berücksichtigt wird.

Abkürzungen: dC: Desoxycytidin; DC: Doxecitin; dCMP: Desoxycytidinmonophosphat; dT: Desoxythymidin; DT: Doxribtimin; dTMP: Desoxythymidinmonophosphat; ISE: Integrierte Wirksamkeitsanalyse; MM: Motorischer Meilenstein.

Funktionsbereich Atmung – Künstliche Beatmung

Zu Behandlungsbeginn (letzter vor ISE-Tag 1 dokumentierter Status) waren 18 Personen (46,2 %) beatmet, zu einem beliebigen Zeitpunkt nach Behandlungsbeginn 23 (59,0 %). Von den 18 zu Behandlungsbeginn beatmeten Personen wurde die Beatmung bei 4 (22,2 %) beendet; 14 (77,8 %) waren bei Beobachtungsende weiterhin beatmet. Von den 14 zu Behandlungsbeginn nicht invasiv beatmeten Personen wurde 1 (7,1 %) auf invasive Beatmung umgestellt. Von 21 Personen ohne Beatmung begannen 5 (23,8 %) nach Behandlungsbeginn eine Beatmung, 2 davon invasiv; bei 1 Person wurde sie wieder eingestellt.

Die Teilgruppen sind klein und über den Ausgangsstatus definiert, Zeitpunkt und Dauer der Ereignisse sind nicht ausgewiesen und ob der Beatmungsstatus bei allen 39 Personen tatsächlich erhoben wurde, ist nicht sichergestellt (siehe Abschnitt „Vorbemerkung zur Einordnung der dargestellten Ergebnisse“ und Kapitel 2.3.2). Vor diesem Hintergrund lässt sich den Zahlen allenfalls Folgendes entnehmen: Beendigungen und Neubeginne halten sich etwa die Waage und die große Mehrheit der beatmeten Personen blieb beatmet. Beginn und Beendigung lagen im ärztlichen Ermessen, sodass die Ergebnisse neben dem Krankheitsverlauf auch die jeweils lokale Versorgungspraxis widerspiegeln dürften; Lücken in der Dokumentation wurden als durchgehende Beatmung gewertet (siehe Kapitel 2.3.2). Eine Richtung der Veränderung des Beatmungsbedarfs unter Behandlung lässt sich daraus nicht ableiten.

Tabelle 11: Ergebnisse zum Endpunkt „Künstliche Beatmung“; ISE-Kernpool, Datenschnitt: 14.03.2025

Studien 101 + 102 **Funktionsbereich Atmung – Künstliche Beatmung[1)]**	**ISE-Kernpool** **N = 39** *n (%)[2)]*
Behandelte mit künstlicher Beatmung **zu Behandlungsbeginn**	18 (46,2)
Art der künstlichen Beatmung zu Behandlungsbeginn	
Nicht-invasiv (z. B. BiPAP, CPAP)[5)]	14 (77,8)
Nach Behandlungsbeginn auf invasiv umgestellt[3)]	1 (7,1)
Invasiv[4)5)]	4 (22,2)
Nach Behandlungsbeginn auf nicht-invasiv umgestellt (z. B. BiPAP, CPAP)	0 (0)
Künstliche Beatmung nach Behandlungsbeginn eingestellt[5)]	4 (22,2)
Künstliche Beatmung bei Beobachtungsende nach Behandlungsbeginn[5)]	14 (77,8)
Behandelte mit künstlicher Beatmung zu beliebigem Zeitpunkt **nach Behandlungsbeginn**	23 (59,0)
Künstliche Beatmung nach Behandlungsbeginn eingestellt[6)]	5 (21,7)
Künstliche Beatmung nach Behandlungsbeginn gestartet	k. A.[7)]
Behandelte ohne künstliche Beatmung **zu Behandlungsbeginn**	21 (53,8)
Künstliche Beatmung nach Behandlungsbeginn begonnen[8)]	5 (23,8)
Art der künstlichen Beatmung nach Behandlungsbeginn	
Nicht-invasiv[9)] (z. B. BiPAP, CPAP)	4 (80,0)
Invasiv[4)9)]	2 (40,0)
Künstliche Beatmung nach Behandlungsbeginn eingestellt[9)]	1 (20,0)

[1)] Behandlungsbeginn = ISE-Tag 1, d. h. die erste dokumentierte Gabe einer berücksichtigten Formulierung (dC/dT, dCMP/dTMP oder DC/DT). Sie liegt bei allen Personen des ISE-Kernpools vor der Aufnahme in Studie 101 bzw. 102; der Behandlungsbeginn ist daher nicht mit dem Studieneintritt oder Studientag 1 der Studie 102 gleichzusetzen (siehe Kapitel 2.2.2). Der Beatmungsstatus zu Behandlungsbeginn entspricht der zuletzt vor der ersten Gabe dokumentierten Angabe; sie kann vom Tag des Behandlungsbeginns stammen, aber auch deutlich länger zurückliegen. In Studie 101 wurde bei fehlender Angabe vor Behandlungsbeginn die erste Angabe danach verwendet; bei den erst in Studie 102 aufgenommenen Personen kann der Status nur aus bei Studieneintritt – unter bereits laufender Behandlung – retrospektiv erhobenen Angaben stammen (siehe Kapitel 2.2.2). Nach Behandlungsbeginn umfasst die gesamte individuelle Beobachtung vom Behandlungsbeginn bis zur letzten dokumentierten Angabe vor dem Datenschnitt vom 14.03.2025 (individuell unterschiedlich lang; Behandlungsdauer zum Datenschnitt vom 15.03.2024 im Median 78,4 Monate (Spannweite: 53,2–157,1), Monate zu 28 Tagen;

zum Datenschnitt vom 14.03.2025 nicht berichtet). Ein Beginn, eine Einstellung oder eine Umstellung der Beatmung zählt, wenn sie zu irgendeinem Zeitpunkt in diesem Zeitraum dokumentiert ist. „Zu beliebigem Zeitpunkt nach Behandlungsbeginn“: Alle Personen, die in diesem Zeitraum beatmet waren – sowohl die bereits zu Behandlungsbeginn beatmeten als auch die danach neu beatmeten. „Bei Beobachtungsende nach Behandlungsbeginn“: Zuletzt dokumentierter Status am Ende der individuellen Beobachtung; ob die Beatmung dazwischen durchgehend bestand, ist nicht berichtet. Wann ein Ereignis eintrat und wie lange der danach bestehende Zustand anhielt, wird nicht ausgewiesen; Gründe für Änderungen der Beatmung wurden nicht erhoben; Dokumentationslücken wurden als durchgehende Beatmung gewertet (siehe Kapitel 2.4). Nicht invasiv: z. B. BiPAP, CPAP; invasiv: Beatmung ohne oder mit Tracheostoma; Hustenassistenz gilt nicht als künstliche Beatmung.

[2)] Da eine behandelte Person mehrmals in den einzelnen Kategorien erscheinen kann, können die Prozentanteile mehr als 100 % ergeben.

[3)] Prozentwerte beziehen sich auf die 14 Personen mit nicht invasiver Beatmung zu Behandlungsbeginn.

[4)] Invasive Beatmung ohne und mit Tracheostoma zusammengefasst. Welche Beatmungszugänge als invasive Beatmung ohne Tracheostoma gezählt wurden, ist in den Unterlagen nicht spezifiziert.

[5)] Prozentwerte beziehen sich auf die 18 Personen mit künstlicher Beatmung zu Behandlungsbeginn.

[6)] Prozentwerte beziehen sich auf die 23 Personen mit künstlicher Beatmung zu beliebigem Zeitpunkt nach Behandlungsbeginn.

[7)] In der ISE für diese Teilgruppe nicht ausgewiesen; rechnerisch mind. 5 Personen (23 Personen mit Beatmung zu beliebigem Zeitpunkt nach Behandlungsbeginn abzüglich 18 Personen mit Beatmung zu Behandlungsbeginn), entsprechend den 5 Personen ohne Beatmung zu Behandlungsbeginn, die danach eine Beatmung begannen. Nähere Informationen liegen nicht vor.

