

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: Givinostat

Datum der Veröffentlichung: 15. Oktober 2025

Version 1.0

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abbildungsverzeichnis	5
Abkürzungsverzeichnis	6
Hintergrund.....	7
1 Fragestellung.....	8
2 Auswahl und Methodik der relevanten Studien	9
2.1 Studienbasis für die Nutzenbewertung	9
2.2 Beschreibung der eingeschlossenen Studie	9
2.3 Endpunkte	18
2.3.1 Mortalität.....	19
2.3.2 Morbidität.....	19
2.3.3 Lebensqualität	31
2.3.4 Sicherheit	32
2.3.5 Erhebungszeitpunkte	35
2.4 Statistische Methoden.....	36
2.5 Verzerrungspotential auf Studien- und Endpunktebene	42
2.6 Indirekter Vergleich	44
3 Ergebnisse der eingeschlossenen Studie	46
3.1 Studiencharakteristika und Studienmedikation	46
3.2 Mortalität	49
3.3 Morbidität	49
3.4 Lebensqualität.....	59
3.5 Sicherheit	59
4 Diskussion der Methodik und Ergebnisse.....	62
4.1 Zulassungsstatus und Zulassungspopulation.....	65
4.2 Mortalität	66
4.3 Morbidität	66
4.4 Lebensqualität.....	68
4.5 Sicherheit	69
5 Anforderung an eine qualitätsgesicherte Anwendung	71
6 Zusammenfassung der Nutzenbewertung	72
Referenzen	73
Anhang	76

Tabellenverzeichnis

Tabelle 1:	Dosierungsempfehlungen gemäß Fachinformation von Givinostat.....	8
Tabelle 2:	Übersicht über die Studienbasis	9
Tabelle 3:	Charakterisierung der Studie EPIDYS.....	9
Tabelle 4:	Für die Nutzenbewertung relevante Protokolländerungen der Studie EPIDYS.....	13
Tabelle 5:	Charakterisierung der Intervention	15
Tabelle 6:	Startdosis und Dosismodifikationen von Givinostat und Placebo unter Protokollversion ≤ 4.0 und ≥ 6.0	17
Tabelle 7:	Zusammenfassung der Endpunktbewertung der Studie EPIDYS.....	18
Tabelle 8:	Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie EPIDYS	35
Tabelle 9:	Verzerrungspotential der Studie EPIDYS	42
Tabelle 10:	Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie EPIDYS.....	43
Tabelle 11:	Allgemeine Angaben; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)	46
Tabelle 12:	Charakterisierung der Studienpopulation; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	47
Tabelle 13:	Angaben zur Dosismodifikationen mit der Studienmedikation; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)	48
Tabelle 14:	4SC (Zeit in Sekunden), MMRM-Analysen anhand beobachteter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) ...	49
Tabelle 15:	RFTF (Zeit in Sekunden), MMRM-Analysen anhand beobachteter Werte zu Woche 48; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) ...	51
Tabelle 16:	10-MWT (Zeit in Sekunden), MMRM-Analysen anhand beobachteter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)....	53
Tabelle 17:	6-MWT (Distanz in Metern), MMRM-Analysen anhand beobachteter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) ...	55
Tabelle 18:	Bestätigter Verlust der Aufstehfähigkeit bis Studienende (Woche 72 / EOS), Überlebenszeitanalyse; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	57
Tabelle 19:	Verlust der Gehfähigkeit ohne Hilfsmittel zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)	59
Tabelle 20:	Ergebnisse der Nebenwirkungen – Zusammenfassung der UE; Studie EPIDYS, Overall Safety Analysis Set (finaler Datenschnitt: 22.02.2022).....	59
Tabelle 21:	UE mit Inzidenz ≥ 10 % oder bei mind. 10 Personen in einem Behandlungsarm; Studie EPIDYS, Overall Safety Analysis Set (finaler Datenschnitt: 22.02.2022).....	60
Tabelle 22:	Zusammenfassende Darstellung der Ergebnisse der Studie EPIDYS der Overall-Population bis Woche 72 / EOS (finaler Datenschnitt: 22.02.2022).....	72

Tabelle 23: 4SC (Zeit in Sekunden), ANCOVA anhand imputierter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)	76
Tabelle 24: RFTF (Zeit in Sekunden), ANCOVA anhand imputierter Werte zu Woche 48; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)	78
Tabelle 25: 10-MWT (Zeit in Sekunden), ANCOVA anhand imputierter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)	80
Tabelle 26: 6-MWT (Distanz in Metern), ANCOVA anhand imputierter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)	82

Abbildungsverzeichnis

Abbildung 1: 4SC (Zeit in Sekunden), MMRM-Analyse anhand beobachteter Werte bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	50
Abbildung 2: RFTF (Zeit in Sekunden), MMRM-Analyse anhand beobachteter Werte bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	52
Abbildung 3: 10-MWT (Zeit in Sekunden), MMRM-Analyse anhand beobachteter Werte bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	54
Abbildung 4: 6-MWT (Distanz in Metern), MMRM-Analyse anhand beobachteter Werte bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	56
Abbildung 5: Bestätigter Verlust der Aufstehfähigkeit bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	58
Abbildung 6: 4SC (Zeit in Sekunden) bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	77
Abbildung 7: RFTF (Zeit in Sekunden) bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	79
Abbildung 8: 10-MWT (Zeit in Sekunden) bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	81
Abbildung 9: 6-MWT (Distanz in Meter) bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]	83

Abkürzungsverzeichnis

4SC	4-Stair Climb
6-MWT	6-Minute Walk Test
10-MWT	10-Meter Walk/Run Test
ANCOVA	Kovarianzanalyse (Analysis of Covariance)
CINRG DNHS	Cooperative International Neuromuscular Research Group – Duchenne Natural History Study
CTCAE	Common Terminology Criteria for Adverse Events
DMD	Duchenne-Muskeldystrophie
EMA	European Medicines Agency
EOS	End of Study
EPAR	European Public Assessment Report
FDA	U.S. Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
k. A.	keine Angabe
KI	Konfidenzintervall
LOCF	Last Observation Carried Forward
LS	Least Squares
MAR	Missing at Random
MMRM	Mixed Model for Repeated Measures
MRS	Magnetresonanztomographie
MRT	Magnetresonanztomografie
MW	Mittelwert
N	Anzahl
NSAA	North Star Ambulatory Assessment
RFTF	Rise From The Floor
PODCI	Pediatric Outcomes Data Collection Instrument
PS	Propensity Score
PT	Preferred Term
pU	pharmazeutischer Unternehmer
RCT	Randomisierte kontrollierte Studie
RR	Relatives Risiko
SAP	Statistischer Analyseplan
SD	Standardabweichung
SGB	Sozialgesetzbuch
SOC	Systemorganklasse
SUE	Schwerwiegendes unerwünschtes Ereignis
UE	Unerwünschtes Ereignis
VerfO	Verfahrensordnung des G-BA

Hintergrund

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

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 G-BA bestimmt gemäß Kapitel 5 § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung des G-BA (VerfO) das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Nutzenbewertung). Zur Bestimmung des Ausmaßes des Zusatznutzens von Givinostat zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR), die Studien, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers (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 Givinostat in seiner Sitzung am 7. Oktober 2025 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 15. Juli 2025 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 15. Oktober 2025 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des IQWiG auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Es wird darüber hinaus eine mündliche Anhörung durchgeführt. Über das Ausmaß des Zusatznutzens beschließt der G-BA innerhalb von drei Monaten nach Veröffentlichung der Nutzenbewertung.

1 Fragestellung

Givinostat (Duvyzat®) 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 G-BA 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 [21]:

Behandlung der Duchenne-Muskeldystrophie (DMD) bei gehfähigen Patienten ab 6 Jahren zusammen mit einer Corticosteroid-Behandlung.

Gemäß Fachinformation wird die zugelassene Dosierung vom Körpergewicht abhängig berechnet (Tabelle 1) und zweimal täglich eingenommen. Bei Auftreten von unerwünschten Ereignissen ist eine zweimalige Dosisreduktion, die ebenfalls Körpergewicht-abhängig berechnet wird, möglich. Die Entscheidung, ob die Behandlung bei Patienten, die ihre Gehfähigkeit verlieren, fortgesetzt werden soll, wird nach ärztlichem Ermessen nach Beurteilung der gesamten Vorteile und Risiken getroffen.

Tabelle 1: Dosierungsempfehlungen gemäß Fachinformation von Givinostat [21]

Gewicht¹⁾	Dosierung	Menge der oralen Suspension
15 kg bis < 20 kg	22,2 mg zweimal täglich	2,5 ml zweimal täglich
20 kg bis < 40 kg	31 mg zweimal täglich	3,5 ml zweimal täglich
40 kg bis < 60 kg	44,3 mg zweimal täglich	5 ml zweimal täglich
Ab 60 kg	53,2 mg zweimal täglich	6 ml zweimal täglich

¹⁾ Basierend auf dem derzeitigen Körpergewicht.

2 Auswahl und Methodik der relevanten Studien

2.1 Studienbasis für die Nutzenbewertung

Tabelle 2: Übersicht über die Studienbasis

Studienname (Nummer)	Studie mit Dossier vom pU eingereicht	Studie vom pU als relevant für die Ableitung des Zusatznutzens erachtet	Studie relevant für die Nutzen- bewertung	Ausschlussgrund
Studien zum Wirkstoff				
EPIDYS (Studie 48) ¹⁾	Ja	Ja	Ja	-
Studien zu externen Kontrollen				
ImagingDMD ²⁾ + CINRG DNHS ²⁾	Ja	Nein	Nein	Siehe Ausführungen in Kapitel 2.6.

¹⁾ Zulassungsrelevante Studie gemäß EPAR.

²⁾ Im Rahmen des Zulassungsverfahrens bei der EMA ergänzend zu den Zulassungsstudien bewertet.

Abkürzungen: CINRG DNHS: Cooperative International Neuromuscular Research Group – Duchenne Natural History Study; EMA: European Medicines Agency; EPAR: European Public Assessment Report; EPIDYS: Epigenetic Rescue of Dystrophin Dysfunction; pU: pharmazeutischer Unternehmer.

Zur Nutzenbewertung für Givinostat herangezogene Unterlagen

- Dossier des pU zu Givinostat [20]
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report [4]
- Studienbericht inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie EPIDYS [13,14,15,16,17]
- SAP und Tabellen zu den Ergebnissen der integrierten Datenanalyse [12,18,19]
- Fachinformation zu Givinostat [21]

2.2 Beschreibung der eingeschlossenen Studie

Die Nutzenbewertung von Givinostat im vorliegenden Anwendungsgebiet basiert auf der Zulassungsstudie EPIDYS. Die Studie und die Intervention werden in Tabelle 3 bzw. Tabelle 5 dargestellt.

Tabelle 3: Charakterisierung der Studie EPIDYS

Charakteris- tikum	Beschreibung
Design und Studienablauf	Randomisierte, doppelblinde, Placebo-kontrollierte, multizentrische Phase-III-Studie zur Bewertung der Wirksamkeit und Sicherheit von Givinostat bei gehfähigen männlichen Patienten mit DMD im Alter von ≥ 6 Jahren. Die Randomisierung erfolgte im Verhältnis 2:1 in die Behandlungsarme „Intervention“ und „Placebo“, stratifiziert nach den eingesetzten Glukokortikoiden: • Deflazacort, tägliche Einnahme • Deflazacort, intermittierende Einnahme • Andere Corticosteroide, tägliche Einnahme • Andere Corticosteroide, intermittierende Einnahme

Charakteris- tikum	Beschreibung
	Die Overall-Population wurde in 2 Populationen eingeteilt: • Off-Target-Population: Baseline-Wert des MFF im M. vastus lateralis $\leq 5\%$ und $> 30\%$ • On-Target-Population: Baseline-Wert des MFF im M. vastus lateralis $> 5\%$ bis $\leq 30\%$ Studienphasen (vorgesehene Studiendauer: 19 Monate) • Screening¹⁾: 4 Wochen vor Randomisierung • Behandlung: 18 Monate Die Studie ist abgeschlossen. Die Patienten konnten im Anschluss an der einarmigen Nachbeobachtungsstudie 51 teilnehmen. Für diese liegen Ergebnisse zum Interimsdatenschnitt (31.12.2023) für einen Zeitraum von ca. 3,3 Jahren vor.
Population	Wesentliche Einschlusskriterien • Gehfähige männliche Patienten. • Alter ≥ 6 Jahre bei Randomisierung. • Charakteristisch klinische Symptome oder Anzeichen einer DMD, die bereits beim Screening vorhanden sind (z. B. proximale Muskelschwäche, Gowers-Zeichen, erhöhter Serum-Kreatinkinase-Spiegel). • Durch Gentest bestätigte DMD-Diagnose. • In der Lage, zwei 4SC-Tests während des Screenings durchzuführen: ○ Ergebnisse müssen ± 1 Sekunde beieinander liegen ○ Mittelwert: ≤ 8 Sekunden • In der Lage den RFTF in ≥ 3 bis < 10 Sekunden während des Screenings durchzuführen. • Manueller Muskeltest des Quadrizeps beim Screening: $\geq$ Grad 3-. • ≥ 6 Monate systemische Corticosteroid-Therapie unmittelbar vor Studienbeginn, in diesem Zeitraum ohne signifikante Änderungen an: ○ Corticosteroid-Typ ○ Dosierungsschema ○ Dosierung (mit Ausnahme von Veränderungen aufgrund Gewichtsänderung) • Begründete Annahme, dass sich Corticosteroid-Dosierung und Dosierungsschema für die Dauer der Studie nicht signifikant ändern werden. Wesentliche Ausschlusskriterien • Behandlung mit einem anderen Prüfpräparat innerhalb von 3 Monaten vor Beginn der Studienbehandlung (Ausnahme: Deflazacort in den USA im Rahmen des „Expanded Access Program“ und in Kanada im Rahmen des „Special Access Program“). • Behandlung mit Idebenon innerhalb von 3 Monaten vor Beginn der Studienbehandlung. • Behandlung mit einem Medikament zur Dystrophin-Wiederherstellung (z. B. Ataluren, Exon-Skipping) innerhalb von 6 Monaten vor Beginn der Studienbehandlung. • Anwendung jeglicher pharmakologischer Therapie, mit Ausnahme von Corticosteroiden, mit möglicher Wirkung auf die Muskelkraft oder -funktion innerhalb von 3 Monaten vor Beginn der Studienbehandlung (z. B. Wachstumshormone): ○ Vitamin D, Calcium und andere Nahrungsergänzungsmittel waren erlaubt, bei gleichbleibender Dosis für mind. 4 Monate vor Erhalt der Studienmedikation; ○ Anwendung von Testosteron als Ersatztherapie zur Behandlung einer verzögerten Pubertät war möglich, bei gleichbleibender Dosis und Einnahmeschema für mind. 6 Monate und der Testosteronspiegel war innerhalb der Normbereiche für das Alter des Patienten. • Durchführung einer Operation, die sich auf die Muskelkraft oder -funktion auswirken könnte, innerhalb von 3 Monaten vor Studienbeginn oder eine geplante Operation zu einem beliebigen Zeitpunkt während der Studie.

Charakteris- tikum	Beschreibung																				
	• Verlust von ≥ 30 Grad Plantarflexion gegenüber dem normalen Bewegungsumfang am Sprunggelenk aufgrund einer Kontraktur (d. h. unveränderter Verlust von mehr als 10 Grad Plantarflexion von der Fußsohle, bei Annahme eines normalen Dorsalflexionsbereichs von 20 Grad). • Änderung der Kontrakturbehandlung, wie z. B. Seriengips, Kontrakturkontrollgeräte, Nachtschienen, Dehnübungen (passiv, aktiv, selbst) innerhalb von 3 Monaten vor der Aufnahme in die Studie oder Erwartung der Notwendigkeit eines solchen Eingriffs während der Studie. • Vorliegen einer anderen klinisch signifikanten Erkrankung, die nach Beurteilung des Prüfpersonals die Sicherheit des Patienten beeinträchtigen könnte, sodass es unwahrscheinlich ist, dass der Behandlungsverlauf oder die Nachsorge abgeschlossen wird, oder die Bewertung der Studienergebnisse beeinträchtigen könnte. • Vorliegen einer Diagnose anderer unkontrollierter neurologischer Erkrankungen oder relevanter unkontrollierter somatischer Erkrankungen, die nicht mit DMD zusammenhängen. • Thrombozytenzahl, Leukozytenzahl und Hämoglobin $< LLN$ während des Screenings (bei Werten außerhalb des Normbereichs sollten Thrombozytenzahl, Leukozytenzahl und Hämoglobin einmalig wiederholt gemessen werden; Ausschluss des Patienten bei wiederholtem Testergebnis $< LLN$). • Symptomatische Kardiomyopathie oder Herzinsuffizienz (NYHA-Klasse III oder IV) oder linksventrikuläre Ejektionsfraktion $< 50\%$ beim Screening. • Aktuelle oder frühere Lebererkrankung oder -beeinträchtigung, einschließlich, aber nicht beschränkt auf, erhöhte Gesamtbilirubin-Werte (d. h. $> 1,5 \times ULN$); Ausnahme war als Ursache die Gilbert-Krankheit oder sekundär durch das Gilbert-Syndrom bedingt. • Unzureichende Nierenfunktion, definiert als Serum-Cystatin C $> 2 \times ULN$ (falls der Wert $> 2 \times ULN$ beträgt, kann einmalig eine erneute Messung vorgenommen werden, falls nach einer erneuten Messung der Wert $> 2 \times ULN$ liegt, wurde der Patient ausgeschlossen). • Nüchtern ermittelte Triglycerid-Werte > 300 mg/dl (3,42 mmol/l) während des Screenings. • Vorliegen eines Baseline-korrigierten QT-Intervalls, nach Fridericias Formel (QTcF) > 450 ms (Mittelwert von 3 aufeinanderfolgenden Messwerten im Abstand von 5 Minuten) oder eine medizinische Vorgeschichte weiterer Risikofaktoren für Torsades de pointes (z. B. Herzinsuffizienz, Hypokaliämie oder Familienanamnese mit langem QT-Syndrom). • Psychiatrische Erkrankung oder soziale Situation, wodurch der Patient unfähig ist, die Muskelfunktionstests und/oder die Verfahren des Studienprotokolls zu verstehen und einzuhalten.																				
Intervention und Zahl der Patientinnen und Patienten	<table> <tr> <td>Gescreent gesamt:</td><td>N = 359</td></tr> <tr> <td>• Randomisiert/Eingeschlossen:</td><td>N = 179</td></tr> <tr> <td> ○ On-Target-Population:</td><td>n = 120</td></tr> <tr> <td> ○ Off-Target-Population:</td><td>n = 59</td></tr> <tr> <td>• Intervention:</td><td>N = 118</td></tr> <tr> <td> ○ On-Target-Population:</td><td>n = 81</td></tr> <tr> <td> ○ Off-Target-Population:</td><td>n = 37</td></tr> <tr> <td>• Kontrolle:</td><td>N = 61</td></tr> <tr> <td> ○ On-Target-Population:</td><td>n = 39</td></tr> <tr> <td> ○ Off-Target-Population:</td><td>n = 22</td></tr> </table>	Gescreent gesamt:	N = 359	• Randomisiert/Eingeschlossen:	N = 179	○ On-Target-Population:	n = 120	○ Off-Target-Population:	n = 59	• Intervention:	N = 118	○ On-Target-Population:	n = 81	○ Off-Target-Population:	n = 37	• Kontrolle:	N = 61	○ On-Target-Population:	n = 39	○ Off-Target-Population:	n = 22
Gescreent gesamt:	N = 359																				
• Randomisiert/Eingeschlossen:	N = 179																				
○ On-Target-Population:	n = 120																				
○ Off-Target-Population:	n = 59																				
• Intervention:	N = 118																				
○ On-Target-Population:	n = 81																				
○ Off-Target-Population:	n = 37																				
• Kontrolle:	N = 61																				
○ On-Target-Population:	n = 39																				
○ Off-Target-Population:	n = 22																				

Charakteristikum	Beschreibung
Ort und Zeitraum der Durchführung; Datenschnitte	Studienzentren 50 Studienzentren in elf Ländern (Belgien, Deutschland, Frankreich, Israel, Italien, Kanada, Niederlande, Serbien, Spanien, USA, Vereinigtes Königreich). Studienzeitraum • Erster Patient, erste Visite: 06.06.2017 • Letzter Patient, letzte Visite: 22.02.2022 Datenschnitte • Interimsdatenschnitt/Futilitätsanalyse: 14.01.2020 (nach 12 Monaten Behandlungsphase der ersten 50 randomisierten Patienten) • Finaler Datenschnitt: 22.02.2022
Endpunkte gemäß Studienprotokoll	Primärer Endpunkt Treppensteigen von 4 Stufen (4SC): Mittlere Änderung nach 18-monatiger Behandlung gegenüber Baseline. Sekundäre Endpunkte • Funktionstests: Mittlere Änderung nach 18-monatiger Behandlung ggü. Baseline: ○ Aufstehen aus der Rückenlage (RFTF) ○ 6-Minuten-Gehtest (6-MWT) ○ Funktionelle Leistungsfähigkeit (NSAA) ○ Muskelkraft, bewertet durch Kniestreckung, und Ellenbogenflexion, erhoben durch handgeführte Myometrie • Bildgebung: Änderung des MFF im M. vastus lateralis • Sicherheit ○ Unerwünschte Ereignisse ○ Atemfunktion gemessen mittels FVC1, FEV1, FVC/FEV1 ○ Herzfunktion mittels EKG ○ Kognitive Funktion mittels Raven-Farb-Matrizen-test²⁾ ○ Körpergewicht und -größe, BMI ○ Bewertung der Akzeptanz/Geschmacksqualität der Suspension zum Einnehmen Explorative Endpunkte • Mittlere Änderung: ○ 10-Meter-Geh-/Lauftest (10-MWT) ○ Körperliche Funktionsfähigkeit (PODCI) ○ Prozent des Sollwerts der vorhergesagten 6-MWT • Zeit bis zu einer anhaltenden Verschlechterung um 10 % im 6-MWT • Bestätigter Verlust der Aufstehfähigkeit (PLOS) • Verlust der Gehfähigkeit ohne Hilfsmittel (LOA) • Untersuchung einer Korrelation zwischen der Wirksamkeit von Givinostat und Krankheitsprogression, DMD-Mutation, dem LTBP4- und Osteopontin-Genotypen • Evaluation möglicher Serum-Biomarker für DMD.

¹⁾ Sofern ein Patient die Ein-/Ausschlusskriterien nicht erfüllte, konnte er nach frühestens 3 Monaten erneut am Screening teilnehmen.

²⁾ Der Endpunkt wurde vom pU in Modul 4 nicht herangezogen. Dieser Einschätzung wird gefolgt.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; DMD: Duchenne-Muskeldystrophie; FEV1: Forced Expiratory Volume per second; FVC: Forced Vital Capacity; LLN: Lower Limit of Normal; LOA: Loss of Ambulation; MFF: Muskelfettanteil; NYHA: New York Heart Association; PLOS: Persistent Loss Of Standing; pU: pharmazeutischer Unternehmer; RFTF: Rise From The Floor; ULN: Upper Limit of Normal.

Protokolländerungen

Es wurden 7 Änderungen des Originalprotokolls vom 07.05.2016 vorgenommen. Es werden die für die Nutzenbewertung relevanten Protokolländerungen ab dem Einschluss des ersten Teilnehmenden (06.06.2017) dargestellt.

Tabelle 4: Für die Nutzenbewertung relevante Protokolländerungen der Studie EPIDYS

Protokollversion	Wesentliche Änderungen
4.0 vom 12.07.2017 (Anzahl bis dahin eingeschlossene Patienten: k. A.)	• Berücksichtigung relevanter Kommentare der FDA zum „May Proceed Letter“ (vom 18.10.2016): Non-Compliance der Studienmedikationseinnahme war kein Ausschlussgrund der Studie. Stattdessen sollte jeder Fall mit dem medizinischen Studienmonitor besprochen werden. Ist die Sicherheit des Teilnehmenden hinreichend gewährleistet, sollte der Teilnehmende nicht aus der Studie ausgeschlossen werden. • Ausschlusskriterium 6: Änderungen der Kriterien für Knöchelgelenkskontrakturen für eine bessere Anwendbarkeit auf die On-Target-Population und zur Erhöhung der Anzahl an für die Studie geeigneten Patienten mit DMD¹⁾. • Hinzufügen eines Gentests zur Bestätigung der DMD-Diagnose (wie im entsprechenden Einschlusskriterium gefordert) während des Screenings.
6.0 vom 19.10.2018 (Anzahl bis dahin eingeschlossene Patienten: Givinostat-Arm: n = 55 (46,7 %), Placebo-Arm: n = 27 (44,3%) ²⁾)	• Reduktion der Anfangsdosis um 1/3, aufgrund auftretender Thrombozytopenien bei 50–60 % der Patienten, die ≥ 8 Wochen die Studienbehandlung erhalten haben. Zusätzlich wurden Angaben zur Dosismodifikation angepasst (siehe auch Tabelle 5 und Tabelle 7). • Hinzufügen eines neuen Ausschlusskriteriums (Triglycerid-Konzentration von > 300 mg/dl (3,42 mmol/l)), aufgrund erhöhter Triglycerid-Werte bei einigen Teilnehmern. Als Schwellenwert wurde der Wert zur Abgrenzung von milder vs. moderater Hypertriglyceridämie gemäß CTCAE (Version 4.03) gewählt. • Bei Anstieg von Triglycerid-Konzentration > 300 mg/dl (3,42 mmol/l) während des Studienverlaufs, vorübergehender Behandlungsunterbrechung. • Zur verbesserten Überwachung von Schilddrüsenhormonen (TSH, FT3, FT4) werden diese in den ersten 3 Monaten monatlich und danach alle 3 Monate bis zum Studienende (Woche 72 / EOS) erhoben. • Umsetzung der IDMC-Empfehlung zur Erhebung des vollständigen Blutbilds. Blutentnahme in den 8 aufeinanderfolgenden Wochen bei sicherheitsbedingter Dosisreduktion aufgrund verringerter Werte an Thrombozyten ($\leq 150 \times 10^9/l$) und/oder Leukozyten ($< 3,0 \times 10^9/l$) und/oder Hämoglobin ($< 10,0$ mg/dl). • Änderung des Einschlusskriteriums „MMT des Quadrizeps beim Screening $\geq$ Grad 3“ auf „MMT des Quadrizeps beim Screening $\geq$ Grad 3-“, um die Aufnahme einer größeren Anzahl von Patienten in die Studie zu ermöglichen. • Zulassen von bis zu 2 Re-Screenings (≥ 3 Monate Abstand). • Erweiterung des Einschlusskriteriums MFF im M. vastus lateralis von „> 10 % bis ≤ 30 %“ auf „> 5 % bis ≤ 30 %“, aufgrund der hohen Screening-Ausschlussrate $\rightarrow$ On-Target-Population. • Ermöglichen der Aufnahme von ≤ 50 Patienten mit MFF im M. vastus lateralis außerhalb des definierten Bereichs (d. h. ≤ 5 % und > 30 %) $\rightarrow$ Off-Target-Population zur Lieferung unterstützender Daten. • MFF-Erhebung durchgehend mit MRT (vorher Algorithmus)³⁾. • Um die Studienbelastung der Patienten/Familien zu minimieren, dürfen einige Sicherheits- / außerplanmäßigen Untersuchungen im Zuhause der Patienten stattfinden. • Zeitpunkt der Screening-Visite 2 wurde aktualisiert, sodass diese 1–14 Tage nach Visite 1 stattfinden kann.

Protokollversion	Wesentliche Änderungen
COVID-Anpassungen (12.03.2020) (Anzahl bis dahin eingeschlossene Patienten: k. A.)	• Länderspezifische Anpassungen waren vorgesehen. • Screening und Randomisierung: Sofern möglich und keine Restriktionen bestehen, sollen die Visiten im Krankenhaus/Studienzentrum stattfinden. Falls ein MRT-Scan nicht verfügbar ist, sollte mit dem Screening bis zur Durchführung des MRT-Scans ausgesetzt werden. Eine Randomisierung ist möglich, selbst wenn die Screening-Visite (V1) außerhalb des Zeitfensters liegt (bspw. bis zu 3 Monate). Jeder Fall sollte individuell besprochen werden. • Eine Verschiebung der geplanten Visite aufgrund der COVID-19-Restriktionen innerhalb des vorgesehen Zeitfensters war erlaubt. Falls dies nicht möglich war, konnte die Studienvisite um zusätzliche 2 Wochen verschoben werden, sofern der Patient ausreichend Studienmedikation zur Verfügung hatte. • Falls die Teilnahme, wie oben beschrieben, dennoch nicht möglich war, war eine Sondervisite bei den Patienten daheim möglich: ○ Folgende Erhebungen waren erlaubt: u. a. Blutentnahme, Erhebung von Vitalzeichen, EKG, Kopie des Patiententagebuchs, Überprüfung des Gesundheitszustands, Erhebung von UE; ○ folgende Untersuchungen wurden ausgeschlossen: u. a. Körpergewicht u. -größe, Funktionstests, respiratorische Funktionstests, PODCI, Urin-Untersuchungen. • Visite zu Woche 72 / Monat 18 (EOS): Es sollten Anstrengungen zur Durchführung der EOS-Visite unternommen werden. Die Studienvisite konnte um bis zu 2 Monate verzögert werden. • Die oben genannten Änderungen sollten nicht als Protokollverletzungen dokumentiert werden.
8.0 vom 08.04.2020 (Anzahl bis dahin eingeschlossene Patienten: k. A.)	• Fortsetzung der Studie EPIDYS auf Grundlage der Interims-/Futilitätsanalyse von Januar 2020. • Fallzahlenanpassung: Änderung der Effektgröße „Veränderung im 4SC (Baseline vs. 18 Monate)“ von 3 Sekunden auf 2 Sekunden. • Streichung der geplanten 2. Interimsanalyse, wie von der FDA empfohlen.

¹⁾ Änderung des Wortlauts von „Kontraktur des Sprunggelenks aufgrund eines fixierten Verlustes von $\geq 10^\circ$ der Dorsalextension aus Neutralstellung, unter Annahme einer normalen Dorsalextension von 20° “ zu „Verlust von $\geq 30^\circ$ Plantarflexion gegenüber dem normalen Bewegungsumfang des Sprunggelenks aufgrund von Kontrakturen (d. h. fixierter Verlust von $\geq 10^\circ$ der Plantarflexion aus Neutralstellung, unter Annahme einer normalen Dorsalextension von 20°)“.

²⁾ Angaben entnommen aus den für das Dossier durchgeführten Subgruppenanalysen („Höhere Dosis ($\leq$ Protokollversion 4.0) vs. Niedrigere Dosis ($>$ Protokollversion 4.0)), bis Protokollversion 4.0 von Version 6.0 abgelöst wurde.

³⁾ Die Erhebung des Kriteriums „MFF im M. vastus lateralis > 10 bis $< 30^\circ$ “ erfolgte über einen Algorithmus, welcher den Wert anhand der kombinierten Ergebnisse des 4SC, 6-MWT sowie dem maximalen Drehmoment der Strecker im Kniegelenk schätzte. Für Patienten, bei denen der geschätzte MFF im M. vastus lateralis zwischen 10 und 15 % lag, sollte der Wert per MRT bestätigt werden.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; CTCAE: Common Terminology Criteria for Adverse Events; DMD: Duchenne-Muskeldystrophie; EOS: End of Study; FDA: U.S. Food and Drug Administration; IDMC: Independent Data Monitoring Committee; k. A.: keine Angabe; MFF: Muskelfettanteil; MMT: Manueller Muskeltest; MRT: Magnetresonanztomografie; PODCI: Pediatric Outcomes Data Collection Instrument; UE: Unerwünschtes Ereignis.