[8)] Prozentwerte beziehen sich auf die 21 Personen ohne künstliche Beatmung zu Behandlungsbeginn.

[9)] Prozentwerte beziehen sich auf die 5 Personen, die nach Behandlungsbeginn eine künstliche Beatmung begannen.

Abkürzungen: BiPAP: Bilevel Positive Airway Pressure; CPAP: Continuous Positive Airway Pressure; dC: Desoxycytidin; DC: Doxecitin; dCMP: Desoxycytidinmonophosphat; dT: Desoxythymidin; DT: Doxribtimin; dTMP: Desoxythymidinmonophosphat; ISE: Integrierte Wirksamkeitsanalyse; k. A.: keine Angabe; pU: pharmazeutischer Unternehmer.

Funktionsbereich Ernährung – Künstliche Ernährung

Von 11 Personen mit Sonde zu Behandlungsbeginn wurde sie bei 3 (27,3 %) entfernt; 8 (72,7 %) wurden bei Beobachtungsende weiterhin über eine Sonde ernährt. Von 28 Personen ohne Sonde zu Behandlungsbeginn erhielten 4 (14,3 %) nach Behandlungsbeginn eine Sonde, die bei 2 von ihnen wieder entfernt wurde. Zu einem beliebigen Zeitpunkt nach Behandlungsbeginn waren damit 15 Personen (38,5 %) sondenernährt, bei 5 von ihnen (33,3 %) wurde die Sonde entfernt.

Auch hier sind die Teilgruppen klein und über den Ausgangsstatus definiert, Einlage und Entfernung lagen in ärztlichem Ermessen, Gründe sind nicht für alle Ereignisse dokumentiert, Dokumentationslücken wurden als durchgehende Sondenernährung gewertet und ob der Sondenstatus bei allen 39 Personen tatsächlich erhoben wurde, ist nicht sichergestellt (siehe Abschnitt „Vorbemerkung zur Einordnung der dargestellten Ergebnisse“ und Kapitel 2.3.2). Vor diesem Hintergrund lässt sich den Zahlen allenfalls entnehmen, dass die Mehrheit der zu Behandlungsbeginn sondenernährten Personen sondenabhängig blieb und den drei Entfernungen bei zu Behandlungsbeginn sondenernährten Personen vier neue Einlagen gegenüberstehen (insgesamt fünf Entfernungen). Eine Richtung der Veränderung des Sondenbedarfs unter Behandlung lässt sich daraus nicht ableiten; die Angaben erlauben nur eine vorsichtige Beschreibung dokumentierter Sondenereignisse.

Tabelle 12: Ergebnisse zum Endpunkt „Künstliche Ernährung"; ISE-Kernpool, Datenschnitt: 14.03.2025

Studien 101 + 102 **Funktionsbereich Ernährung – Künstliche Ernährung**[1)]	**ISE-Kernpool** **N = 39** *n (%)*[2)]
Behandelte mit nasogastraler/PEG-Sonde **zu Behandlungsbeginn**	11 (28,2)
Sonde nach Behandlungsbeginn entfernt[3)]	3 (27,3)
Grund für die Entfernung der Sonde nach Behandlungsbeginn[4)5)]	
Nicht näher spezifizierte Verbesserung	0 (0)
Verbesserung bzgl. Dysphagie und oraler Nahrungsaufnahme	1 (33,3)
Verbesserung bzgl. Dysphagie	1 (33,3)
Verbesserung bzgl. oraler Nahrungsaufnahme	0 (0)
Andere	0 (0)
Sondenbezogenes UE	0 (0)
Künstliche Ernährung bei Beobachtungsende nach Behandlungsbeginn[3)]	8 (72,7)
Behandelte ohne nasogastrale/PEG-Sonde **zu Behandlungsbeginn**	28 (71,8)
Künstliche Ernährung nach Behandlungsbeginn begonnen[6)]	4 (14,3)
Grund für Beginn der künstlichen Ernährung[7)]	
Ergänzung zur oralen Nahrungsaufnahme	3 (75,0)
Dysphagie	1 (25,0)
Andere	1 (25,0)
Nicht berichtet	0 (0)
Sonde nach Behandlungsbeginn entfernt[7)]	2 (50,0)
Grund für die Entfernung der Sonde nach Behandlungsbeginn[8)]	
Nicht näher spezifizierte Verbesserung	1 (50,0)
Verbesserung bzgl. Dysphagie und oraler Nahrungsaufnahme	0 (0)
Verbesserung bzgl. Dysphagie	0 (0)
Verbesserung bzgl. oraler Nahrungsaufnahme	0 (0)
Andere	0 (0)
Sondenbezogenes UE	1 (50,0)
Behandelte mit nasogastraler/PEG-Sonde zu beliebigem Zeitpunkt **nach Behandlungsbeginn**	15 (38,5)
Sonde nach Behandlungsbeginn entfernt[9)]	5 (33,3)
Grund für die Entfernung der Sonde nach Behandlungsbeginn[4)10)]	
Nicht näher spezifizierte Verbesserung	1 (20,0)
Verbesserung bzgl. Dysphagie und oraler Nahrungsaufnahme	1 (20,0)
Verbesserung bzgl. Dysphagie	1 (20,0)
Verbesserung bzgl. oraler Nahrungsaufnahme	0 (0)
Andere	0 (0)
Sondenbezogenes UE	1 (20,0)

[1)] Behandlungsbeginn = ISE-Tag 1, d. h. die erste dokumentierte Gabe einer berücksichtigten Formulierung (dC/dT, dCMP/dTMP oder DC/DT). Sie liegt bei allen Personen des ISE-Kernpools vor der Aufnahme in Studie 101 bzw. 102; der Behandlungsbeginn ist daher nicht mit dem Studieneintritt oder Studientag 1 der Studie 102 gleichzusetzen (siehe Kapitel 2.2.2). Künstliche Ernährung: Ernährung über eine nasogastrale oder gastrostomische Sonde (in der Tabelle als nasogastrale/PEG-Sonde bezeichnet); „Sonde" und „künstliche Ernährung" werden gleichbedeutend verwendet. Der Sondenstatus zu Behandlungsbeginn entspricht der zuletzt vor der ersten Gabe dokumentierten Angabe; sie kann vom Tag des Behandlungsbeginns stammen, aber auch deutlich länger zurückliegen. In Studie 101 wurde bei fehlender Angabe vor Behandlungsbeginn die erste Angabe danach verwendet; bei den erst in Studie 102 aufgenommenen Personen kann der Status nur aus bei Studieneintritt – unter bereits laufender Behandlung – retrospektiv erhobenen Angaben stammen (siehe Kapitel 2.2.2). Nach Behandlungsbeginn umfasst die gesamte individuelle Beobachtung vom Behandlungsbeginn bis zur letzten dokumentierten Angabe vor dem Datenschnitt vom 14.03.2025 (individuell unterschiedlich lang; Behandlungsdauer zum Datenschnitt vom 15.03.2024 im Median 78,4 Monate (Spannweite: 53,2–157,1), Monate zu 28 Tagen; zum Datenschnitt vom 14.03.2025 nicht berichtet). Eine Einlage oder Entfernung der Sonde zählt, wenn sie zu irgendeinem Zeitpunkt in diesem Zeitraum dokumentiert ist. „Zu beliebigem Zeitpunkt nach Behandlungsbeginn": Alle Personen, die in diesem Zeitraum eine Sonde hatten – sowohl die bereits zu Behandlungsbeginn sondenernährten als auch die

mit Einlage danach. „Bei Beobachtungsende nach Behandlungsbeginn“: Zuletzt dokumentierter Status am Ende der individuellen Beobachtung; ob die Sondenernährung dazwischen durchgehend bestand, ist nicht berichtet. Gründe für Einlage und Entfernung sind Mehrfachnennungen und nicht für alle Ereignisse dokumentiert. Wann ein Ereignis eintrat und wie lange der danach bestehende Zustand anhielt, wird nicht ausgewiesen; Dokumentationslücken wurden als durchgehende Sondenernährung gewertet (siehe Kapitel 2.4).