Charakterisierung der Intervention

Tabelle 5: Charakterisierung der Intervention

Intervention	Kontrolle
Givinostat - Givinostat (Givinostat-Hydrochlorid: 8,86 mg/ml) wurde als Suspension (10 mg/ml) zur Verfügung gestellt. - Die Dosierung erfolgte Körpergewicht-abhängig und wurde peroral zweimal täglich nach dem Essen verabreicht¹⁾. - Fortführung der Behandlung mit Glukokortikoiden. <u>Überwachung von Laborparametern²⁾</u> Wöchentliche Überwachung des vollständigen Blutbilds für 8 aufeinanderfolgende Wochen bei Dosisreduktion bedingt durch: - Thrombozytenzahl $\leq 150 \times 10^9/l$ und/oder - Leukozytenzahl $< 3,0 \times 10^9/l$ und/oder - Hämoglobin $< 10,0 \text{ mg/dl}$	Placebo - Placebo wurde als Suspension (10 mg/ml) zur Verfügung gestellt; Verabreichung wie bei der Intervention.¹⁾ - Fortführung der Behandlung mit Glukokortikoiden. - Überwachung von Laborparametern entsprechend der Interventionsgruppe.
Dosierungsangaben (Startdosis und Dosismodifikationen) Angaben zur Startdosis und Dosisreduktionen in Abhängigkeit der Protokollversion (≤ 4.0 bzw. ≥ 6.0) sind in Fehler! Verweisquelle konnte nicht gefunden werden. (s. u.) zu finden. Für die Dosismodifikationen³⁾ sind folgende Hinweise, in Abhängigkeit der Protokollversion, zu beachten. <u>Dosisreduktion für Patienten, die bis inklusive Protokollversion 4.0 in die Studie aufgenommen wurden</u> 2 Dosisreduktionen möglich - Dosisreduktion um 1/3 der Startdosis war notwendig bei Thrombozytenzahl $\leq 150 \times 10^9/l$ (mind. 2 konsekutive Messungen), aber über $> 50 \times 10^9/l$ ($\neq$ Grenzwert für Abbruch der Studienmedikation). - Bei Dosisunterbrechungen aufgrund von UE konnte die Behandlung erneut angefangen werden. Eine optionale Dosisreduktion um 1/3 der Dosis, die zum Zeitpunkt des Auftretens des UE eingenommen wurde, war bei erneuter Einnahme der Studienmedikation möglich. Voraussetzung war: Thrombozytenzahl und/oder Leukozytenzahl und/oder Hämoglobin im Normbereich und/oder Triglyceride $< 300 \text{ mg/dl}$ ($3,42 \text{ mmol/l}$) oder bei Vorliegen einer milden Diarrhö. - Eine weitere Dosisreduktion um 20 % der aktuellen Dosis aufgrund von Sicherheitsbedenken war möglich. <u>Dosisreduktion unter Protokollversion 6.0, 7.0 oder 8.0</u> - Dosisreduktion war nur möglich, sofern die Stopp-Kriterien zum Abbruch der Studienmedikation nicht erfüllt waren. - Nur eine Dosisreduktion war erlaubt: Reduktion um mind. 20 % der Dosis zum Zeitpunkt der temporären Unterbrechung aufgrund von UE. - Bei Dosisunterbrechungen aufgrund von UE konnte die Behandlung erneut angefangen werden. Voraussetzung war: Thrombozytenzahl und/oder Leukozytenzahl und/oder Hämoglobin im Normbereich und/oder Triglyceride $< 300 \text{ mg/dl}$ ($3,42 \text{ mmol/l}$) bzw. Vorliegen einer milden Diarrhö. - Bei Thrombozytenzahl $\leq 150 \times 10^9/l$ (mind. 2 konsekutive Messungen), sofern die Stopp-Kriterien nicht erfüllt waren, musste eine Reduktion um mind. 20 % der Dosis erfolgen.	
Abbruch der Studienmedikation - Schwerwiegende medikamentös-bedingte Diarrhö (z. B. Anstieg auf ≥ 7 Stuhlgänge pro Tag). - Jegliches medikamentenbedingte SUE. - QTcF $> 500 \text{ ms}$.⁴⁾ - Thrombozytenzahl $\leq 50 \times 10^9/l$.⁵⁾ - Leukozytenzahl $\leq 2,0 \times 10^9/l$.⁵⁾ - Hämoglobin $\leq 8,0 \text{ g/dl}$.⁵⁾ Bei Abbruch der Studienmedikation wurde der Patient durch das ärztliche Prüfpersonal bis zum Abklingen oder einer vertretbaren Stabilisierung des UE weiter beobachtet, alle hierzu relevanten Informationen wurden dokumentiert. Erst danach fand die EOS-Visite statt.	

Intervention	Kontrolle
Unterbrechung der Studienmedikation • Moderate oder schwere Diarrhö (z. B. Anstieg auf ≥ 4 Stuhlgänge pro Tag). • Thrombozytenzahl $< 75 \times 10^9/l$, aber $> 50 \times 10^9/l$.⁶⁾ • Leukozytenzahl $< 3,0 \times 10^9/l$, aber $> 2,0 \times 10^9/l$.⁶⁾ • Hämoglobin $< 10,0$ g/dl, aber $> 8,0$ g/dl.⁶⁾ • Triglyceride > 300 mg/dl (3,42 mmol/l) in nüchternem Zustand.⁶⁾	
Nicht erlaubte Begleitmedikation • Jegliche weitere Studienmedikation abgesehen von Givinostat (Ausnahme: Deflazacort in den USA und Kanada im Zuge von Expanded/Special-Access-Programmen). • Jegliches Produkt zur Restaurierung von Dystrophin. • Jegliche pharmakologische Behandlung, die eine mögliche Wirkung auf Muskelstärke bzw. Funktion haben könnte (Ausnahme: Corticosteroide).	
Erlaubte Begleitmedikation • Vor dem Screening: Impfung gegen Varizellen aufgrund gleichzeitiger Glukokortikoid-Gabe. Bei Verabreichung von (abgeschwächten) Lebendimpfstoffen konnte das ärztliche Prüfpersonal eine Überwachung veranlassen. • Systemische Corticosteroide. • Nahrungsergänzungsmittel (Vitamin D, Kalzium etc.), sofern die Einnahme mindestens 3 Monate unmittelbar vor Studienbeginn stabil ist. • Testosteron (als Testosteronersatztherapie zur Behandlung einer verzögerten Pubertät), sofern ○ Behandlungsdosis und -schema mindestens 6 Monate vor Studienbeginn stabil sind, und ○ der Testosteronspiegel innerhalb des altersspezifischen Referenzbereichs liegt. • Arzneimittel zur Behandlung von während der Studienbehandlung auftretenden Krankheiten waren grundsätzlich erlaubt (z. B. Antiemetika, Antidiarrhoika, Antipyretika, Antihistaminika, Analgetika und Antibiotika).	
Folgende Begleittherapien erforderten während der gesamten Studie besondere Vorsicht • Arzneimittel, die das QT-Intervall verlängern. • P-Glykoprotein-Inhibitoren. • Lebendimpfstoffe. • Arzneimittel, die Substrate des organischen Kationentransporters 2 (OCT2) sind. • Thrombozytenaggregationshemmer, Thrombolytika und Antikoagulantien.	

¹⁾ Das ärztliche Studienpersonal sollte die Prüfmedikation an die Erziehungsberchtigten abgeben.

²⁾ Unabhängig einer Reduktion der Dosis war im ersten Monat eine wöchentliche und im zweiten und dritten Monat eine zweiwöchentliche Erhebung beschrieben.

³⁾ Bei einer Gewichtszunahme bleibt die Dosis unverändert; bei einer Gewichtsabnahme wird die Dosis angepasst.

⁴⁾ Durchschnittlicher QTc-Wert aus drei kurz aufeinanderfolgenden EKG-Messungen, berechnet mittels Fridericia-Formel $QTc = QT/RR^{1/3}$.

⁵⁾ Bestätigt durch Wiederholung des Tests am nächsten Werktag. Die Behandlung sollte bis zur Auswertung des zweiten Tests unterbrochen werden und nur bei einem akzeptablen Wert der zweiten Messung wieder aufgenommen werden. Lag der Wert erneut unter dem definierten Grenzwert, sollte die Studienmedikation dauerhaft abgebrochen werden.

⁶⁾ Temporäre Unterbrechung der Studienmedikation. Erneute Erhebung des Parameters sollte nach einer Woche (bei Triglyceriden nach 2 Wochen) erfolgen. Weitere wöchentliche Messungen (bei Triglyceriden alle 2 Wochen) sollten bis zur Normalisierung des Werts (bei Triglyceriden < 300 mg/dl [3,42 mmol/l]) erfolgen.

Abkürzungen: EOS: End of Study; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Tabelle 6: Startdosis und Dosismodifikationen von Givinostat und Placebo unter Protokollversion ≤ 4.0 und ≥ 6.0

Körpergewicht (kg)	≥ 10 bis $< 12,5$	$\geq 12,5$ bis < 20	≥ 20 bis < 25	≥ 25 bis < 30	≥ 30 bis < 40	≥ 40 bis < 50	≥ 50 bis < 60	≥ 60 bis < 70	≥ 70
Startdosis gemäß Protokollversion ≤ 4.0									
Startdosis (mg) BID	20	25	30	35	40	50	55	60	70
Suspension (ml) BID	2,0	2,5	3,0	3,5	4,0	5,0	5,5	6,0	7,0
Startdosis gemäß Protokollversion ≥ 6.0									
Startdosis (mg) BID	13,3	16,7	20	23,3	26,7	33,3	36,7	40	46,7
Suspension (ml) BID	1,3	1,7	2,0	2,3	2,7	3,3	3,7	4	4,7
Dosismodifikationen gemäß Protokollversion ≤ 4.0									
Reduzierte Dosis (mg) BID	13,3	16,7	20	23,3	26,7	33,3	36,7	40	46,7
Suspension (ml) BID	1,3	1,7	2,0	2,3	2,7	3,3	3,7	4	4,7
Dosismodifikationen gemäß Protokollversion ≥ 6.0									
Reduzierte Dosis (mg) BID	10,6	13,4	16	18,6	21,4	26,6	29,4	32	37,4
Suspension (ml) BID	1,1	1,3	1,6	1,9	2,1	2,7	2,9	3,2	3,7

Abkürzungen: BID: zweimal täglich.

2.3 Endpunkte

In diesem Kapitel wird die Eignung der Endpunkte hinsichtlich Operationalisierung, Patientenrelevanz und Validität beurteilt. Dazu wurden das Herstellerdossier, der zugehörige Studienbericht, das Studienprotokoll, der SAP und die in diesen Quellen zitierte Literatur herangezogen. Ergänzend wurden eigene Recherchen durchgeführt. 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 7 stellt das Ergebnis dieser Bewertung zusammenfassend dar.

Tabelle 7: Zusammenfassung der Endpunktbewertung der Studie EPIDYS

Endpunkt	Kategorie	Berücksichtigung im Dossier des pU	Berücksichtigung in der Nutzenbewertung
Todesfälle ¹⁾	Mortalität	Ja	Ja
Treppensteigen von 4 Stufen (4SC) ²⁾	Morbidität	Ja	Ergänzend ³⁾
Aufstehen aus der Rückenlage (RFTF)		Ja	Ergänzend ³⁾
Bestätigter Verlust der Aufstehfähigkeit		Ja	Ergänzend ³⁾
10-Meter-Geh-/Lauftest (10-MWT)		Ja	Ergänzend ³⁾
6-Minuten-Gehtest (6-MWT)		Ja	Ergänzend ³⁾
Verlust der Gehfähigkeit ohne Hilfsmittel		Ja	Ergänzend ³⁾
Funktionelle Leistungsfähigkeit (NSAA)		Ja	Nein
Körperliche Funktionsfähigkeit (PODCI) ⁴⁾		Ja	Nein
Muskelfettanteil		Ergänzend	Nein
Unerwünschte Ereignisse	Sicherheit	Ja	Ja
Körpergewicht		Ergänzend	Nein
Körpergröße		Ergänzend	Nein
BMI		Ergänzend	Nein

¹⁾ Todesfälle wurden über die Sicherheit erhoben.

²⁾ Primärer Endpunkt.

³⁾ Der Endpunkt wird aufgrund von Limitationen in der statistischen Auswertung ergänzend dargestellt.

⁴⁾ Der pU stellt den Endpunkt in der Kategorie „Lebensqualität“ dar.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; NSAA: North Star Ambulatory Assessment; PODCI: Pediatric Outcomes Data Collection Instrument; pU: pharmazeutischer Unternehmer; RFTF: Rise From The Floor.

2.3.1 Mortalität

Mortalität

Der Endpunkt „Mortalität“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Todesfälle wurden über die Sicherheit ab dem Zeitpunkt der Einverständniserklärung der Erziehungsberechtigten bis zur Follow-up-Visite erfasst, unabhängig von der Todesursache oder einer Korrelation mit dem Studienmedikament.

Todesfälle wurden anhand des Overall Safety Analysis Set erfasst.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz

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

Validität

Die Erhebung wird als valide angesehen.

2.3.2 Morbidität

Treppensteigen von 4 Stufen (4SC)

Der Endpunkt „Treppensteigen von 4 Stufen“ wird in der Nutzenbewertung aufgrund von Limitationen in der statistischen Auswertung (siehe Kapitel 2.4) ergänzend dargestellt.

Operationalisierung

Beschreibung

Beim „4-Stair Climb“ (4SC) handelt es sich um einen „Timed Function Test“, der die Fähigkeit 4 Treppenstufen hochzusteigen untersucht.

Der Endpunkt konnte wie folgt operationalisiert werden:

- Zeit (in Sekunden) für das Treppensteigen von 4 Stufen;
- Geschwindigkeit (4 Stufen/Gesamtzeit) zum Treppensteigen von 4 Stufen

Ein standardisiertes Vorgehen wurde in einem Manual beschrieben. Es war eine feste Reihenfolge der verschiedenen (Funktions-)Tests vorgegeben (4SC, NSAA in umgekehrter Reihenfolge (inklusive 10-MWT und RFTF), Manuelle Muskeltests für bilaterale Knie-Extension und Ellenbogen-Flexion, 6-MWT und gegebenenfalls Lungenfunktionstests zu Monat 12).

Der Ablauf, der im Folgenden beschrieben ist, galt für alle Funktionstests (u. a. 4SC, RFTF, 10-MWT, 6-MWT): Der Test wurde in der gleichen Umgebung, von in Funktionstests geschultem Personal (bspw. Physiotherapeuten) und von der gleichen Person durchgeführt und ausgewertet. Eine Videoaufzeichnung zu Qualitätszwecken erfolgte zur Screening-Visite (V1), Randomisierungsvisite (V3), Studienvisite 13 zu Monat 12 (V13) und der letzten Studienvisite / EOS-Visite.

Die Erziehungsberechtigten sollten während der Testdurchführung nicht anwesend sein.

Um die Verblindung zu gewährleisten, unterschied sich das Studienpersonal zur Überwachung der Sicherheit vom Personal zur Erhebung der Wirksamkeitsendpunkte. Das Studienpersonal musste diesen Maßnahmen zur Aufrechterhaltung der Verblindung zustimmen. Um die Verblindung zu gewähren, sollten die Testergebnisse nicht an die Erziehungsberechtigten weitergegeben werden. Das Personal für die Testdurchführung durfte keine Fragen zur Einnahme der Studienmedikation

oder zu unerwünschten Ereignissen stellen, und war nicht an der Auswertung der Sicherheitsdaten beteiligt.

Aufgrund der COVID-19-Pandemie wurde das Visitenfenster für die Visite zu Woche 72 (Monat 18) von ± 7 Tage auf ± 2 Monate erweitert. Es wird berichtet, dass es aufgrund temporärer Schließungen von Studienzentren zu Einschränkungen in der Erhebung der Endpunkte kam. Die Testdurchführung sollte vor Ort im Studienzentrum stattfinden.

Testdurchführung

Der Test erfolgt an einer 4-stufigen Treppe (15 cm / Stufe) mit einer Plattform am Treppende. An den Seiten der Treppe befanden sich zwei verstellbare Geländer.

Der Proband soll sich mittig vor die Treppe stellen (≤ 20 cm Abstand), mit beiden Armen an der Seite (Treppengeländer darf nicht gehalten werden). Sobald die Aufforderung zum Start gegeben wird, soll die Stoppuhr gestartet werden. Die Zeit wird gestoppt (auf die Zehntelsekunde genau), wenn der Proband aufrecht auf der Plattform steht, mit beiden Armen an der Seite. Der Patient wurde dazu aufgefordert, so schnell wie möglich 4 Treppenstufen zu steigen mit den Armen an den Seiten, möglichst ohne das Treppengeländer anzufassen; die Nutzung der Handläufe war erlaubt.

Beurteilung der Testdurchführung

Neben der benötigten Zeit in Sekunden und der Geschwindigkeit wurde der Test mittels folgender Skala durch die Bewerterin / den Bewerter eingeordnet:

1. Nicht in der Lage, vier Treppenstufen zu steigen (Zeit wurde nicht erfasst).
2. Treppenstufen werden im Nachstellschritt gestiegen (beide Füße werden auf eine Stufe gestellt, bevor zur nächsten Stufe übergegangen wird), **beide Hände** befinden sich am Geländer (ein- oder beidseitig) bzw. eine Hand ist am Geländer, während die andere sich am Bein abstützt.
3. Treppenstufen werden im Nachstellschritt gestiegen, **eine Hand** befindet sich am Geländer oder stützt sich am Bein ab.
4. Treppenstufen werden im Nachstellschritt gestiegen, **benötigt kein** Geländer bzw. kein Abstützen am Bein.
5. Treppenstufen werden im Wechselschritt gestiegen, Geländer bzw. Abstützen am Bein **werden benötigt**.
6. Treppenstufen werden im Wechselschritt gestiegen, Geländer bzw. Abstützen am Bein wird **nicht** benötigt.

Zudem sollte durch das Prüfpersonal eingeschätzt werden, inwiefern der Patient sich optimal oder suboptimal angestrengt hat. Der Test konnte bis zu 2 Mal wiederholt werden und nur das schnellste Testergebnis wurde aufgezeichnet und die Beurteilung der Testdurchführung dokumentiert. Falls der Proband nicht in der Lage ist, den Test in < 30 s zu absolvieren, kann der Test abgebrochen werden. Eine Wiederholung war nicht notwendig.

Es wurden mit dem Dossier Responderanalysen für die Dauer bzw. benötigte Zeit und Geschwindigkeit sowie Analysen für die kontinuierlichen Daten eingereicht. Eine deskriptive Auswertung war für die Gradeinteilung der Bewertungsskala präspezifiziert.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar. Aufgrund der COVID-19-Pandemie waren Studienzentren teilweise geschlossen, weshalb eine Erhebung der Endpunkte zu den im Studienprotokoll definierten Zeitpunkten für einen Teil der Studienpopulation gefährdet war. Daher wurde das Visitenfenster für die Visite zu Woche 72 (Monat 18) von ± 7 Tage auf ± 2 Monate erweitert. Daraus ergeben sich Unsicherheiten, inwiefern der tatsächliche Zeitpunkt der Erhebung oder der präspezifizierte Zeitpunkt (EOS) in die Analyse einging (siehe Kapitel 2.4).

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevant angesehen.

Validität

Die Validität des Endpunkts „Treppensteigen von 4 Stufen“ (mittels 4SC) wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als valide eingeschätzt [6,7,8,10].

Die inhaltliche Interpretation einer Geschwindigkeit zur Beurteilung einer Verbesserung/Verschlechterung des Funktionszustands stellt sich schwierig dar. Im vorliegenden Anwendungsgebiet erscheint relevant, ob ein Patient in der Lage ist selbstständig und ohne Hilfsmittel den Test zu absolvieren. Von klinischer Relevanz zur Beurteilung der Veränderung der Funktion ist im vorliegenden Anwendungsgebiet eher die Dauer bzw. benötigte Zeit (in Sekunden), die der Betroffene benötigt aufzustehen, als die Geschwindigkeit [9].

Zur Herleitung einer Responseschwelle wird vom pU ein verteilungsbasierter Ansatz (1/3 Standardabweichung (SD) zu Baseline) herangezogen. Die Bestimmung einer klinischen Relevanzschwelle bei zeitabhängigen Funktionstests kann aufgrund erhöhter Variabilität in den Baseline-Werten und der geringen Veränderung gegenüber Baseline, die gemessen wird, schwierig sein [5]. Zudem erscheinen gemessene Änderungen in der Funktionalität, die über den Endpunkt gemessen werden, im vorliegenden Anwendungsgebiet altersbedingt [2], was bei der Bestimmung einer klinischen Relevanzschwelle zu berücksichtigen ist.

Daher werden die Ergebnisse der kontinuierlichen Daten des Endpunkts „Treppensteigen von 4 Stufen“ in der Operationalisierung der benötigten Zeit (in Sekunden) zum Durchführen des Tests gegenüber den Responderanalysen bevorzugt.

Die Ergebnisse der benötigten Zeit (in Sekunden) der kontinuierlichen Daten der Overall-Population werden ergänzend dargestellt (Angaben und Limitationen in der Auswertung siehe Kapitel 2.4).

Aufstehen aus der Rückenlage (RFTF)

Der Endpunkt „Aufstehen aus der Rückenlage“ wird in der Nutzenbewertung aufgrund von Limitationen in der statistischen Auswertung (siehe Kapitel 2.4) ergänzend dargestellt.

Operationalisierung

Beschreibung

Der Test „Rise From The Floor“ (RFTF) ist ein „Timed Function Test“, der die Zeit (in Sekunden) erhebt, die ein Patient mit DMD benötigt, um aus der Rückenlage am Boden bis in den aufrechten Stand zu gelangen. Der Test wurde im Rahmen des „North Star Ambulatory Assessment“ (NSAA) durchgeführt (Item 12).

Es liegen zwei Operationalisierungen vor:

- Zeit (in Sekunden), die benötigt wird, um aus der Rückenlage am Boden in den aufrechten Stand zu gelangen.
- Geschwindigkeit (Erhebung/Sekunde), die benötigt wird, um aus der Rückenlage in den aufrechten Stand zu gelangen.

Ein standardisiertes Vorgehen wurde in einem Manual beschrieben. Es wurde eine feste Reihenfolge der Tests vorgegeben sowie Maßnahmen zum standardisierten Vorgehen, zur Qualitätssicherung und anlässlich der COVID-19-Pandemie (siehe Operationalisierung zum Endpunkt „Treppensteigen von 4 Stufen“).

Im Rahmen der Überprüfung der Inter-Rater-Reliabilität, u. a. für die Erhebungsinstrumente 6-MWT und NSAA, sollten stets zwei Personen, die in der Durchführung und Bewertung der Tests geschult sind, anwesend sein. Die Bewertung der Testdurchläufe sollte in Echtzeit erfolgen und nicht im Nachgang durch die Videoaufnahmen bewertet werden.

Testdurchführung

Für die Startposition lag der Patient auf dem Rücken, mit den Armen an seinen Seiten und den Beinen zusammen auf dem Boden. Der Patient wurde nun gebeten, nach dem Startsignal so schnell wie möglich in einen aufrechten Stand mit den Armen an der Seite zu gelangen. Dabei wurde die Zeit zwischen dem Startsignal und dem aufrechten Stand mit den Armen an den Seiten gemessen. Sollte der Patient nach 30 Sekunden nicht erfolgreich aufgestanden sein, wurde ein Stuhl angeboten, mit dem Hinweis, die Sitzfläche als Hilfe zu verwenden. Die Benutzung eines Kissens war nicht erlaubt.

Beurteilung der Testdurchführung

Neben der Dauer und der Geschwindigkeit in der Durchführung des RFTF, wurde die Durchführung mit folgender Skala durch die Bewerterin / den Bewerter beurteilt:

- Grad 1: Nicht in der Lage, aus der Rückenlage aufzustehen, auch nicht mit Hilfe eines Stuhls. Es wurde keine Zeit aufgezeichnet.
- Grad 2: Gowers-Zeichen mit Hilfsmittel – Möbel oder externe Unterstützung werden benötigt, um aus der Rückenlage in eine stehende Position zu gelangen. Es wurde keine Zeit aufgezeichnet.
- Grad 3: Vollständig ausgeprägtes Gowers-Zeichen – Dreht sich auf den Bauch und "klettert" mit den Armen an den Beinen hoch.
- Grad 4: Partiiell ausgeprägtes Gowers-Zeichen – Dreht sich auf den Bauch und stützt sich beim Aufrichten mit einer Hand auf dem Bein ab.
- Grad 5: Dreht sich auf die Seite und stützt sich beim Aufrichten mit einer oder beiden Händen auf dem Boden ab, berührt aber nicht die Beine.
- Grad 6: Steht auf, ohne sich zu drehen oder sich beim Aufrichten mit den Händen auf den Beinen abzustützen.

Da der Endpunkt während der Erhebung des NSAA erfasst wurde, wurde die Durchführung zusätzlich anhand der NSAA-Skala bewertet:

- 2 Punkte: Kein Gowers-Zeichen.
- 1 Punkt: Mindestens eine Komponente des Gowers-Zeichen liegt vor (Umdrehen auf den Bauch, Vierfüßlerstand, Abstützen mit den Händen auf Boden/Beinen, breiter Stand mit gestreckten Knien).
- 0 Punkte
 - 0a: Benötigt Hilfsmittel (z. B. Stuhl, Wand).
 - 0b: Nicht in der Lage den Test durchzuführen.

Gemäß Manual entspricht „0b“ auf der NSAA-Skala einem RFTF-Grad 1 und „0a“ auf der NSAA-Skala einem RFTF-Grad 2.

Eine Wiederholung des Tests war bei unzureichender Compliance des Patienten, beurteilt durch die Bewerterin / den Bewerter, und zur Überprüfung der dokumentierten Zeit und der Beurteilung der Gradeinteilung möglich. Es sollten der bestmögliche Grad und die Bestzeit dokumentiert werden, selbst wenn die Bestzeit und der bestmögliche Grad zu unterschiedlichen Testdurchläufen während der Visite erzielt wurden.

Es wurden Analysen zu Dauer und Geschwindigkeit zu den kontinuierlichen Daten eingereicht. Neben der Auswertung der kontinuierlichen Daten wurden post hoc durchgeführte Responderanalysen für die Zeit (Sekunden) mit einer Responseschwelle von 1/3 SD der Baseline-Werte (verteilungsbasierter Ansatz) und für die Geschwindigkeit eine Responseschwelle von 0,023 Erhebungen/Sekunde eingereicht.

Bewertung

Die Operationalisierung wird als weitestgehend geeignet angesehen. Die Durchführung des RFTF erfolgte im Rahmen der Erhebung des NSAA-Item 12. Im Manual wird angegeben, dass die Bewertung der Testdurchführung zwischen der RFTF-Skala und der Skala des NSAA-Items 12 konsistent sein soll. Angaben, wie mit diskrepanten Ergebnissen umgegangen wird, konnten nicht identifiziert werden.

Patientenrelevanz

Der Endpunkt wird in der vorliegenden Operationalisierung als patientenrelevant angesehen.

Validität

Die Validität des Endpunkts wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als valide eingeschätzt [6,7,8,10].

Responderanalysen mit einer anhand eines verteilungsbasierten Ansatzes abgeleiteten Response-schwelle werden als nicht geeignet angesehen (siehe Ausführungen zur Bewertung der Validität zum Endpunkt „Treppensteigen von 4 Stufen“).

Die Ergebnisse zur benötigten Zeit (in Sekunden) der Overall-Population werden ergänzend dargestellt (Angaben und Limitationen in der Auswertung siehe Kapitel 2.4).

Bestätigter Verlust der Aufstehfähigkeit

Der Endpunkt „Bestätigter Verlust der Aufstehfähigkeit“ wird in der Nutzenbewertung aufgrund von Limitationen in der statistischen Auswertung (siehe Kapitel 2.4) ergänzend dargestellt.

Operationalisierung

Beschreibung

Der Beginn eines Verlusts der Aufstehfähigkeit war definiert als die erste Visite, bei der dies beobachtet wurde. Ein bestätigter Verlust der Aufstehfähigkeit lag vor, wenn der Patient bei mindestens 2 aufeinanderfolgenden Visiten (frühestens nach 6 Monaten) nicht in der Lage war, aufzustehen. Für den Verlust der Aufstehfähigkeit mussten bei den Visiten folgende Scores vorliegen:

- RFTF-Grad 1 **und**
- NSAA-Score (Item 12) „0“ (bzw. 0b = Nicht in der Lage den Test aufgrund körperlicher Einschränkungen durchzuführen).

Gemäß Manual sollten die Angaben des RFTF mit NSAA-Item 12 konsistent sein. Informationen zum Vorgehen bei diskrepanten Angaben zwischen RFTF und NSAA konnten nicht identifiziert werden.

Wenn eine Person die Kriterien bis Studienende (Visite Woche 72 / EOS) nicht erfüllte, wurde zum Zeitpunkt der letzten Erhebung (mit imputierten Daten) zensiert, d. h. zur Visite in Woche 72.

Fehlende Werte wurden in Klasse 1 und Klasse 2 differenziert. Fehlende Werte der Klasse 2 wurden zur Identifikation von Patienten verwendet, die den Test aufgrund körperlicher Einschränkungen nicht durchführen konnten (siehe Kapitel 2.4).

Mit Modul 4 wurden Überlebenszeitanalysen für den Endpunkt „Bestätigter Verlust der Aufstehfähigkeit“ der Overall-Population eingereicht.

Bewertung

Die Operationalisierung wird als überwiegend geeignet angesehen. Zu den Limitationen der Erhebungsinstrumente NSAA und RFTF, siehe zugehörigen Abschnitt zu den Endpunkten.

Gemäß Manual sollten die Angaben des RFTF mit NSAA-Item 12 konsistent sein. Es ist unklar, wie mit diskrepanten Ergebnissen umgegangen wurde.

Patientenrelevanz

Der Verlust der Aufstehfähigkeit wird als patientenrelevant angesehen.

Validität

Zur Bewertung der Validität der Erhebungsinstrumente RFTF und NSAA siehe entsprechende Abschnitte zu den Endpunkten.

Der pU definiert den Endpunkt als persistierenden Verlust der Stehfähigkeit, bei dem der Patient an zwei aufeinanderfolgenden Visiten nicht in der Lage war aufzustehen. Für die vom pU gewählte Operationalisierung sollte bei einer Studiendauer von 18 Monaten der Verlust der Stehfähigkeit an mehr als 2 aufeinanderfolgenden Visiten zu messen sein. In der vorliegenden Operationalisierung handelt es sich daher, anders als vom pU als persistierender/ dauerhafter Verlust definiert, um die Erfassung des bestätigten Verlusts der Aufstehfähigkeit.