[2)] Da eine behandelte Person mehrmals in den einzelnen Kategorien erscheinen kann, können die Prozentanteile mehr als 100 % ergeben.

[3)] Prozentanteile beziehen sich auf die 11 Personen mit Sonde zu Behandlungsbeginn.

[4)] Gründe liegen nicht für alle Entfernungen vor; Mehrfachnennungen sind möglich.

[5)] Prozentanteile beziehen sich auf die 3 Personen, bei denen die Sonde nach Behandlungsbeginn entfernt wurde.

[6)] Prozentanteile beziehen sich auf die 28 Personen ohne Sonde zu Behandlungsbeginn.

[7)] Prozentanteile beziehen sich auf die 4 Personen, die nach Behandlungsbeginn eine Sonde erhielten.

[8)] Prozentanteile beziehen sich auf die 2 Personen, bei denen eine nach Behandlungsbeginn eingelegte Sonde wieder entfernt wurde.

[9)] Prozentanteile beziehen sich auf die 15 Personen mit Sonde zu beliebigem Zeitpunkt nach Behandlungsbeginn.

[10)] Prozentanteile beziehen sich auf die 5 Personen, bei denen die Sonde nach Behandlungsbeginn entfernt wurde.

Abkürzungen: dC: Desoxycytidin; DC: Doxecitin; dCMP: Desoxycytidinmonophosphat; dT: Desoxythymidin; DT: Doxribtimin; dTMP: Desoxythymidinmonophosphat; ISE: Integrierte Wirksamkeitsanalyse; PEG: Perkutane endoskopische Gastrostomie; UE: Unerwünschtes Ereignis.

7.4 Sicherheit

Bei der Interpretation der Sicherheitsergebnisse ist zu berücksichtigen, dass der Auswertungsbeginn (und in Folge auch der Auswertungszeitraum) der Personen aus Studie 102, die zuvor nicht an Studie 101 teilgenommen haben (n = 10), nicht bekannt ist (siehe Kapitel 2.4). Für die anderen Personen (n = 29) begann die Sicherheitsnachbeobachtung mit der ersten Studienmedikationsgabe in Studie 101.

Alle 39 Personen hatten mindestens ein UE, 23 (59,0 %) mindestens ein SUE; bei keiner Person führte ein UE zum Abbruch. Die häufigsten UE waren „Diarrhö“ (84,6 %), „Fieber“ (46,2 %), „COVID-19“ (43,6 %), „Infektion der oberen Atemwege“ (41,0 %), „Rhinorrhoe“ (38,5 %) und „Erbrechen“ (33,3 %); „Alanin-Aminotransferase erhöht“ trat bei 28,2 %, „Aspartat-Aminotransferase erhöht“ bei 23,1 % auf. Die häufigsten SUE waren „Akute respiratorische Insuffizienz“ und „Pneumonie“ (jeweils 12,8 %) sowie „Oberschenkelfraktur“ (10,3 %).

Für schwere UE (CTCAE-Grad ≥ 3) liegen für den Kernpool keine Angaben vor. Ebenso fehlt für den Kernpool eine Auswertung ohne erkrankungsbezogene Ereignisse: Der pU hat in Modul 4 zwar erkrankungsbezogene Ereignisse unter Einbeziehung seiner medizinischen Fachabteilung spezifiziert, legt eine entsprechende Auswertung jedoch nur für die anwendungsgebietsbezogene ISS-Population (N = 50) vor (siehe Kapitel 2.3.4). Die dargestellten Gesamtraten der UE und SUE enthalten damit auch Ereignisse, die auf die Grunderkrankung zurückgehen.

„Diarrhö“ betraf die meisten Personen und erscheint als Nebenwirkung plausibel; Transaminasenerhöhungen traten bei etwa einem Viertel auf. Die SUE betreffen überwiegend Atmung, Schlucken und Ernährung und damit möglicherweise Manifestationen der Grunderkrankung bzw. supportiver Maßnahmen zur Beatmung, die nicht als erkrankungsbezogene Ereignisse herausgerechnet wurden; auch die Oberschenkelfrakturen könnten mit der eingeschränkten Mobilität zusammenhängen. Auswertungen nach SOC liegen nicht vor.

Dass kein UE zum Abbruch führte, ist vor dem Hintergrund der Selektion auf Verträglichkeit in Studie 102 zu sehen (Ausschluss bei Lebererkrankung, kein Übergang nach Behandlungsabbruch; siehe Kapitel 2.2.3). Die 2 Personen, die die Behandlung in Studie 101 wegen erhöhter Leberwerte abgebrochen hatten, sind im Kernpool gesamt (N = 50) enthalten und entfallen im Kernpool (N = 39) nur wegen ihres Symptombeginns nach dem 12. Lebensjahr; Abbrüche vor Studieneinschluss sind dagegen nicht erfasst, sodass hepatische Nebenwirkungen unterrepräsentiert

sein könnten. Da zudem ein großer Teil der behandlungsbegleitenden Zeit nur retrospektiv dokumentiert ist (siehe „Vorbemerkung zur Einordnung der dargestellten Ergebnisse“), könnten UE vor Eintritt in Studie 102 unvollständig erfasst sein; ohne expositionsadjustierte Raten lassen sich die UE-Anteile nicht ins Verhältnis zur Beobachtungszeit setzen.

Tabelle 13: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE; ISS-Kernpool, Datenschnitt: 15.03.2024

Studien 101 + 102 **Zusammenfassung der UE[1)]** *Personen mit mindestens einem …*	**ISS-Kernpool** **N = 39** *n (%)*
UE (ergänzend dargestellt)	39 (100
schweren UE (CTCAE-Grad ≥ 3)	k. A.
SUE	23 (59,0)
UE, das zum Abbruch der Studienmedikation führte	0 (0)

[1)] Die Kodierung erfolgte nach MedDRA-Version 26.0.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; ISS: Integrierte Sicherheitsauswertung; k. A.: keine Angabe; MedDRA: Medical Dictionary for Regulatory Activities; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Unerwünschte Ereignisse

Eine Kodierung nach SOC ist für die UE nicht verfügbar.

Tabelle 14: UE mit Inzidenz ≥ 10 % nach Preferred Term; ISS-Kernpool, Datenschnitt: 15.03.2024

Studien 101 + 102 **UE** *Preferred Term[1)]*	**ISS-Kernpool** **N = 39** *n (%)*
Diarrhö	33 (84,6)
Fieber	18 (46,2)
COVID-19	17 (43,6)
Infektion der oberen Atemwege	16 (41,0)
Rhinorrhoe	15 (38,5)
Erbrechen	13 (33,3)
Husten	11 (28,2)
Kopfschmerzen	11 (28,2)
Alanin-Aminotransferase erhöht	11 (28,2)
Abdominalschmerz	10 (25,6)
Gastroenteritis	9 (23,1)
Aspartat-Aminotransferase erhöht	9 (23,1)
Atemwegsinfektion	8 (20,5)
Erhöhte Kreatinphosphokinase im Blut	8 (20,5)
Ohreninfektion	7 (17,9)
Pneumonie	7 (17,9)
Grippe	7 (17,9)
Schmerzen im Oropharynx	7 (17,9)

Studien 101 + 102 **UE** *Preferred Term*[1)]	**ISS-Kernpool** **N = 39** *n (%)*
Ausschlag	7 (17,9)
Anionenlücke erhöht	7 (17,9)
Milchsäure im Blut erhöht	7 (17,9)
Oberschenkelfraktur	5 (12,8)
Akute respiratorische Insuffizienz	5 (12,8)
Grippeähnliche Erkrankung	5 (12,8)
Dyspnoe	5 (12,8)
Basophilenzahl erhöht	5 (12,8)
Dysphagie	5 (12,8)
Thrombozytenzahl erhöht	5 (12,8)
Schmerz in einer Extremität	4 (10,3)
Harnwegsinfektion	4 (10,3)
Schmerzen Oberbauch	4 (10,3)
Depression	4 (10,3)
Muskuläre Schwäche	4 (10,3)
Kohlendioxid erniedrigt	4 (10,3)
Gastroösophageale Refluxerkrankung	4 (10,3)
Nasopharyngitis	4 (10,3)
Skoliose	4 (10,3)
Tachykardie	4 (10,3)

[1)] Die Kodierung erfolgte nach MedDRA-Version 26.0 und liegt nur nach Preferred Term vor. Eine Kodierung nach Systemorganklasse ist nicht verfügbar.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; ISS: Integrierte Sicherheitsauswertung; UE: Unerwünschtes Ereignis.