Unklar ist, inwiefern durch einen Wachstumsschub der pädiatrischen Patienten und/oder durch therapeutische Maßnahmen ein temporäreres besseres Testergebnis erzielt werden kann. Dieser Einfluss wird als gering angesehen, da durch die vorliegende Operationalisierung ein progredienter Krankheitsverlauf widerspiegelt wird.

Gemäß Manual sollten die Angaben des RFTF zu Dauer und Gradeinteilung mit denen des NSAA konsistent sein. Es ist unklar, inwiefern es sich bei der potentiellen Doppelerfassung um eine qualitätssichernde Maßnahme handelt.

Der Endpunkt „Funktionelle Leistungsfähigkeit“ wird aufgrund von Limitationen in der Validität des NSAA nicht in der Nutzenbewertung dargestellt. Die Durchführung des RFTF erfolgte im Rahmen der Erhebung des NSAA-Item 12. Im Manual wird angegeben, dass die Bewertung der Testdurchführung zwischen der RFTF-Skala und der Skala des NSAA-Item 12 konsistent sein soll. Angaben, wie mit diskrepanten Ergebnissen umgegangen wird, konnten nicht identifiziert werden. Daher wird die Erfassung des Endpunkts „Bestätigter Verlust der Aufstehfähigkeit“ weitestgehend als valide angesehen.

Der Endpunkt wird in der Nutzenbewertung aufgrund von Limitationen in der statistischen Auswertung (siehe Kapitel 2.4) ergänzend dargestellt.

Gehfähigkeit

10-Meter-Geh-/Lauftest (10-MWT)

Der Endpunkt „10-Meter-Geh-/Lauftest“ wird in der Nutzenbewertung aufgrund von Limitationen in der statistischen Auswertung (siehe Kapitel 2.4) ergänzend dargestellt.

Operationalisierung

Beschreibung

Der „10-Meter Walk/Run Test“ (10-MWT) ist ein „Timed Function Test“, der die Zeit erhebt, die ein Patient benötigt, um 10 Meter zu gehen oder zu laufen. Der Test wurde im Rahmen des NSAA durchgeführt (Item 17).

Es liegen zwei Operationalisierungen vor:

- Benötigte Zeit für eine Geh-/Laufstrecke von 10 Metern in Sekunden
- Geh-/Laufgeschwindigkeit über 10 Meter in Meter/Sekunde

Ein standardisiertes Vorgehen wurde in einem Manual beschrieben. Es wurde eine feste Reihenfolge der Tests vorgegeben (siehe Operationalisierung des 4SC) und es sollten Untersuchungen zur Qualität/Reliabilität vorgenommen werden (siehe Abschnitt zu RFTF).

Bewertung der Testdurchführung

Neben der Dauer und der Geschwindigkeit in der Durchführung des 10-MWT wurde der Test mit folgender Skala durch die Bewerterin / den Bewerter beurteilt:

- Grad 1: Nicht in der Lage selbständig zu laufen (keine Erfassung der Zeit).
- Grad 2: Nicht in der Lage selbständig zu gehen, aber mit Hilfe von Knie-Knöchel-Fuß-Orthese oder mit Unterstützung möglich (keine Erfassung der Zeit).
- Grad 3: Zeigt hochgradig angepassten, breitbeinigen und lordotischen Gang; Gehgeschwindigkeit kann nicht erhöht werden.
- Grad 4: Mäßig angepasster Gang; kann Geschwindigkeit aufnehmen, aber nicht rennen.
- Grad 5: Kann Gehgeschwindigkeit erhöhen, läuft jedoch mit Doppelstützphase, d. h. Laufen ohne Flugphase.
- Grad 6: Rennen mit Flugphase.

Eine erneute Durchführung des Tests war möglich, um die Ergebnisse zu überprüfen. Es sollten die schnellste Durchführung und der höchste Beurteilungsgrad dokumentiert werden, auch wenn beides an unterschiedlichen Testwiederholungen während der Visite erzielt wurde.

Neben der Auswertung der kontinuierlichen Daten wurden post hoc durchgeführte Responderanalysen für die Zeit (Sekunden) mit einer Responseschwelle von 1/3 SD der Baseline-Werte und für die Geschwindigkeit von 0,212 Meter/Sekunde eingereicht.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar (siehe Ausführungen zu 4SC und RFTF).

Patientenrelevanz

Der Endpunkt „10-Meter-Geh-/Lauftest“ wird im vorliegenden Anwendungsgebiet als patientenrelevant angesehen.

Validität

Die Validität des Endpunkts wurde in vorherigen Nutzenbewertungsverfahren bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als valide eingeschätzt [6,7,8,10].

Responderanalysen mit einer anhand verteilungsbasierten Ansatzes abgeleiteten Response-schwelle werden nicht als geeignet angesehen (siehe Ausführungen zur Bewertung der Validität der Endpunkte „Treppensteigen von 4 Stufen“ bzw. „Aufstehen aus der Rückenlage“).

Die Ergebnisse zur benötigten Zeit (in Sekunden) der kontinuierlichen Daten der Overall-Population werden aufgrund von Limitationen in der statistischen Auswertung (siehe Kapitel 2.4) ergänzend dargestellt.

6-Minuten-Gehtest (6-MWT)

Der Endpunkt „6-Minuten-Gehtest“ wird in der Nutzenbewertung aufgrund von Limitationen in der statistischen Auswertung (siehe Kapitel 2.4) ergänzend dargestellt.

Operationalisierung

Beschreibung

Der „6-Minute Walk Test“ (6-MWT) dient der Untersuchung der körperlichen Funktion und wird durch das Studienpersonal erhoben. Er erfasst die Distanz, die eine Person innerhalb von 6 Minuten gehen kann.

Der Test wurde gemäß McDonald et al. (2013) [24] durchgeführt. Bei der Publikation handelt es sich um eine für pädiatrische Patienten mit DMD adaptierte Version der standardisierten Handlungsanweisung der American Thoracic Society aus dem Jahr 2002.

Für den 6-MWT sollen die Studienteilnehmer Schuhe tragen, wobei für jeden Test die gleiche Art von Schuhen getragen werden soll. Erlaubt sind Turn- oder Tennisschuhe; Flipflops, „Crocs“ oder Gummistiefel sind nicht erlaubt. Orthesen durften während des Tests nicht getragen werden, mit Ausnahme von Einlegesolen. Wenn eine Einlegesole beim ersten Test verwendet wurde, musste sie auch bei jedem weiteren Test verwendet werden.

Die Patienten sollen in einem Rollstuhl zu der Teststrecke gebracht werden, damit sie vor dem Testbeginn ausgeruht sind. Eine Erholung von 10 Minuten war vor Testbeginn vorgesehen. Eine Durchführung des Tests war nur bei Freigabe durch das ärztliche Personal möglich.

Wenn der Patient nach einem Sturz verletzt war oder nicht mehr in der Lage aufzustehen, wurden die Gesamtzeit und die Gesamtstrecke dokumentiert und der Test abgebrochen. Bei vorzeitigem Abbruch des Tests und sofern keine Angaben zum Testergebnis vorliegen, wurde der Test als fehlend dokumentiert (siehe Kapitel 2.4).

Neben der Auswertung der kontinuierlichen Daten wurden post hoc durchgeführte Responderanalysen mit einer post hoc definierten Responseschwelle von 30 Metern definiert.

Bewertung

Die Operationalisierung ist weitestgehend nachvollziehbar (siehe Ausführungen zu 4SC und RFTF).

Patientenrelevanz

Der Endpunkt „6-Minuten-Gehtest“ wird im vorliegenden Anwendungsgebiet als patientenrelevant angesehen.

Validität

Der 6-MWT wurde zur Messung der funktionalen körperlichen Fähigkeiten entwickelt. Es handelt sich um ein standardisiertes und etabliertes Testverfahren, das zu Diagnostik und Verlaufskontrolle bei einer Vielzahl von Indikationen, u. a. auch bei DMD, zum Einsatz kommt. Für die Studie wurde eine adaptierte Version des 6-MWT verwendet, die anhand Patienten mit DMD (N = 21) und gesunden Patienten (N = 34) zwischen 4 und 12 Jahren validiert wurde [24,25]. Der Test wurde ursprünglich für erwachsene Personen entwickelt. Die Richtlinie der European Medicines Agency (EMA) zu Endpunkten in klinischen Studien bei DMD erwähnt den 6-MWT bei gehfähigen Patienten, jedoch unter Vorbehalt aufgrund u. a. folgender Limitationen: inter- und intrapersonelle Variabilität, Einfluss des Alters zu Baseline, Einfluss eines Wachstumseffekts. [5]

Post hoc legt der pU Responderanalysen mit einer Responseschwelle von 30 Metern vor, die mittels verteilungsbasierter Verfahren berechnet wurde. Die klinische Relevanzschwelle wird als nicht geeignet angesehen. Es werden ankerbasierte Verfahren zur Ableitung einer „Minimal Importance Difference“ (MID) bevorzugt. Hinzu kommen Limitationen bezüglich altersabhängiger Veränderungen in der Funktionalität der Betroffenen, die zu berücksichtigen sind.

Die Analysen zu kontinuierlichen Daten der Overall-Population werden ergänzend dargestellt (Limitationen zur statistischen Auswertung siehe Kapitel 2.4).

Verlust der Gehfähigkeit ohne Hilfsmittel

Der Endpunkt wird in der Nutzenbewertung aufgrund von Limitationen in der statistischen Auswertung (siehe Kapitel 2.4) ergänzend berücksichtigt.

Operationalisierung

Beschreibung

Ein Verlust der Gehfähigkeit ohne Hilfsmittel lag vor, wenn zu einer Visite zu einem beliebigen Studienzeitpunkt vor oder zu Studienende (letzte Studienvisite zu Monat 18 / Woche 72 bzw. EOS-Visite) folgende Kriterien erfüllt waren:

- Der Patient war aufgrund körperlicher Einschränkungen nicht in der Lage, den 6-MWT durchzuführen **und**
- der Patient war nicht in der Lage, den 10-MWT in ≤ 30 Sekunden ohne Unterstützung oder Hilfsmittel zu absolvieren ($10\text{-MWT} \leq \text{Grad } 2$).

Die Kriterien mussten bei allen darauffolgenden Visiten bis Studienende zu Monat 18 oder EOS vorliegen.

Fehlende Werte des 6-MWT und 10-MWT sollten identifiziert und nach den Vorgaben zum Umgang mit fehlenden Werten imputiert (siehe Kapitel 2.4) werden.

Fehlende Werte wurden in Klasse 1 und 2 differenziert. Fehlende Werte der Klasse 1 waren definiert als fehlende Werte aus anderen Gründen oder durch COVID-19 bedingt. Fehlende Werte der Klasse 2 wurden zur Identifikation von Patienten verwendet, die den Test aufgrund von Gehunfähigkeit / körperlicher Einschränkungen nicht durchführen konnten (siehe Kapitel 2.4).

Post hoc wurden mit dem Dossier Responderanalysen für den Anteil der Patienten mit Verlust der Gehfähigkeit ohne Hilfsmittel eingereicht.

Bewertung

Zur Bewertung der Endpunkte „10-Meter-Geh-/Lauftest“ und „6-Minuten-Gehtest“ siehe entsprechende Abschnitte weiter oben in diesem Kapitel.

Die Operationalisierung wird als bedingt geeignet angesehen. Bei Vorliegen eines Verlusts der Gehfähigkeit zur letzten Studienvisite war keine erneute Bestätigung notwendig.

Angaben, weshalb in der vorliegenden Operationalisierung anstelle eines 10-MWT-Score 1, stattdessen ein 10-MWT-Score von ≤ 2 und eine Mindestzeit vorgegeben wurde, ist unklar.

Patientenrelevanz

Der Verlust der Gehfähigkeit wird als patientenrelevant angesehen.

Validität

Bei dem Endpunkt „Verlust der Gehfähigkeit ohne Hilfsmittel“ handelt es sich um einen kombinierten Endpunkt aus den Endpunkten „10-Meter-Geh-/Lauftest“ und „6-Minuten-Gehtest“. Zur Bewertung der beiden Endpunkte wird auf die entsprechenden Abschnitte in diesem Kapitel verwiesen.

Gemäß den Studienunterlagen lag ein Verlust der Gehfähigkeit vor, wenn der Patient aufgrund körperlicher Einschränkungen nicht in der Lage war, den 6-MWT durchzuführen und der Patient nicht in der Lage war den 10-MWT in ≤ 30 Sekunden zu absolvieren oder dies nicht ohne Unterstützung oder Hilfsmittel möglich war ($\triangleq 10\text{-MWT-Grad } 2$). Eine Begründung für die vorliegende Operationalisierung liegt nicht vor; zudem ist unklar, weshalb für den 10-MWT ein zeitliches Limit für die benötigte Zeit zum Absolvieren des Tests und bspw. für einen Gehverlust $10\text{-MWT} \leq \text{Grad } 2$ gewählt wurde.

Ein dauerhafter Verlust der Gehfähigkeit ohne Hilfsmittel lag vor, wenn zu einer Visite zu einem beliebigen Studienzeitpunkt vor oder zu Studienende (Woche 72 / EOS) die präspezifizierten

Kriterien als erfüllt galten. Ein Verlust der Gehfähigkeit, der durch konsekutive Messungen bestätigt wird, sollte einem dauerhaften Verlust der Gehfähigkeit vorangehen. In der vorliegenden Operationalisierung liegt ein dauerhafter Verlust der Gehfähigkeit bereits bei einer einmaligen Erhebung zum Studienende entweder zur letzten Studienvisite (Monat 18) oder bei der Visite bei vorzeitigem Abbruch der Studie vor. Für diese Patienten liegt somit kein bestätigter Verlust der Gehfähigkeit vor.

Der Endpunkt wird in der Nutzenbewertung aufgrund von Limitationen in der statistischen Auswertung ergänzend berücksichtigt.

Funktionelle Leistungsfähigkeit (NSAA)

Der Endpunkt „Funktionelle Leistungsfähigkeit“ wird in der Nutzenbewertung aufgrund nicht-gegebener Validität nicht berücksichtigt.

Operationalisierung

Beschreibung

Die funktionelle Leistungsfähigkeit wurde anhand der klinischen Bewertungsskala „North Star Ambulatory Assessment“ (NSAA) durch eine in der Erhebung des Endpunkts geschulten Person erhoben, die die Gehfähigkeit männlicher Patienten mit DMD beurteilt. Die Bewertungsskala umfasst 17 Items, u. a. Aufgaben zum ein- und zweibeinigen Stehen (für jeweils 3 Sekunden), Gehen, Aufstehen aus verschiedenen Positionen, Klettern auf eine Kiste, Kopfanheben in Rückenlage, Hüpfen, 10 Meter laufen. Die Erhebungsinstrumente RFTF (Item 12) und 10-MWT (Item 17) sind Bestandteil des NSAA.

Gemäß Studienunterlagen wurde die Reihenfolge der Items innerhalb des NSAA geändert, um eine effizientere Durchführung des Tests zu gewährleisten. Beim Test „Auf- und Absteigen von Stufen“ wurde folgende Anweisung gegeben: den rechten Fuß eine Stufe hoch, den rechten Fuß eine Stufe wieder runter, den linken Fuß eine Stufe hoch, den linken Fuß wieder runter. Bei den Tests „Sich setzen“ und „Aufstehen vom Boden“ konnte der Kopf gehoben werden, um Ermüdungserscheinungen zu vermeiden.

Die Durchführung der 17 Items wurde durch die Bewerterin / den Bewerter anhand einer 3-Punkte-Skala (0–2 Punkte) bewertet. Der Gesamtscore bildete sich aus der Summe der erzielten Scores der 17 Items und konnte einen Wert zwischen 0 und 34 Punkten annehmen, wobei höhere Werte eine bessere funktionelle Leistungsfähigkeit anzeigen.