Schwerwiegende unerwünschte Ereignisse

Eine Kodierung nach SOC ist für die SUE nicht verfügbar.

Tabelle 15: SUE, die bei ≥ 5 % der Personen aufgetreten sind, nach Preferred Term; ISS-Kernpool, Datenschnitt: 15.03.2024

Studien 101 + 102 **SUE** *Preferred Term*[1)]	**ISS-Kernpool** **N = 39** *n (%)*
Akute respiratorische Insuffizienz	5 (12,8)
Pneumonie	5 (12,8)
Oberschenkelfraktur	4 (10,3)
Nephrolithiasis	2 (5,1)
Pneumonie grippal	2 (5,1)
Fieber	2 (5,1)
Bronchitis	2 (5,1)

Studien 101 + 102 **SUE** *Preferred Term*[1)]	**ISS-Kernpool** **N = 39** *n (%)*
COVID-19	2 (5,1)
Brustkorbschmerz	2 (5,1)
Chronische respiratorische Insuffizienz	2 (5,1)
Dysphagie	2 (5,1)
Dyspnoe	2 (5,1)
Schlechtes Gedeihen	2 (5,1)
Gastrostomie	2 (5,1)
Aspirationspneumonie	2 (5,1)
Respiratorische Insuffizienz	2 (5,1)
Virale Infektion der oberen Atemwege	2 (5,1)

[1)] Die Kodierung erfolgte nach MedDRA-Version 26.0 und liegt nur nach Preferred Term vor. Eine Kodierung nach Systemorganklasse ist nicht verfügbar.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; ISS: Integrierte Sicherheitsauswertung; SUE: Schwerwiegendes unerwünschtes Ereignis.

7.5 Indirekter Vergleich zum Gesamtüberleben

In den Studienunterlagen wurden nicht randomisierte Vergleiche ohne Brückenkomparator auf Basis von individuellen Patientendaten, die in Sekundärquellen publiziert worden sind, durchgeführt.

In Modul 4 wurde ein „Target Trial“ zum Endpunkt „Mortalität“ durchgeführt und Ereigniszeitanalysen (Zeit bis zum Tod seit Symptombeginn, Zeit bis zum Tod seit Behandlungsbeginn) eingereicht.

Als Informationsquellen für die Durchführung der TTE wurde in den Unterlagen des pU insbesondere die Ausführungen in Modul 4 identifiziert. Ergänzend wurde ein Dokument mit Spezifikationen für die Analysen in Modul 4 identifiziert, jedoch kein SAP oder Studienprotokoll für die TTE. Es wurde in der Gesamtschau auf Basis der vorliegenden Unterlagen kein präspezifiziertes Vorgehen für die TTE festgestellt.

Als bewertungsgegenständlich wird der indirekte Vergleich gegenüber der Zulassungspopulation (Personen mit einem Alter ≤ 12 Jahren bei Symptombeginn) angesehen.

Datenquellen für die externe Kontrolle

Zur TTE wurde eine externe Vergleichskohorte herangezogen, die aus unbehandelten Personen mit TK2d besteht und auf Basis von publizierten individuellen Patientendaten aus zwei separaten Datenquellen (Updated-UPD und Studie 107) gebildet wurde. Bei der Erstellung der Datenquellen wurde nach Angaben des pU sichergestellt, dass sich Personen nicht doppelt in den Datenquellen befinden. Im Rahmen der Erstellung der Datenquelle wurden Autorinnen und Autoren der Primärpublikationen kontaktiert, um Unklarheiten beseitigen zu können.

Zur Erstellung der Datenquelle „Updated-UPD“ wurde eine systematische Literaturrecherche (SLR) nach publizierten Fallserien und Fallberichten durchgeführt. Die SLR wurde am 17.06.2019 aktualisiert und erneut durchgeführt sowie um neue identifizierte Daten ergänzt. Das Ziel war die Erstellung einer externen Kontrollgruppe für die retrospektive Kohortenstudie Studie 101 (siehe auch Tabelle 2). Als Einschlusskriterium wurde definiert, dass individuelle Patientendaten

vorliegen müssen sowie eine genetische Bestätigung von Mutationen im TK2-Gen. Es wurden Variablen definiert, die erhoben werden sollten, sofern diese vorhanden waren. Diese betrafen das Alter zu verschiedenen Zeitpunkten (Symptombeginn, Tod, Verlust von Mobilitätsfunktionen, Beatmung), Geschlecht und TK2-Mutation. Weiterhin sollten wenn möglich biologische Gewebeproben entnommen werden.

In der ersten durchgeführten SLR wurden drei retrospektive Kohortenstudien identifiziert, die individuelle Patientendaten in den jeweiligen Supplements bereitstellen sowie weitere publizierte Einzelfallberichte. In der ersten Literaturrecherche wurden insgesamt 103 Patientendaten identifiziert („101-UPD"). In der aktualisierten Literaturrecherche konnten 25 weitere Patientendaten identifiziert werden, sodass durch diese Datenquelle Informationen zu insgesamt 128 Personen mit TK2d vorliegen („Updated-UPD"). 3 Personen, die identifiziert worden sind als „unbehandelt", haben im Rahmen einer parallelen klinischen Studie, die nicht im Zusammenhang mit der Nutzenbewertung steht, eine Intervention erhalten. Diese wurden zum Zeitpunkt des Erhalts der Intervention für die Ereigniszeitanalyse zensiert. Die identifizierten Fallberichte und retrospektiven Studien wurden im Zeitraum zwischen 2001 und 2019 publiziert. Weitere Informationen zur Aktualität der Fallberichte liegen auf Basis der vorliegenden Unterlagen nicht vor.

Eine weitere Datenquelle ist eine retrospektive Datensammlung von Fallberichten im Rahmen der retrospektiven Studie 107 (siehe auch Tabelle 2). Mit Protokollamendment 2 wurde die Chart-Review-Studie um die Erhebung von unbehandelten Personen mittels retrospektiver Datenerhebung erweitert. Als Einschlusskriterien wurde eine Einverständniserklärung, eine bestätigte genetische Mutation im TK2-Gen sowie das Vorhandensein von medizinischen Aufzeichnungen mit einem Vitalstatus (verstorben oder lebend) beschrieben.

Insgesamt wurden 171 individuelle Patientendaten identifiziert. Nach Bereinigung von Überschneidungen und der Zulassungspopulation von Personen mit einem Alter von ≤ 12 Jahren bei Symptombeginn wurden 116 Personen in die Datenquelle aufgenommen.

Bewertung der Datenquellen

Die systematische Suche nach möglichen Datenquellen im Anwendungsgebiet wird als sachgerecht bewertet.

Eine retrospektive Datenerhebung ist mit inhärenten Limitationen verbunden. In der vorliegenden Datenquelle liegt eine unklare Aktualität der Daten vor. Es konnten keine konkreten Angaben identifiziert werden, aus welchem Zeitraum die individuellen Patientendaten identifiziert worden sind. Ob eine retrospektive (bspw. post mortem) Zuordnung zu der Grunderkrankung bei Personen im Kontrollarm erfolgte, kann auf Basis der vorliegenden Unterlagen nicht ausgeschlossen werden. Bei der vorliegenden Datensituation muss berücksichtigt werden, dass eine erstmalige umfassende Charakterisierung der Erkrankung im Jahr 2001 erfolgte [7]. Es liegt eine sehr hohe Anzahl an fehlenden Daten im vorliegenden Datensatz vor und es ist unklar, inwiefern krankheitsdynamische Charakteristika valide erhoben werden konnten. Für die externe Kontrolle liegt laut pU keine „wahre" Baseline vor. Es konnten nicht für alle erhobenen Daten deskriptive Auswertungen identifiziert werden, sodass krankheitsspezifische Informationen nicht abschließend bewertet werden können (bspw. Genotyp der TK2-Mutation). Dies gilt sowohl für die Interventions- als auch für die Vergleichspopulation. Dies erschwert die Interpretation der vorliegenden Datensituation.

Die Erhebung des Alters bei Symptombeginn basiert auf retrospektiven Angaben des Prüfpersonals ohne konkrete Operationalisierung (bspw. qualifizierendes Symptom). Eine weitere Limitation ist eine unklare bzw. nicht näher beschriebene Definition und Operationalisierung von „Unbehandelten".

Auswahl der Interventionspopulation (Doxecitin und Doxribtimin)

Die ISE für den Endpunkt „Gesamtüberleben“ wird vom pU als Vergleichspopulation herangezogen. Zur Einschätzung der Interventionspopulation zur Mortalität siehe auch Kapitel 2.3.1. Grundsätzliches zur Zusammenfassung der Studiendaten findet sich in Kapitel 2.2.2.

Die Ein- und Ausschlusskriterien der retrospektiven Studie 101 für die Interventionspopulation beinhalten, dass eine Visite zwischen dem 01.06.2018 und dem 15.12.2018 stattgefunden haben muss. Weitere Ein- und Ausschlusskriterien beinhalten die genetische Testung auf TK2-Mutation sowie der Ausschluss einer polygenetischen oder anderen genetischen Erkrankung. Durch das Zusammenführen von mehreren Datenquellen mussten bei der Erstellung der Datenquellen potentielle Doppelterfassungen von individuellen Patientendaten berücksichtigt werden. Weiterhin wurde die Zulassungspopulation auf Personen mit einem Alter bei Symptombeginn von unter 12 Jahren berücksichtigt. Insgesamt wurden nach Bereinigung von Überschneidungen und Berücksichtigung der Zulassungspopulation N = 82 identifiziert. Die letzte Beobachtung einer Person erfolgte am 28.03.2019 und der Studienbericht ist vom 31.03.2023.

Für den indirekten Vergleich wurde eine „Evaluable Population“ (EP) definiert, die aus dem Full Analysis Set (N = 82) hergeleitet wird. Dazu werden Personen aus der Analyse ausgeschlossen, bei denen das Alter bei Symptombeginn nicht oder kein letzter bekannter Vitalstatus (Tod oder Lebensalter nach Symptombeginn) verfügbar war.

Auswahl der Vergleichspopulation (historische Verlaufskohorte)

In die Grundpopulation gingen wie oben beschrieben 116 Personen ein (Full Analysis Set).

Es wurden Ein- und Ausschlusskriterien für die Vergleichspopulation spezifiziert. Ausgeschlossen wurden Personen, deren Alter bei Symptombeginn (n = 9) oder das letzte Lebensalter (n = 17) zum Zeitpunkt der Datenerhebung unbekannt war. Um einen potentiellen Immortal Time Bias zu vermeiden und in der Population zu berücksichtigen, wurden unbehandelte Personen, die vor dem Alter gestorben sind, in dem die jüngste behandelte Person die Behandlung begonnen hat, ausgeschlossen (n = 13). In einem weiteren Schritt wurden Personen ausgeschlossen, für die keine vollständigen Datensätze vorlagen. Die Definition von „vollständigen Datensätzen“ bezieht sich damit auf die als relevant bezeichneten Confounder und verwendeten Baseline-Charakteristika, die als „Surrogatparameter“ für Confounder herangezogen worden sind (siehe dazu Abschnitt „Bewertung der verwendeten Confounder“), sodass für die adjustierte „Inverse Probability of Treatment Weighting“ (IPTW)-Analyse vom pU insgesamt 56 Personen berücksichtigt/herangezogenen worden sind.

Der Beobachtungszeitraum beginnt zum Zeitpunkt des Alters bei Symptombeginn. Es finden sich keine Informationen über begleitende Therapien oder generell zur Therapieentscheidung (bspw. palliativ).

Erhebung von Endpunkten

Die Erhebung des Gesamtüberlebens für die Interventionspopulation wird in Kapitel 2.3.1 beschrieben. Für die Vergleichspopulation wurde das Gesamtüberleben mittels der individuellen Patientendaten der Sekundärquellen hergeleitet. Es wurden weitere Endpunkte zur Morbidität erhoben und Vergleiche (Funktionsbereiche: Motorik, Atmung, Ernährung) durchgeführt. Es liegt ein sehr hoher Anteil an fehlenden Daten vor. Da die ISE-Population in der Nutzenbewertung nicht berücksichtigt wird, erfolgt keine eingehende Bewertung von den Vergleichen zu diesen Endpunkten.

Einordnung und Bewertung der Populationen

Der pU zieht als externe Vergleichspopulation identifizierte Sekundärquellen heran, die individuelle Patientendaten in den Anhängen bereitstellen. In der Folge kann ein Publikationsbias nicht ausgeschlossen werden, da somit insbesondere lediglich publizierte Fälle und Fallserien eingeschlossen werden konnten.

Der Selektionsmechanismus aus den Grundpopulationen wird beschrieben, jedoch kann dieser nicht vollständig nachvollzogen werden. Es wird davon ausgegangen, dass fehlende Daten einen Einfluss auf die Selektionsmechanismen haben. Die Verfügbarkeit von Daten war insbesondere von Relevanz für die Auswahl von Confoundern (nähere Beschreibung siehe unten) und hatte somit direkten Einfluss auf die Population, die zur adjustierten Analyse herangezogen worden ist („EP-IPTW“).

Die retrospektive Beobachtungsstudie zur Intervention mit Doxecitin und Doxribtimin Studie 101 hatte als Einschlusskriterium das Vorliegen einer Visite zwischen dem 01.06.2018 und dem 15.12.2018 vor dem Ende der Datenerhebung. Dadurch, dass eine Wirksamkeits- oder Sicherheitserhebung mit der Intervention vorliegen musste, konnten Personen, die die Intervention erhalten haben, kein Ereignis im Gesamtüberleben bis zum Zeitpunkt der Datenerhebung erleiden. Ein Immortal Time Bias kann für „behandelte“ Personen aus der retrospektiven Kohortenstudie nicht ausgeschlossen werden. Dies betrifft ungefähr die Hälfte der gesamten Interventionspopulation. Eine vergleichbare Limitation konnte für die weiteren Datenquellen der Interventionspopulation nicht identifiziert werden.

Die Erkrankung der TK2d zählt zu den primären mitochondrialen Myopathien und wurde als eigenständige Erkrankung erstmalig im Jahr 2001 beschrieben. Bei 41,1 % der Personen im externen Kontrollarm und bei 13 % im Interventionsarm lag das Geburtsjahr vor dem Jahr 2000. Die Zeit bis zur Diagnose seit Symptombeginn betrug im Median 11,40 Monate im Interventionsarm und 38,10 Monate im externen Kontrollarm und die Zeit bis zur Diagnose seit Geburt lag im Interventionsarm bei 37,65 Monaten und im Vergleichsarm bei 62,03 Monaten.