Die Bewertung der Testdurchführung erfolgte durch die Bewerterin / den Bewerter für jedes Item wie folgt:

- 2 Punkte: Patient führt den Test normal aus, ohne offensichtliche Einschränkungen in der Aktivität.
- 1 Punkt: Eingeschränkte Durchführung, aber der Patient erreicht das Ziel ohne physische/fremde Hilfe.
- 0 Punkte: Patient kann die Aufgabe nicht selbstständig ausführen.

Es liegen post hoc durchgeführte Analysen zu den kontinuierlichen Daten der Overall-Population vor sowie Responderanalysen für eine Verbesserung/Verschlechterung mit einer Response-schwelle von mindestens 15 % der Skalenspannweite (5,1 Punkte) für die Overall-Population.

Bewertung

Die Operationalisierung ist nachvollziehbar. Der NSAA umfasst mit 2 Items den RFTF und den 10-MWT. Diese Funktionen sind somit im Sinne einer Doppelerfassung sowohl über den NSAA als auch über die beiden Endpunkte dargestellt.

Patientenrelevanz

Einschränkungen der funktionalen Leistungsfähigkeit, insbesondere des Gehvermögens, werden als patientenrelevant angesehen.

Validität

Eckdaten zu psychometrischen Eigenschaften wurden in verschiedenen Studien zur Reliabilität (Scott et al. (2012) [28], Mazzone et al. (2009) [23]) bzw. zur Änderungssensitivität (Mayhew et al. (2011) [22], Muntoni et al. (2019) [26]) ermittelt.

In der Studie von Scott et al. (2012) [28] konnte eine hohe Test-Retest-Reliabilität mit 100 % und ein hoher Intraclass-Korrelations-Koeffizient (ICC) für den Gesamtscore von 0,95 ermittelt werden, für die einzelnen Items wurde aber eine große Variabilität beobachtet (Kohorte 1 (6 Probanden; 5 Untersucher/innen): ICC = 0,0 für das Item „Vom Boden Aufstehen“; ICC = 1,0 für das Item „Eine Stufe hochsteigen (rechts)“). Zudem war die Anzahl an Probanden (n = 5 bzw. n = 3) gering. Der ICC für die Items „Hüpfen“ und „10 Meter laufen“ betrug 0,8 bzw. 0,6. Eine Änderungssensitivität wurde anhand einer transformierten/linearisierten Skala (0–100 Punkte) [22] bzw. der kumulative Anteil an Personen mit einem Funktionsverlust von 2 oder 1 Punkten auf 0 Punkte anhand Longitudinaldaten zum natürlichen Krankheitsverlauf untersucht [26]. In der Studie EPIDYS wurde jedoch die Ordinalskala (0–34 Punkte) verwendet und die mittlere Änderung zum Studienende gegenüber Baseline analysiert. Es wurden keine Daten zur Beurteilung der Änderungssensitivität anhand der Ordinalskala eingereicht.

Aufgrund der Doppelerfassung der vom NSAA umfassten Items 10-MWT und RFTF und der limitierten Angaben zu Reliabilität, Validität und Änderungssensitivität wird der Endpunkt in der Nutzenbewertung nicht dargestellt.

Körperliche Funktionsfähigkeit (PODCI)

Der Endpunkt „Körperliche Funktionsfähigkeit“ wird in der Nutzenbewertung aufgrund unzureichender Operationalisierung und unklarer Validität nicht berücksichtigt.

Operationalisierung

Beschreibung

Der Endpunkt wurde mit dem „Pediatric Outcomes Data Collection Instrument“ (PODCI) der „Pediatric Orthopedic Society of North America“ (POSNA) erhoben, einem Instrument zur Bewertung der allgemeinen Gesundheit, der Schmerzen und der Fähigkeit, an Alltagsaktivitäten sowie lebhaften Aktivitäten teilzunehmen. Es wurde für Kinder und Jugendliche im Alter von 2 bis 18 Jahren mit Knochen- und Muskelerkrankungen entwickelt. Bei Kindern ≤ 10 Jahren füllten die Eltern die Elternversion des Fragebogens aus, ältere Kinder füllten den Fragebogen selbst aus.

Der Fragebogen besteht aus 6 Subskalen mit 86 Items. Für die vorliegende Studie werden 5 Subskalen verwendet, die anhand einer Likert-Skala von 1 bis 6 bzw. 1 bis 5 bewertet werden, wobei höhere Werte einen höheren Grad der Beeinträchtigungen darstellen. Die Subskala „Therapieerwartungen“ wurde nicht angewandt. Der Bezugszeitraum variierte und es werden drei Möglichkeiten angegeben: „Während der letzten Woche“, „Während der letzten 12 Monate“, „In diesem Moment“.

Die in der Studie verwendeten Subskalen sind:

- Obere Extremitäten (8 Items)
- Transfer/Basismobilität (11 Items)
- Sport / Physikalische Funktion (13 Items)
- Schmerz/Wohlbefinden (3 Items)
- Zufriedenheit (5 Items)

Die Subskalen werden in einen standardisierten Score umgewandelt. Es können bis zu 100 Punkte erreicht werden. Der standardisierte globale Funktionsscore wurde als standardisierter Score der 4 Subskalen „Obere Extremitäten“, „Transfer/Basismobilität“, „Sport / Physikalische Funktion“, „Schmerz/Wohlbefinden“ berechnet:

$$\text{Globaler Funktionsscore} = \text{Summe der vier standardisierten Subskalen} / 4$$

Eine Berechnung des globalen Funktionsscores konnte nur erfolgen, wenn für alle 4 Subskalen Werte vorlagen.

Gemäß SAP war eine separate Auswertung der standardisierten Scores der Subskalen und des globalen Funktionsscores für die Eltern- und Patientenversion vorgesehen. Im SAP V6.1. Addendum (14.07.2022), nach Datenbanksperrung (22.02.2022) aber vor Entblindung, wurde das Vorgehen angepasst, da auffiel, dass Patienten nicht konsistent dem Alter entsprechend eine Fragebogenversion erhielten. Stattdessen haben Patienten bzw. deren Erziehungsberechtigten, unabhängig des Alters, je nach Visite entweder die Version für Patienten oder Erziehungsberechtigte ausgefüllt. Aus diesem Grund wurden anstelle der präspezifizierten Analyse, getrennt nach der Version für den Patienten bzw. Erziehungsberechtigten, post hoc beide Versionen zusammen ausgewertet. Für Patienten ≤ 10 Jahre, die die selbstberichtete Version > 10 Jahre, erhielten, wurden die Daten als fehlende Werte der Klasse 1 (= andere Gründe als körperliche Einschränkungen bzw. Gehunfähigkeit; COVID-19) definiert und imputiert.

Bei der Version für Erziehungsberechtigte wurden Items mit einem Score von 5 Punkten (= zu jung für diese Aktivität) als fehlend eingeordnet und nicht zur Berechnung der Subskala miteinbezogen. Mit SAP V6.1 Addendum (14.07.2022) fiel auf, dass bei einzelnen Items hierfür ein Score von „6“ vergeben wurde und bei der Subskala „Zufriedenheit“ die Bezeichnung variierte („Kind ist zu jung“). Um ein einheitliches Vorgehen zu ermöglichen, wurden alle Items, bei der die Beantwortung die Formulierung „zu jung“ enthält, als fehlend definiert.

Post hoc wurden Responderanalysen mit einer Responseschwelle von ≥ 15 % der Skalenspannweite für den Gesamtscore und die Subskalen für die On-Target-Population und die Overall-Population eingereicht.

Bewertung

Die Operationalisierung kann nur bedingt nachvollzogen werden. Es wurden beide Versionen des Fragebogens (selbstberichtete Version für Patienten und fremdberichtete Version für Erziehungsberechtigte) im Studienverlauf zu unterschiedlichen Studienvisiten für einzelne Patienten ($n = k$. A.) verwendet, wodurch zum Teil keine konsistente Erhebung möglich war. Nach Studienende wurden die ausgefüllten Fragebögen von Patienten ≤ 10 Jahren, die in einer Visite die selbstberichtete Fragebogenversion > 10 Jahre erhielten, nicht für die Auswertung berücksichtigt.

Patientenrelevanz

Der Endpunkt wird vorliegend als patientenrelevanter Endpunkt angesehen.

Der pU ordnet den Fragebogen der Lebensqualität zu. Der Fragebogen umfasst eine globale Funktionsskala und eine aus 5 Items bestehende Skala zur „Zufriedenheit“. Es ist nicht nachvollziehbar, dass der Endpunkt alle Aspekte der multidimensionalen gesundheitsbezogenen Lebensqualität adäquat abbilden kann. Der Fragebogen wird daher der Kategorie „Morbidität“ zugeordnet.

Validität

Die Validität des Erhebungsinstruments „PODCI“ wurde in vorherigen Nutzenbewertungsverfahren [6,7,8,10] bereits bewertet und wird im zugrundeliegenden Anwendungsgebiet als nur eingeschränkt valide eingeschätzt, auf Grundlage einer Studie von Daltroy et al (1998) [3] anhand 470 Patienten im Alter von 2 bis 18 Jahren mit verschiedenen muskuloskelettalen Erkrankungen.

Mit dem Dossier reicht der pU Validierungsstudien ein: anhand a) 24 Patienten mit DMD für die Version für Erziehungsberechtigte (Ergebnisse nur für die globale Funktionsskala sowie die

Subskalen „Transfer/Basismobilität“ und „Sport / Physikalische Funktion“ berichtet) [11], b) 570 gehfähigen Patienten mit Zerebralparese für die Version für Erziehungsberechtigte [1] und c) 417 pädiatrischen Patienten mit Osteogenesis Imperfecta für beide Versionen [27].

Insgesamt zeigen Untersuchungen aus den Validierungsstudien unterschiedliche psychometrische Eigenschaften der Fragebogenversionen für Erziehungsberechtigte und Heranwachsende auf. Die vollständige Erhebung der einzelnen Subskalen ist vom Alter der Patienten abhängig. Für die Version der Erziehungsberechtigten in der Indikation „Zerebralparese“ wird eine Zusammenfassung der Items auf 4 Subskalen befürwortet. Zur Untersuchung der Konstruktvalidität beider Versionen zeigten sich nur für die physikalischen Subskalen moderate bis hohe Korrelationen. Zur Änderungssensitivität liegen keine konsistenten Ergebnisse vor.

Die nicht einheitliche Erhebung durch eine Fragebogenversion über den Studienverlauf und die im Anschluss vom pU post hoc durchgeführten Anpassungen der Auswertung des Fragebogens (Nicht-Berücksichtigung und Imputation der erhobenen Daten für Patienten ≤ 10 Jahre mit beantworteter Fragebogenversion > 10 Jahre) wird als kritisch gesehen. Hinzu kommen Limitationen aufgrund limitierter Daten der psychometrischen Eigenschaften beider Versionen und deren Vergleichbarkeit. Es liegen keine Informationen zur Anzahl an betroffenen Patienten vor.

Aufgrund unzureichender Operationalisierung, unzureichender Validität und Änderungssensitivität sowie unklarer Patientenrelevanz der Subskala „Zufriedenheit“ wird der PODCI in der Nutzenbewertung nicht berücksichtigt.

Muskelfettanteil

Der Endpunkt „Muskelfettanteil“ wird in der Nutzenbewertung aufgrund nicht gegebener Patientenrelevanz nicht berücksichtigt.

Operationalisierung

Beschreibung

Der Muskelfettanteil verschiedener Muskeln (M. vastus lateralis, M. biceps femoris long head, M. semitendinosus, MM. quadriceps femoris und Hamstring-Muskulatur) wurde mittels Magnetresonanztomografie (MRT) und/oder Magnetresonanzspektroskopie (MRS) gemessen. Es handelt sich hierbei um bildgebende Verfahren.

Bewertung

Die Operationalisierung ist nachvollziehbar.

Patientenrelevanz/Validität

Der Endpunkt „Muskelfettanteil“ wird vom pU ergänzend in Modul 4 dargestellt. Der pU gibt den Endpunkt als Surrogat für den Verlust der Steh- und Gehfähigkeit an. Diese Endpunkte wurden in der Studie EPIDYS durch andere Messinstrumente (RFTF, 10-MWT, 6-MWT) direkt durch ein standardisiertes Vorgehen erhoben. Bildgebende Verfahren, wie MRT/MRS, sind nicht per se patientenrelevant. Geeignete Validierungsstudien für eine Surrogatvalidierung wurden nicht eingereicht. Aufgrund der fehlenden Patientenrelevanz wird die Validität des Endpunkts nicht beurteilt.

2.3.3 Lebensqualität

Es wurde kein geeignetes Messinstrument zur Erhebung der Lebensqualität eingereicht.

2.3.4 Sicherheit

Unerwünschte Ereignisse

Der Endpunkt „Unerwünschte Ereignisse“ wird in der Nutzenbewertung berücksichtigt.

Operationalisierung

Beschreibung

Als unerwünschtes Ereignis (UE) wurden alle unerwünschten medizinischen Ereignisse gewertet, die nicht zwingend kausal mit dem Studienmedikament zusammenhängen. Medizinische Zustände, die bereits vor Studienstart bestanden, wurden nur als UE gewertet, wenn sie sich nach Einnahme des Studienmedikaments verschlechterten. Abnormale Laborbefunde oder Testergebnisse waren nur UE, wenn sie klinische Anzeichen oder Symptome induzierten, klinisch signifikant waren oder eine Behandlung erforderten.

UE wurden ab Erhalt der ersten Dosis der Studienmedikation bis zur letzten Dosis der Studienmedikation (Studienende oder vorzeitiger Abbruch der Studie) erhoben. Wenn ein Patient nicht in die einarmige Verlängerungsstudie (Studie 51) wechselte, wurden UE bis zu 4 Wochen nachbeobachtet und es fand im Anschluss eine finale Follow-up-Visite statt.

Es finden sich keine Angaben, ob Ereignisse, die der Grunderkrankung zugeordnet werden, nicht als UE ausgewertet wurden.

Die Klassifizierung der Nebenwirkungen erfolgte anhand des „Medical Dictionary for Regulatory Activities“ (MedDRA), Version 24.1, nach Systemorganklasse (SOC) und Preferred Term (PT). Die Einteilung in Schweregrade erfolgte gemäß „Common Terminology Criteria for Adverse Events“ (CTCAE), Version 4.03.

Schwerwiegende unerwünschte Ereignisse (SUE) wurden wie folgt definiert:

- Tödlich (resultiert im Tod);
- lebensbedrohlich (bezogen auf den Zeitpunkt des Eintretens des Ereignisses, nicht auf Ereignisse, die in anderen Schweregraden hypothetisch/möglicherweise lebensbedrohlich sind);
- stationärer Krankenhausaufenthalt oder Verlängerung eines bestehenden Krankenhausaufenthaltes erforderlich;
- anhaltende/erhebliche Beeinträchtigung oder Behinderung;
- angeborene Anomalie oder Geburtsfehler;
- medizinisch signifikantes Ereignis oder Erfordernis eines medizinischen Eingriffs, um eine der oben genannten Folgen zu verhindern (vermutete Übertragung infektiöser Erreger durch das Studienmedikament wurden als medizinisch signifikantes Ereignis eingeschlossen).

Weitere wichtige medizinische Ereignisse, welche nicht direkt lebensbedrohlich sind oder zum Tod oder zu einem Krankenhausaufenthalt führten, wurden als SUE betrachtet, wenn sie den Patienten gefährdeten und/oder einen medizinischen oder chirurgischen Eingriff erforderten, um Auswirkungen zu verhindern, die als SUE definiert wurden. Ein vorab geplanter Krankenhausaufenthalt war kein SUE.

Es wurden keine UE von besonderem Interesse definiert.