Der Zeitpunkt des kalendarischen Auftretens der Erkrankung kann inhärente Auswirkungen auf die Positivität der zu vergleichenden Populationen haben. Es ist zum Zeitpunkt der Nutzenbewertung unklar, ob Personen im externen Kontrollarm retrospektiv der TK2-Erkrankung zugeordnet worden sind, während die Zuordnung im Interventionsarm während des Krankheitsverlaufs erfolgte. Dies steht im Kontext der genetischen Testung auf TK2d. Der Zeitpunkt der genetischen Diagnose im Krankheitsverlauf in der Vergleichspopulation ist zum Zeitpunkt der Nutzenbewertung unklar.

Die vorliegende Evidenzbasis führt dazu, dass ein synthetischer Beobachtungsstart definiert werden muss, da die Ein- und Ausschlusskriterien sowie die teilweise konsekutiv durchgeführten retrospektiven Studien zur Intervention keinen vergleichbaren Beobachtungsstart definieren können. Insgesamt ist die Wahl des Alters bei Symptombeginn nachvollziehbar, jedoch die konkrete Operationalisierung durch das Fehlen einer konkreten Definition (bspw. Ptosis oder progressive Myopathie) mit Limitationen verbunden (z. B. unklare Symptomatik bei Erstsymptomen), die dazu führen können, dass kein vergleichbarer Beobachtungsstart vorliegt. Durch die spätere Diagnose seit Symptombeginn, die Zeit von Geburt bis zur Diagnose sowie des Geburtsjahrs kann nicht ausgeschlossen werden, dass Personen in der externen Kontrolle zu deutlich fortgeschrittenen Krankheitsstadien oder post mortem diagnostiziert worden sind im Vergleich zu der Interventionspopulation. Die Baseline-Charakteristika hinsichtlich der Krankheitsschwere und Krankheitsdynamik sind innerhalb der Interventionsstudien heterogen und für die externe Kontrollgruppe zu einem frühestmöglichen Zeitpunkt erhoben worden (wobei hier unklar ist worauf sich frühestmöglich bezieht).

Identifikation, Auswahl und Erhebung von Confoundern

Literaturrecherche und Auswahl

Es wurde eine SLR nach Confoundern im Anwendungsgebiet mit Fokus auf die Mortalität durchgeführt. In die SLR wurden insbesondere retrospektive Beobachtungsstudien ab dem Jahr 2000 eingeschlossen, welche zum Teil auch als externe Vergleichspopulation herangezogen worden sind. Es wurden 10 Studien identifiziert, die Prognosefaktoren zum Gesamtüberleben enthielten. Darauf aufbauend wurden Interviews mit Experten durchgeführt, die die identifizierten Confounder nach ihrer Relevanz bewerteten. Zudem konnten im Rahmen der Interviews weitere Confounder genannt werden.

Eine Übersicht der identifizierten Confounder sowie der Relevanzbewertung ist in Tabelle 16 dargestellt.

Tabelle 16: Durch den pU identifizierte und bewertete Confounder

Confounder	Operationalisierung	In SLR Identifiziert	Ranking	PS-Modell
Alter bei Symptombeginn	kontinuierlich	ja	1	ja
Krankheitsschwere	-	ja	2	nein
Genotyp	-	ja	-[1)]	nein
FVC	-	ja	3	nein
BMI	-	ja	4	nein
6MWT	-	ja	5	nein
NSAA	-	ja	6	nein
mtDNA-Level	-	ja	7	nein
Beeinträchtigung der respiratorischen Funktion	-	unklar	7	nein
Infantiler Phänotyp	-	unklar	9	nein
Intermediate Phänotyp	-	unklar	9	nein
Kardiomyopathie (Phänotyp)	-	unklar	9	nein
Enzephalopathie (Phänotyp)	-	unklar	9	nein
Cardiac Involvement	-	unklar	9	nein
Serum-Kreatinkinase	-	unklar	9	nein
Palliatives Therapiesetting	-	unklar	9	nein
Respiratorische Insuffizienz	-	unklar	9	nein
Genetische Testung	-	unklar	9	nein
Missense-Mutation (Genotyp)	-	unklar	18	nein
Kreatin-Level	-	ja	18	nein
Involvierte medizinische Fachdisziplinen[2)]	-	unklar	18	nein
Epilepsie	-		18	nein
Nonsense-Mutation	-	unklar	22	nein
Schwerwiegende neurologische Mutationen	-	unklar	22	nein
Ort der medizinischen Versorgung	-	unklar	22	nein
Symptomatische Therapiemaßnahmen	-	unklar	22	nein

Confounder	Operationali-sierung	In SLR Identifiziert	Ranking	PS-Modell
Geographische Herkunft		unklar	22	ja[3)]
Zeitpunkt der Diagnose	-	unklar	22	nein
Dysphagie	-	unklar	22	nein
Progressive Muskelschwäche	-	unklar	22	nein
Sekundäre Erkrankungen	-	unklar	22	nein
Qualitative Unterschiede in symptomatischer Behandlung	-	unklar	22	nein
Therapie beim Fortschreiten der Erkrankung	-	unklar	22	nein
Spezifische Begleitmedikation	-	unklar	33	nein
Orthopädische Erkrankung	-	unklar	33	nein
Spanische Herkunft (Founder Variante)	-	unklar	35	nein
GDF-15-Level	-	ja	35	nein
Verlust der Funktionsmutation	-	unklar	37	nein

[1)] Unterteilung in mehrere Confounder im Rahmen des Rankings (Missense-Mutation; Nonsense-Mutation; Verlust der Funktionsmutation; Schwerwiegende neurologische Mutationen).
[2)] Zusätzlich durch Experten genannt.
[3)] Geographische Region des Wohnorts wurde operationalisiert und nicht die geographische Herkunft.

Abkürzungen: 6MWT: 6-Minute Walk Test; BMI: Body-Mass-Index; FVC: Forcierte Vitalkapazität; GDF-15: Growth Differentiation Factor 15; mtDNA: mitochondriale DNA; NSAA: North Star Ambulatory Assessment; PS: Propensity Score; pU: pharmazeutischer Unternehmer; SLR: Systematische Literaturrecherche.

Eine Voraussetzung zur Adjustierung für Confounder im Rahmen eines nicht randomisierten Vergleichs ist die Verfügbarkeit der entsprechenden Daten im Datensatz. Dementsprechend wurden Confounder herangezogen, bei denen mindestens 70 % der Daten vorliegen. Die Datenverfügbarkeit ist in Tabelle 17 dargestellt.

Tabelle 17: Datenverfügbarkeit von als relevant bewerteten Confoundern

Confounder	Verfügbarkeit	Behandelte N = 82 *n (%)*	Unbehandelte N = 92 *n (%)*
Alter bei Symptombeginn	fehlend	0	0
	verfügbar	82 (100)	92 (100)
Krankheitsschwere	fehlend	Operationalisierung nicht möglich	
	verfügbar		
Genotyp	fehlend	k. A.	
	verfügbar		
FVC	fehlend	61 (74,4)	92 (100)
	verfügbar	21 (25,6)	0 (0)
Art der künstlichen Beatmung	fehlend	52 (63,4)	79 (85,9)
	verfügbar	30 (36,6)	13 (14,1)
BMI	fehlend	44 (53,7)	92 (100)
	verfügbar	38 (46,3)	0 (0)

Confounder	Verfügbarkeit	Behandelte N = 82 *n (%)*	Unbehandelte N = 92 *n (%)*
6MWT	fehlend	57 (69,5)	92 (100)
	verfügbar	25 (30,5)	0 (0)
NSAA	fehlend	72 (87,8)	92 (100)
	verfügbar	10 (12,2)	0 (0)
mtDNA-Level	fehlend	k. A.	
	verfügbar		
100MRT	fehlend	65 (79,3)	92 (100)
	verfügbar	17 (20,7)	0 (0)
Geschlecht	fehlend	0 (0)	0 (0)
	verfügbar	82 (100)	92 (100)
Geographische Region des Wohnorts	fehlend	0 (0)	25 (27,2)
	verfügbar	82 (100)	67 (72,8)
Geburtsjahr	fehlend	0 (0)	33 (35,9)
	verfügbar	82 (100)	59 (64,1)
Indexpatient	fehlend	5 (6,1)	26 (28,3)
	verfügbar	77 (93,9)	66 (71,7)

Abkürzungen: 100MRT: 100-Metre Run Test; 6MWT: 6-Minute Walk Test; BMI: Body-Mass-Index; k. A.: keine Angabe; FVC: Forcierte Vitalkapazität; mtDNA: mitochondriale DNA; NSAA: North Star Ambulatory Assessment.