Um die Verblindung zu gewährleisten, unterschied sich das Studienpersonal zur Überwachung der Sicherheit vom Personal, das die Wirksamkeitsendpunkte erhob. Das Studienpersonal musste dieser Maßnahmen zur Aufrechterhaltung der Verblindung zustimmen. Eine weitere Handlungsanweisung war u. a. keine Weitergabe von Sicherheitsdaten an das Personal, das die Wirksamkeitsendpunkte erhob. Das Personal für die Testdurchführung durfte keine Fragen zur Einnahme der Studienmedikation oder UE stellen.

Bewertung

Die Operationalisierung ist nachvollziehbar. Der pU legt keine zusätzlichen Auswertungen unter Nichtberücksichtigung von erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung vor. Es ist nicht auszuschließen, dass Ereignisse der Grunderkrankung mit in die Erfassung der UE eingingen.

Patientenrelevanz

Das Auftreten von UE ist patientenrelevant. Die Patientenrelevanz von Laborparametern als Teil der UE ist unklar.

Validität

Die Erfassung von UE wird als valide bewertet.

Körpergröße, Körpergewicht, BMI

Die Endpunkte „Körpergröße“, „Körpergewicht“ und „BMI“ werden in der Nutzenbewertung aufgrund unklarer Validität nicht berücksichtigt.

Operationalisierung

Beschreibung

Anthropometrische Parameter wurden in der Studie als Sicherheitsendpunkte erhoben.

Körpergröße

Die Körpergröße der Patienten wurde – abhängig von der Fähigkeit sicher und unabhängig zu stehen – als Stehgröße oder Armspanne gemessen.

Für die Stehgröße wurde der Patient angewiesen gerade zu stehen, mit den Fersen so nah an der Wand wie möglich. Falls der Patient nicht in der Lage ist auf flachen Füßen zu stehen, muss der Abstand zwischen Boden und Ferse gemessen und von der Gesamtgröße abgezogen werden.

Die Armspanne wird sitzend erhoben. Dabei wird der Abstand zwischen den Mittelfingerspitzen gemessen, wobei die Arme seitlich waagerecht abgespreizt werden, die Handflächen nach vorne gerichtet. Wenn nötig kann eine weitere Person dabei helfen, die Arme in der korrekten Position zu halten. Wenn Muskelkontrakturen vorhanden sind, soll versucht werden, die Arme möglichst gerade zu halten. Dann wird entlang der Kontur des Armes gemessen; die Messung kann bis zur Drosselrinne (Sulcus jugularis) erfolgen und gedoppelt werden.

Körpergewicht

Im Rahmen der körperlichen Untersuchung wurde das Körpergewicht erhoben.

BMI

Im Rahmen der körperlichen Untersuchung wurde der BMI erhoben, der das Körpergewicht im Verhältnis zur Körpergröße beschreibt und mit folgender Formel berechnet wird:

$$\text{BMI} = \text{Körpergewicht (kg)} / (\text{Körpergröße (cm)} / 100)^2$$

Z-Scores für diese 3 Endpunkte (alters- (und geschlechts-)adjustierte Analysen) konnten nicht identifiziert werden.

Bewertung

Die Operationalisierung der anthropometrischen Parameter ist bedingt nachvollziehbar, da wichtige Informationen zur Erhebung fehlen. Es liegen nur zur Erhebung der Körpergröße Informationen vor, Angaben zur Erhebung des Körpergewichts konnten nicht identifiziert werden.

Es wurde nicht angegeben, zu welchem Tageszeitpunkt die Messungen durchgeführt wurden oder ob eine Reihenfolge vorgegeben war (analog zu den Funktionstests), oder wie genau das Gewicht (z. B. Messwiederholungen) erhoben wurde.

Zudem liegen keine alters- (und geschlechts-)adjustierte Analysen (z-Scores) vor. Der pU reicht post hoc durchgeführte Analysen der kontinuierlichen Daten ein.

Patientenrelevanz

Die anthropometrischen Parameter werden bei Erkrankungen, die die körperliche Entwicklung beeinflussen, als patientenrelevant angesehen. Aufgrund der empfohlenen Langzeiteinnahme von Glukokortikoiden, wird die Entwicklung der Patienten eingeschränkt.

Die Endpunkte „Körpergröße“ und „Körpergewicht“ werden in der vorliegenden Indikation entsprechend als patientenrelevant angesehen.

Bei dem Endpunkt „BMI“ handelt es sich um einen zusammengesetzten Endpunkt, der sich aus den Endpunkten „Körpergröße“ und „Körpergewicht“ zusammensetzt. Auf eine Darstellung der Ergebnisse des Endpunkts „BMI“ wird verzichtet.

Validität

Für die Endpunkte „Körpergröße“ und „Körpergewicht“ wurden keine alters- (und geschlechts-) adjustierten z-Scores berechnet. Die z-Scores spiegeln die Anzahl der Standardabweichungen eines Werts von den nach Alter und Geschlecht standardisierten normalen Durchschnittswerten wider. Vergleichbare Auswertungen konnten nicht identifiziert werden.

Eine alters- und geschlechtsadjustierte Auswertung mittels z-Scores wird bevorzugt. Die Endpunkte „Körpergröße“ und „Körpergewicht“ werden daher nicht in der Nutzenbewertung dargestellt.

2.3.5 Erhebungszeitpunkte

Eine Übersicht der Erhebungszeitpunkte der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte findet sich in Tabelle 8.

Tabelle 8: Erhebungszeitpunkte der berücksichtigten und ergänzend dargestellten Endpunkte in der Studie EPIDYS

Studienvisite	Base-line	M 3	M 6	M 9	M 12	M 15	M 18 EOS ¹⁾	M 19 FUV
		W 12	W 24	W 36	W 48	W 60	W 72	W 76
	Zeitfenster (Tage)	0	± 7 ²⁾	± 7 ²⁾	± 7 ²⁾	± 7 ²⁾	± 7 ²⁾³⁾	± 7 ⁴⁾
Endpunkt								
Todesfälle ⁵⁾	X							
Treppensteigen von 4 Stufen (4SC) ⁶⁾⁷⁾	X	X	X	X	X	X	X	
Aufstehen aus der Rückenlage (RFTF)	X	X	X	X	X	X	X	
10-Meter-Geh-/Lauftest (10-MWT)	X	X	X	X	X	X	X	
6 Minuten-Gehtest (6-MWT)	X	X	X	X	X	X	X	
Bestätigter Verlust der Aufstehfähigkeit	X	X	X	X	X	X	X	
Verlust der Gehfähigkeit ohne Hilfsmittel	X	X	X	X	X	X	X	
Unerwünschte Ereignisse	X							

¹⁾ EOS-Visite, auch bei vorzeitigem Abbruch der Studie. Patienten wurden gebeten innerhalb von 2 Wochen nach Erhalt der letzten Studiendosis an der Visite teilzunehmen.

²⁾ In der Protokolländerung zur COVID-19-Pandemie wurde das Visitenfenster von ± 7 Tage auf ± 7 Tage zuzüglich + 14 Tage erweitert. Diese Angabe findet sich nicht in den finalen Studienunterlagen.

³⁾ Während der COVID-19-Pandemie wurde das Visitenfenster auf ± 2 Monate erweitert.

⁴⁾ Nachbeobachtungsvisite 4 Wochen nach Erhalt der letzten Dosis der Studienmedikation.

⁵⁾ Todesfälle wurden im Rahmen der Sicherheit erhoben.

⁶⁾ Primärer Endpunkt.

⁷⁾ Der Endpunkt wurde als Ein-/Ausschlusskriterium während der Screening-Phase (4 Wochen vor Randomisierung) im Abstand von 2 Wochen zweimal wiederholt.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; FUV: Follow-up-Visite; M: Monat; RFTF: Rise From The Floor; W: Woche.

2.4 Statistische Methoden

Der finale SAP V6.0 vom 13.04.2022 und Addendum 1 (14.07.2022) wurden nach der letzten Studienvsiste (22.02.2022) erstellt, aber gemäß Angaben des pU vor Entblindung. SAP V5.0 wurde am 03.02.2022 erstellt, es wurden keine bewertungsrelevanten Änderungen zwischen den Versionen festgestellt.

Analysepopulationen

Overall-Population

- Alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation und mindestens einen nicht-fehlenden Post-Baseline-Wert des 4SC oder einen fehlenden Post-Baseline-Wert des 4SC aufgrund einer Gehunfähigkeit oder anderen körperlichen Einschränkungen, aufweisen. Die Overall-Population besteht aus der On- und der Off-Target-Population.
- Die primäre Auswertungspopulation für die Wirksamkeitsendpunkte war gemäß pU die On-Target-Population (s. u.). Analysen in der Overall-Population konnten laut SAP supportiv durchgeführt werden.
 - On-Target-Population: Subpopulation der Overall-Population, mit einer Muskeldegeneration zu Baseline, die ein beeinträchtigendes Maß erreichte und die voraussichtlich innerhalb eines Jahres eine signifikante Verschlechterung der Muskelfunktion aufweisen werden – definiert als Muskelfettanteil im M. vastus lateralis zu Baseline $> 5\%$ und $\leq 30\%$.
 - Off-Target-Population: Alle Teilnehmenden der Overall-Population, die einen Muskelfettanteil im M. vastus lateralis zu Baseline $\leq 5\%$ und $> 30\%$ aufweisen.

Overall Safety Analysis Set

- Alle Teilnehmenden, die mindestens eine Dosis der Studienmedikation erhalten haben. Die Auswertung erfolgt nach der tatsächlich erhalten Studienmedikation.
- Entspricht der Overall-Population.

Datenschnitte

- Präspezifizierte Interimsanalyse: 20.01.2020 (s. Fallzahlplanung)
- Finaler Datenschnitt: 22.02.2022

Fallzahlplanung

Die Interimsanalyse führte zur Fortsetzung der Studie und einer Reduktion der Fallzahl von 213 (192 Personen zuzüglich 10 % Abbruchquote) auf 110 Personen in der primären Analysepopulation (On-Target-Population) aufgrund einer deutlich reduzierten SD (von 6 auf 3 Sekunden). Auch die Effektgröße „Veränderung im 4SC (Baseline vs. 18 Monate)“ wurde von 3 auf 2 Sekunden verringert, obwohl eine Anpassung der Effektgröße bei einer verblindeten Fallzahlrekalkulation nicht präspezifiziert war. Die Abbruchquote von 10 % auf 8 % reduziert. Es wird nicht von einer Auswirkung auf die Ergebnissicherheit ausgegangen. Zusätzlich sollten bis zu 50 Personen (ca. 35 % der Overall-Population) in die Off-Target-Population eingeschlossen werden.

Präspezifizierte Subgruppenanalysen

Folgende a priori definierten Subgruppenanalysen waren für die Sicherheitsendpunkte des Overall Safety Analysis Set beschrieben:

- Gewichtsadjustierte Exposition (mg/kg): $< \text{Median}$ vs. $\geq \text{Median}$
- Startdosis (höhere Dosis gemäß Protokollversion ≤ 4 vs. niedrigere Dosis gemäß Protokollversion > 4)

Es zeigten sich keine für die Nutzenbewertung relevanten Subgruppeninteraktionen.

Präspezifizierung geplanter und durchgeführter Analysen

Gemäß SAP sollten alle statistischen Analysen zu den Wirksamkeitsendpunkten zu Studienende (Woche 72 (= Monat 18) bzw. EOS) anhand der On-Target-Population durchgeführt werden. Analysen anhand der Overall-Population konnten supportiv durchgeführt werden, waren aber im SAP nicht spezifisch für Wirksamkeitsendpunkte definiert. Mit dem Dossier wurden post hoc durchgeführte Analysen der Wirksamkeitsendpunkte für die Overall-Population eingereicht.

Timed Function Tests (4SC, RFTF, 10-MWT) und 6-MWT

„Treppensteigen von 4 Stufen“ (Zeit in Sekunden) zu Woche 72 gegenüber Baseline war als primärer Endpunkt definiert.

Es wurden anhand der On-Target-Population für die Erhebungsinstrumente 4SC, RFTF, 10-MWT und 6-MWT Kovarianzanalyse (ANCOVA)-Modelle durchgeführt, adjustiert nach „Alter zum Studieneinschluss“, dem Stratifizierungsfaktor „Glukokortikoid-Behandlungsregime“ („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“) sowie den Baseline-Werten aller „Timed Function Tests“ (4SC, RFTF, 10-MWT) und des 6-MWT, zur Berechnung der Mittelwertdifferenz zum Studienende (Woche 72 / EOS) gegenüber Baseline. Fehlende Werte der Wirksamkeitsendpunkte wurden imputiert (siehe Abschnitt „Fehlende Werte und Daten-Imputationen“ in diesem Kapitel). Es sollte der zweiseitige p-Wert berichtet werden.

Mit dem Dossier wurden post hoc für die Overall-Population die oben beschriebenen ANCOVA für die Erhebungsinstrumente 4SC, RFTF, 10-MWT und 6-MWT eingereicht. Auf Anfrage der U. S. Food and Drug Administration (FDA) wurden zusätzlich „Mixed Models for Repeated Measures“ (MMRM) adjustiert für die gleichen oben beschriebenen Faktoren bzw. Kovariablen durchgeführt, jedoch ohne Ersetzung fehlender Werte.

Eine Rationale zur Adjustierung nach allen Baseline-Werten (4SC, RFTF, 10-MWT, 6-MWT) anstelle nur des Baseline-Werts des jeweiligen Endpunkts konnte nicht identifiziert werden.

Je nach Endpunkt fehlen Werte von 20 bis 30 % der Patienten. Der pU legt für beide Analysen (MMRM und ANCOVA) keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor.

Angaben zur Plausibilität der Annahme zufällig fehlender Werte, die den MMRM-Analysen zugrunde liegen, konnten nicht identifiziert werden.

Bestätigter Verlust der Aufstehfähigkeit

Gemäß SAP waren Überlebenszeitanalysen mittels Cox-Proportional-Hazard-Modell für die On-Target-Population präspezifiziert. Post hoc wurden dieselben Analysen unter Berücksichtigung der präspezifizierten Variablen für die Overall-Population berechnet.

Mit dem Dossier wurden, anders als beim Endpunkt „Verlust der Gehfähigkeit ohne Hilfsmittel“, keine Analysen zur Berechnung eines „Relativen Risikos“ (RR) eingereicht, obwohl diese mit ähnlichen Behandlungszeiten der Studienarme bei dem vorliegenden chronischen Krankheitsbild als adäquat eingeschätzt werden.

Verlust der Gehfähigkeit ohne Hilfsmittel

Laut SAP war nur eine deskriptive Auswertung für die On-Target-Population vorgesehen. Post hoc wurden zwei Analysen für die Berechnung des adjustierten RR und p-Werts mit Modul 4 eingereicht. Die Berechnung sollte jeweils mittels logistischer Regression erfolgen, zum einen unter der Berücksichtigung der präspezifizierten Variablen (inklusive Baseline-Werte des 4SC, RFTF,

10-MWT und 6-MWT) und zum anderen unter Berücksichtigung des Stratifizierungsfaktors „Glukokortikoid-Behandlungsregime“ und „Alter zu Studieneinschluss“ (ohne Baseline-Werte des 4SC, RFTF, 10-MWT und 6-MWT).

Es wurden keine Angaben zum Umgang zur Nullzellenkorrektur gemacht. Es wird von einer Addition des Faktors 0,5 auf alle Zellen bei keinem Ereignis ausgegangen.

Wenn das Modell nicht konvergierte erfolgte die Berechnung mittels Wald-Methode.

Sicherheit

Eine deskriptive Auswertung war gemäß SAP für das Overall Safety Analysis Set (entspricht der Overall-Population) beschrieben. Post hoc wurden, analog zum Endpunkt „Verlust der Gehfähigkeit ohne Hilfsmittel“, zwei Analysen zur Berechnung des RR eingereicht.

Fehlende Werte und Daten-Imputation

Gemäß SAP wurden fehlende Werte der Erhebungsinstrumente 4SC, RFTF, 10-MWT bzw. 6-MWT anhand präspezifizierter Imputationsregeln ersetzt.