Folgender Confounder wurde für den indirekten Vergleich für das Propensity-Score (PS)-Modell herangezogen:

- Alter bei Symptombeginn

Folgende Baseline-Charakteristika wurden zusätzlich in das PS-Modell einbezogen:

- Geschlecht
- Geographische Region des Wohnorts
- Geburtsjahr
- Indexpatient

„Geographische Region des Wohnorts“, „Geburtsjahr“ sowie „Indexpatient“ sollen als Surrogatparameter versorgungsbezogene Confounder abbilden. Als „Indexpatient“ ist eine Person definiert, die als erste Person in der unmittelbaren familiären Vorgeschichte von der Erkrankung betroffen ist.

Bewertung der verwendeten Confounder

Die Durchführung einer SLR und die anschließende Einordnung und Identifizierung zusätzlicher Confounder durch Experten wird als sachgerecht bewertet.

Es wurden nicht die Confounder herangezogen, die für das PS-Modell von Relevanz sind, sondern es gehen Datenerhebungen (Confounder, Baseline-Charakteristika) mit einer vorliegenden Datenverfügbarkeit von ≥ 70 % in das PS-Modell ein.

Die Auswahl der Confounder wird als nicht sachgerecht bewertet.

Statistische Methoden für den indirekten Vergleich

Der pU reicht eine Sensitivitätsanalyse zur primären Analyse ein, da die primäre Analyse laut Angaben in Modul 4 keine ausreichende Überlappung zeigte. Die Bewertung bezieht sich somit auf die in Modul 4 aufbereitete Sensitivitätsanalyse.

Eine statistische Adjustierung kann die genannten Limitationen (Immortal Time Bias für die Interventionspopulation; Confounder-Auswahl; nicht dargelegte Positivität; hoher Anteil fehlender Daten) zur Vergleichbarkeit der Populationen nicht beheben und die Ableitung eines Behandlungseffekts erscheint nicht sachgerecht. Im Folgenden wird die statistische Methodik hinter dem indirekten Vergleich dennoch beschrieben und anschließend eingeordnet.

Schätzung des PS-Modells

Die PS wurden mittels logistischer Regression unter Einbezug der gewählten Confounder berechnet. Zur PS-Adjustierung zieht der pU ein IPTW unter Verwendung von stabilisierten Gewichten heran. In der Sensitivitätsanalyse legt der pU „Overlap Weighting" vor, weil die primäre Analyse (zum „Average Treatment Effect" (ATE)) nicht für ausreichend Überlappung geführt hat. Durch das „Overlap Weighting" wird der „Average Treatment Effect in the Overlap" (ATO) geschätzt.

Die Schätzung des ATO ist definiert als die Schätzung des Behandlungseffekts für die Population, bei der eine Überlappung vorliegt. Dies entspricht der Schätzung eines Behandlungseffekts für eine Population, bei der eine klinische Equipoise (Gleichgewicht) auf Basis der verwendeten Confounder vorliegt. Dies kann interpretiert werden als eine Subgruppe („Subset") der Gesamtpopulation, für welche die Daten auf eine exakte Balance und optimale Varianz hindeuten [8].

Die Verwendung des IPTW-ATO führt dazu, dass Personen höher gewichtet werden, wenn der PS um 0,5 liegt und zu einer Runtergewichtung, sofern der PS sich näher an 0 oder näher an 1 orientiert. Die Gewichte sind per Definition zwischen 0 und 1 [6].

In der Gesamtschau kann der ATO unter Verwendung von IPTW als der Vergleich von zwei Populationen mit der optimalen internen Validität interpretiert werden.

Zur Modelldiagnostik zieht der pU graphische Darstellungen zur Überlappung, Balanciertheit und der Darstellung der Gewichte heran.

Bewertung der Confounder-Adjustierung und der statistischen Methoden

Es kann auf Basis der vorliegenden Informationen nicht davon ausgegangen werden, dass es sich beim vorliegenden PS-Modell um ein Modell handelt, das alle oder hinreichend viele Confounder mit Einfluss auf das Outcome oder die Therapiezuteilung enthält („Strong Ignorable Treatment Assignment (SITA)-Bedingung nicht erfüllt).

Bei Verwendung des ATO (mit nicht geeignetem PS-Modell) handelt es sich um keinen verallgemeinerbaren Behandlungsschätzer, sondern um einen Behandlungsschätzer mit der optimalen internen Validität. Das geschätzte mediane Gesamtüberleben in der externen Vergleichspopulation beträgt im Median 3 Jahre und würde somit im Median unter der Zeit von Geburt bis zur Diagnose liegen (62,03 Monate). Dies hängt damit zusammen, dass die Baseline-Charakteristika auf Basis der vorliegenden Informationen nicht die gewichtete EP-IPTW-ATO-Population beschreiben, sondern die naive EP-IPTW-Population. Der Zeitpunkt der genetischen Diagnose im Krankheitsverlauf für unbehandelte Personen ist weiterhin unklar. Dies hätte weitere Implikationen auf die Bewertung der Positivität der Populationen.

Das Alter bei erstmaliger Behandlung in der Interventionspopulation beträgt im Median 4,26 Jahre (Spannweite: 0,7–35,5) (EP-IPTW) und das mediane Gesamtüberleben für die externe Vergleichspopulation beträgt durch die IPTW-ATO-Adjustierung 3 Jahre (95%-KI: [1,1; 23,3]).

Durch die verwendete Gewichtungsmethodik deutet dies auf einen Immortal Time Bias zuungunsten der externen Vergleichspopulation hin.

Der pU reicht eine graphische Darstellung der PS-Verteilungen nach dem „Overlap Weighting“ mittels IPTW ein. Die Interpretation gestaltet sich auf Basis der vorliegenden Angaben schwierig und erscheint unter Berücksichtigung der antizipierten Orientierung um einen PS von 0,5 als wenig plausibel. Eine Darstellung der unadjustierten PS konnte nicht identifiziert werden. Die Balanciertheit wurde mittels standardisierter Mittelwertdifferenz (SMD) dargelegt. Unter der Annahme der exakten Balance (SMD = 0) liegen dennoch SMD ≥ 0 vor, die jedoch größtenteils unter 0,1 liegen.

Das Alter bei Symptombeginn wird in der Literatur und durch die Einordnung der befragten Experten als der wichtigste Confounder bewertet. Allerdings ist die Operationalisierung des Confounders auf Basis der vorliegenden Unterlagen nicht vollständig nachvollziehbar. Die Bestimmung des Zeitpunkts erfolgte auf Basis der vorliegenden Unterlagen retrospektiv, wobei die Studienunterlagen keine Determinanten zur Herleitung (bspw. initiale klinische Präsentation und aufgetretenes qualifizierendes Symptom) darlegen. Es muss berücksichtigt werden, dass das Alter bei Symptombeginn aus den publizierten individuellen Patientendaten der Sekundärquellen entnommen worden ist.

Der pU identifizierte im Rahmen der Confounder-Recherche krankheitsbezogene und krankheitsdynamische Confounder, die insbesondere aufgrund der Datenverfügbarkeit nicht berücksichtigt werden konnten. Die Genotyp-Phänotyp-Korrelation ist laut verfügbarer Literatur nicht eindeutig, jedoch scheinen unterschiedliche potentielle Outcomes hinsichtlich des Genotyps vorzuliegen. Es liegt jedoch keine belegte Monokausalität vor, sodass eine Adjustierung nur für den Genotyp oder der Aktivität des Enzyms ebenfalls keine hinreichende Adjustierung darstellen würde. Jedoch kann nicht nachvollzogen werden, weshalb keine deskriptive Darstellung der identifizierten Genotypen in den Interventionsstudien erfolgte.

Die genannten Limitationen zur Erfüllung der Grundbedingungen (Fallzahl, Vollständigkeit der Confounder, Positivität, „Time Zero“) führen dazu, dass der berechnete Effektschätzer auf Basis eines ATO-Schätzers für Aussagen zur Wirksamkeit nicht sinnvoll interpretierbar ist. Dieser bezieht sich nicht mehr auf die gesamte hergeleitete Vergleichspopulation, sondern auf eine konstruierte Population, die nicht klar beschreibbar ist bzw. zu der keine Beschreibung vorliegt. Es ist unklar, ob der ATO-Ansatz für Populationen mit grundsätzlichen Positivitätsproblemen im Kontext einer Nutzenbewertung geeignet ist. Dieser gewichtet die Populationen so, dass diese in einem klinischen Equipoise (Gleichgewicht) sind. Der geschätzte Effekt bezieht sich auf die Personen, die die größtmögliche interne Validität aufweisen. Hierbei bleibt jedoch nicht nachvollziehbar, welche Population abgebildet wird und wie diese das Anwendungsgebiet widerspiegelt. Die Wahl des ATO als „Target of Inference“ erfolgte, weil die Schätzung des ATE durch IPTW nicht möglich war. Dies deutet in der Gesamtschau auf inhärente statistische Positivitätsprobleme hin.

Anhand der vorgelegten Unterlagen kann nicht abschließend beurteilt werden, inwiefern sich die erzeugte Pseudopopulation von der Gesamtpopulation unterscheidet und ob sich die Ergebnisse für die Pseudopopulation auf die Gesamtpopulation übertragen lassen. Der pU stellt die Hazard Ratios und p-Werte dar, ohne dabei auf die methodischen Besonderheiten der Post-hoc-Herleitung und die Einschränkungen der Interpretierbarkeit und insbesondere der Verallgemeinerbarkeit einzugehen. In dem beigefügten Analyseplan wird die gewählte Gewichtungsmethodik und das „Target of Inference“ nicht beschrieben. Nähere Informationen zu den nicht konvergierten Modellen konnten nicht identifiziert werden. Eine ergebnisgesteuerte Berichterstattung kann nicht ausgeschlossen werden. Als weitere Limitation wird die geringe Stichprobengröße in beiden Populationen gesehen, welche die Validität von PS-Verfahren mittels Gewichtung durch IPTW einschränkt.

Die herangezogenen statistischen Methoden zur Berechnung der PS erfordern zum Teil sehr starke Annahmen an die Datensituation. Die fehlende Berücksichtigung von relevanten Confoundern und die Unklarheiten zum Beobachtungsstart („Time Zero“), des Immortal Time Bias sowie die nicht dargelegte Positivität können nicht durch komplexe statistische Methoden mit unbegründeten Entscheidungsstrukturen zur Modellwahl aufgewogen werden. Auf Basis der vorliegenden Unterlagen führen die eingereichten Analysen zu dem Modell, das am besten zu der vorliegenden Datensituation und der herangezogenen Entscheidungsstruktur passt. Eine datengetriebene Modellauswahl kann auf Basis der vorliegenden Unterlagen nicht ausgeschlossen werden. Im Kontext der Nutzenbewertung ist ein verallgemeinerbares Modell von Relevanz.

Es ist auf Basis der vorliegenden Unterlagen unklar, ob eine deskriptive unadjustierte Darstellung des natürlichen Verlaufs der Erkrankung zur Kontextualisierung des gesehenen Behandlungseffekts sinnvoll ist, da durch den vorliegenden Immortal Time Bias das potentielle Outcome in der Interventionspopulation nicht sinnvoll dargestellt werden kann.

EP-imputiert-IPTW: Zeit bis zum Tod seit Behandlungsbeginn

Es wurde eine weitere Analyse identifiziert, die für die unbehandelte Population einen Zeitpunkt eines synthetischen Behandlungsbeginns imputiert. Bei der Imputation werden für jede unbehandelte Person alle möglichen Datumsangaben zum Behandlungsbeginn in Schritte von jeweils einem Monat Abstand zwischen dem Alter bei Symptombeginn und dem Datum des Todes bzw. des letzten bekannten Lebensstatus berücksichtigt. Es wird nach dem Zufallsprinzip sowie der Einteilung nach Kategorien (Alter bei Symptombeginn ≤ 2 Jahre und > 2 Jahre bis ≤ 12 Jahre) für jede behandelte Person eine unbehandelte Person ausgewählt. Die weitere Analyse erfolgte analog zu der bereits beschriebenen Analyse mittels IPTW und der Schätzung des ATO. Die Modelldiagnostik erfolgte mittels SMD und graphischer Darstellung der PS nach dem Gewichtungsverfahren. Die Positivitätsannahme für die adjustierte Analyse gilt als nicht dargelegt, da für die adjustierte Population keine ausreichende Überlappung vorliegt. Die vorgelegten Auswertungen sind im Sinne der Nutzenbewertung weder interpretierbar noch verallgemeinerbar.

Zusammenfassende Bewertung der Methodik des indirekten Vergleichs

Im Rahmen der Evidenzsynthese konnten mehrere indirekte Vergleiche identifiziert werden. Der pU konzipiert einen „Target Trial“ und zieht diesen für eine Ableitung des Zusatznutzens heran. Die Bewertung in der Nutzenbewertung bezieht sich auf den indirekten Vergleich in der zulassungskonformen Population in der Endpunktkategorie „Mortalität“.

Es liegen insbesondere retrospektive Daten sowohl für die Interventions- als auch für die Kontrollpopulation vor. Die Unterschiede in den Ein- und Ausschlusskriterien, in den Daten und Mechanismen für die Aufnahme in die klinische Studie sowie in der Vergleichspopulation begünstigen einen Selektionsbias und resultieren in potentiell systematisch unterschiedlichen Populationen für einen nicht randomisierten Vergleich ohne Brückenkomparator. Durch das Vorliegen einer Visite in einem bestimmten Zeitkorridor für die Interventionspopulation der Studie 101 liegt ein Immortal Time Bias für mindestens 50 % der Interventionspopulation vor.

Der pU legt eine systematische Literaturrecherche nach Confoundern sowie anschließender Einordnung der Relevanz durch Experten vor. Es wurden nicht alle als relevant bewerteten Confounder ins PS-Modell aufgenommen.

Auf Basis der vorgelegten Datensituation (berichtete Baseline-Charakteristika, Immortal Time Bias, Confounder-Auswahl, „Time Zero“, Aktualität, Berichterstattung des PS-Modells) kann nicht davon ausgegangen werden, dass eine ausreichende statistische Positivität der beiden Populationen für eine adjustierte Analyse vorhanden ist. Die Erfüllung der notwendigen SITA-Bedingung (kein ungemessenes Confounding und bspw. ein PS zwischen 0 und 1) wird als nicht dargelegt angesehen. Der Effekt im Kontrollarm kann auf Basis der vorliegenden Unterlagen

nicht als kontrafaktisches Ergebnis (auf Basis einer kontrafaktischen Analyse; „Counterfactual Outcome“) für Patientinnen und Patienten des Studienarms gesehen werden.

Die Interpretation eines berechneten Behandlungsunterschieds erscheint nicht sachgerecht. In der Gesamtschau sind die vom pU vorgelegten Ergebnisse nicht für die Quantifizierung eines Zusatznutzens geeignet und werden nicht für die Nutzenbewertung herangezogen.