Fehlende Werte wurden in 2 Klassen eingeteilt:

- Klasse 1: Andere Gründe, COVID-19
- Klasse 2: Gehunfähigkeit, körperliche Einschränkungen

Es wurden zunächst fehlende kontinuierliche Werte (4SC, RFTF, 10-MWT, 6-MWT) imputiert:

- Klasse 1: Andere Gründe, COVID-19
Mittelwert aller nicht-fehlenden Werte des Endpunkts des jeweiligen Zeitpunkts aller Patienten des zugehörigen Behandlungsarms innerhalb des gleichen Glukokortikoid-Behandlungsregimes und Gruppeneinteilung (On- oder Off-Target-Population).
- Klasse 2: Gehunfähigkeit, körperliche Einschränkungen
Gemäß SAP Non-Responder-Imputation: Entweder Imputation mit dem Wert „0“ oder als doppelter Maximalwert aller gemessenen nicht-fehlenden Werte aller Patienten (abhängig vom Endpunkt).

Zunächst wurden fehlende Werte der Klasse 1 ersetzt, im Anschluss fehlende Werte der Klasse 2.

Bei vorzeitigem Abbruch der Studie wurde entsprechend der oben beschriebenen Kategorien imputiert. Fehlt der Wert der letzten Visite vor Abbruch der Studie und wurde dieser als Klasse 2 definiert, so wurden alle nach Abbruch der Studie folgenden fehlenden Werte ebenfalls als Klasse 2 imputiert. Wurde während der letzten Visite vor Studienabbruch der fehlende Wert als Klasse 1 definiert, wurden alle nach Abbruch der Studie nachfolgenden fehlenden Werte als Klasse 1 imputiert.

Für den 10-MWT-Score und den RFTF-Score wurden fehlende Werte wie folgt differenziert:

- Klasse 1: Frage, ob der Test durchgeführt wurde, mit „nein“ dokumentiert; oder Frage, ob der Test durchgeführt wurde, mit „ja“ beantwortet und Grad > 2 dokumentiert, aber Angaben zur benötigten Zeit (Sekunden) fehlen.
- Klasse 2: Frage, ob der Test durchgeführt wurde, mit „ja“ beantwortet und Grad 1 oder Grad 2 dokumentiert, aber Angaben zur benötigten Zeit (Sekunden) fehlen.

Für den 6-MWT wurde zusätzlich wie folgt differenziert:

- Klasse 1: Frage, ob der Test durchgeführt wurde, wurde mit „nein“ dokumentiert.
- Klasse 2: Frage, ob der Test durchgeführt wurde, wurde mit „ja“ beantwortet und es wurde angegeben, dass der Test vorzeitig abgebrochen wurde. Angaben zur benötigten Zeit und zu zurückgelegter Strecke zum Testende fehlen.

Das beschriebene Vorgehen wird als kritisch gesehen: Wie bereits die FDA/EMA [4,29] diskutiert haben, handelt es sich um singuläre Imputation, die zu einer Reduktion des Standardfehlers führen können. Wenn die Variabilität der Daten unterschätzt wird, kann das zu einer Verringerung der p-Werte und einer Verkleinerung der Konfidenzintervalle (KI) führen. Die Imputation mit dem doppelten Maximalwert ist außerdem abhängig von den beobachteten Daten und kann nicht grundsätzlich als eine Art Non-Responder-Imputation gewertet werden. Sie kann außerdem bei großen Werten zu einer Veränderung der Verteilung führen, für die die ANCOVA ggf. nicht mehr geeignet sein könnte. Die im Anschluss an die Imputationen nach Klasse 1 bzw. Klasse 2 folgende Imputation mittels „Last Observation Carried Forward“ (LOCF) führt außerdem dazu, dass sich Patienten im Zeitverlauf nicht weiter verschlechtern können, was bei einer progredienten Erkrankung nicht realistisch erscheint. Es konnte darüber hinaus kein Hinweis gefunden werden, ob im ANCOVA-Modell mögliche heterogene Varianzen aufgrund der 2:1-Randomisierung oder auch durch die Imputationen berücksichtigt wurden.

Innerhalb der MMRM werden keine Imputationen vorgenommen und eine Varianzreduktion sowie eine Veränderung der Verteilung der Daten wird nicht angenommen. In Modul 4 wird der Umgang mit fehlenden Werten bei der MMRM-Auswertung nicht weiter spezifiziert. Es wird angenommen, dass der pU bei fehlenden Werten von einer „Missing at Random“ (MAR)-Annahme ausgeht. Die bei den MMRM getroffene Annahme, dass fehlende Werte zufällig fehlen, erscheint bei fehlenden Werten der Klasse 2 nicht zutreffend. Angaben zur Anzahl an fehlenden Werten Klasse 1 und Klasse 2 liegen nicht vor, weshalb eine abschließende Beurteilung nicht möglich ist. Die MMRM-Analyse wird gegenüber der Auswertung mit ANCOVA trotzdem bevorzugt.

Bestätigter Verlust der Aufstehfähigkeit

Für den Endpunkt „Bestätigter Verlust der Aufstehfähigkeit“ wurden fehlende Werte identifiziert und entsprechend Klasse 1 oder Klasse 2 zugeordnet. Fehlende Werte der Klasse 2 (körperliche Einschränkungen, Gehunfähigkeit) wurden als Ereignis gewertet.

Im SAP wird angegeben, dass bei Überlebenszeitanalysen, sofern das Ereignis nicht eintraf, zum letzten Erhebungszeitpunkt zu Visite Woche 72 zensiert wurde (inklusive imputierter Daten).

Mit SAP V6.1 Addendum (14.07.2022), nach Datenbanksperrung, wurde der Umgang mit fehlenden Werten für den RFTF und NSAA-Item 12 vereinheitlicht und dabei auf das Vorgehen zur Imputation fehlender Werte des NSAA verwiesen.

Zur Bildung des NSAA-Gesamtscores wurden mehrere Vorgehen zur Imputation beschrieben. Für den Endpunkt „Bestätigter Verlust der Aufstehfähigkeit“ wird davon ausgegangen, dass das im SAP-Addendum referenzierte Vorgehen der Beschreibung der Imputation bei Fehlen eines Items gemeint ist.

Für jeden Score wurden fehlende Werte als Klasse 1 definiert, wenn:

- Frage, ob der Test durchgeführt wurde, mit „Nein“ beantwortet und folgende Begründungen: „Durchführung des Tests abgelehnt“ oder „Andere“.
- Imputation erfolgte ähnlich wie bei kontinuierlichen Werten der Klasse 1, es wurde jedoch der Modalwert der verfügbaren Daten gebildet. Sollte ein Modalwert nicht eindeutig bestimmbar sein (mehrere Modalwerte), wurde der niedrigste gewählt.

Für jeden Score wurden fehlende Werte als Klasse 2 definiert, wenn:

- Frage, ob der Test durchgeführt wurde mit „Nein“ beantwortet und als Begründung „nicht in der Lage den Test aufgrund körperlichen Einschränkungen durchzuführen“.
- Imputation mit dem Wert „0“.

Falls sich ein Score von 0 ergibt, wird wie folgt umgegangen:

- Score 0 + fehlender Wert der Klasse 2 entspricht Score 0b (Durchführung nicht möglich aufgrund körperlicher Einschränkung).
- Score 0 + fehlender Wert der Klasse 1 entspricht Score 0a (Hilfsmittel zur Durchführung benötigt).

Die Imputation fehlender Werte der Klasse 1 durch Bildung eines Modalwerts wird, ähnlich wie beim Umgang fehlender kontinuierlicher Werte der Klasse 1, kritisch gesehen (siehe Ausführungen weiter oben). Die Rationale zur Wahl des niedrigsten Modalwerts bei mehreren verfügbaren Werten ist unklar. Das Vorgehen zur Identifikation fehlender Werte der Klasse 2 kann nachvollzogen werden.

Im SAP wird angegeben, dass bei Überlebenszeitanalysen, sofern das Ereignis nicht eintraf, zum letzten Erhebungszeitpunkt zensiert wurde, d. h. zur Visite in Woche 72. Fehlende Werte wurden imputiert, aber es ist unklar, ob auch fehlende Werte nach vorzeitigem Abbruch der Studie ergänzt wurden und inwiefern ein Abbruch der Studie ein nicht informatives Ereignis darstellt, welches i. d. R. in einer Überlebenszeitanalyse zensiert werden würde.

Weitere Auswertung ohne Imputationen wurden nicht vorgelegt.

Verlust der Gehfähigkeit ohne Hilfsmittel

Für den Endpunkt „Verlust der Gehfähigkeit ohne Hilfsmittel“ wurden fehlende Werte des 10-MWT und 6-MWT wie weiter oben beschrieben identifiziert und Klasse 1 oder Klasse 2 zugeordnet. Fehlende Werte der Klasse 2 (körperliche Einschränkungen, Gehunfähigkeit) wurden als Ereignis gewertet.

Für den 10-MWT wurde bei fehlenden Werten der Klasse 1 zusätzlich wie folgt vorgegangen:

- Baseline: Score > 2, fehlende Angaben zur benötigten Zeit: Imputation als fehlender Wert Klasse 1.
- Bei Visiten Post-Baseline mit verfügbaren Werten zur benötigten Zeit von ≤ 30 Sekunden, wurde der zuletzt verfügbare Wert imputiert, um zu bestimmen, ob der Patient Hilfsmittel zur Durchführung benötigte.

Das Vorgehen zur Identifikation fehlender Werte der Klasse 2 kann nachvollzogen werden. Fehlende Werte wurden imputiert, aber es ist unklar, ob auch fehlende Werte nach vorzeitigem Abbruch der Studie imputiert wurden.

Einschätzung der statistischen Auswertungen

Timed Function Tests (4SC, RFTF, 10-MWT) und 6-MWT

Mit dem Dossier wurden die für die On-Target-Population präspezifizierten Analysen zur benötigten Zeit (in Sekunden) bzw. für den 6-MWT die zurückgelegte Distanz (Meter) mittels adjustierter ANCOVA-Modelle zum Studienende für die Overall-Population eingereicht. Fehlende Werte wurden zunächst klassifiziert und anschließend imputiert. Zusätzlich wurden post hoc durchgeführte Analysen anhand beobachteter Werte mittels MMRM durchgeführt.

Es wird angenommen, dass die MMRM-Analysen unter MAR-Annahme durchgeführt wurden. Die Annahme, dass fehlende Werte zufällig fehlen, erscheint bei fehlenden Werten der Klasse 2 nicht zutreffend. Für beide Analysen war der Anteil an fehlenden Werten in beiden Behandlungsgruppen mit 20–30 % je nach Endpunkt ähnlich. Angaben zur Anzahl fehlender Werte der Klasse 1 und Klasse 2 bzw. differenziert nach Gründen liegen nicht vor. Eine abschließende Beurteilung der MMRM-Analysen ist daher nicht möglich.

Ergebnisse der Analysen mittels ANCOVA und MMRM wurden zu allen Visitenzeitpunkten und zum Studienende im Dossier dargestellt, wobei unklar ist, wie das Studienende definiert wurde. Gemäß SAP war eine separate Auswertung zur letzten Studienvisite in Woche 72 und bei vorzeitigem Studienabbruch zur EOS-Visite vorgesehen. Für die letzte Studienvisite zu Woche 72 wurde das Visitenfenster aufgrund COVID-19-bedingter Restriktionen von ± 7 Tage auf ± 2 Monate erweitert. Angesichts der beschriebenen Limitationen der letzten Studienvisite (Woche 72 / EOS) werden die Ergebnisse zu Woche 60 bevorzugt. Wegen der COVID-19-Pandemie bestand die Möglichkeit, die Studienvisite (Woche 60 ± 7 Tage) auf zuzüglich + 14 Tage zu erweitern. Es ist unklar, inwiefern Patienten nur in die Analyse eingingen, wenn diese innerhalb des Zeitfensters ± 7 Tage zu Woche 60 an der Studienvisite teilnahmen.

Zusammenfassend kann die vom pU für die MMRM-Analysen angegebene Anzahl an verfügbaren Werten zur Studienvisite zu Woche 60 und zum Studienende (Woche 72 / EOS) nicht nachvollzogen werden.

Aufgrund der fehlenden Angaben zu fehlenden Werten differenziert nach Gründen, der Limitationen zur unklaren Operationalisierung der letzten Studienvisite sowie der Unklarheiten zu den durch COVID-19 bedingten Anpassungen der Visitenzeitfenster werden die Ergebnisse der MMRM-Analysen zu Woche 60 (für RFTF zu Woche 48, aufgrund geringen Rücklaufs < 70 % zu Woche 60) ergänzend dargestellt.

Die präspezifizierten ANCOVA anhand imputierter Daten werden aufgrund zusätzlicher Limitationen im Umgang mit fehlenden Werten und der möglichen Auswirkungen auf die Varianzen bzw. den Umgang mit heterogenen Varianzen im Modell allgemein aufgrund der 2:1-Randomisierung aus Transparenzgründen im Anhang der Nutzenbewertung dargestellt.

Bestätigter Verlust der Aufstehfähigkeit

Es wurden post hoc für die Overall-Population durchgeführte Überlebenszeitanalysen des Endpunkts „Bestätigter Verlust der Aufstehfähigkeit“ anhand imputierter Daten bis Woche 72 / EOS eingereicht. Angaben fehlen differenziert nach Gründen für fehlende Werte (ca. 20–30 % in beiden Behandlungsarmen). Die Imputation fehlender Werte der Klasse 1 sollte analog zum Umgang mit fehlenden Werten des NSAA durch Bilden eines Modalwerts erfolgen. Dies wird, ähnlich wie beim Umgang mit fehlenden kontinuierlichen Werten der Klasse 1, kritisch gesehen (siehe Ausführungen weiter oben). Die Rationale zur Wahl des niedrigsten Modalwerts bei mehreren verfügbaren Werten ist unklar. Im SAP wird angegeben, dass bei Überlebenszeitanalysen, sofern das Ereignis nicht eintraf, zum letzten Erhebungszeitpunkt zensiert wurde, d. h. zur Visite in Woche 72. Es ist unklar, ob fehlende Werte nach vorzeitigem Abbruch der Studie ergänzt wurden und inwiefern ein Abbruch der Studie ein nicht informatives Ereignis darstellt, welches i. d. R. in einer Überlebenszeitanalyse zensiert werden würde.

Mit dem Dossier wurden, anders als beim Endpunkt „Verlust der Gehfähigkeit ohne Hilfsmittel“, keine Analysen zur Berechnung eines RR eingereicht, obwohl diese aufgrund ähnlicher Behandlungszeiten der Studienarme bei dem vorliegenden chronischen Krankheitsbild als adäquat eingeschätzt werden.

Die Überlebenszeitanalysen anhand der Overall-Population werden in der Nutzenbewertung ergänzend dargestellt.

Verlust der Gehfähigkeit ohne Hilfsmittel

Es wurden post hoc durchgeführte Analysen anhand der Overall-Population zu Woche 72 / EOS eingereicht. Die Analyse wurde mit imputierten Daten durchgeführt, wobei Angaben differenziert nach Gründen zu fehlenden Werten fehlen (ca. 20 % je Endpunkt). Es ist unklar, ob Imputationen

nach Studienabbruch, ähnlich wie bei den weiter oben beschriebenen Endpunkten, vorgenommen wurden und wie diese zur Studienvisite Woche 72 / EOS berücksichtigt wurden (siehe Ausführungen weiter oben).

Die post hoc berechneten adjustierten RR mittels logistischer Regression unter Berücksichtigung der präspezifizierten Kovariablen werden in der Nutzenbewertung ergänzend dargestellt.

Sicherheit

In der Nutzenbewertung wurden, wie beim Endpunkt „Verlust der Gehfähigkeit ohne Hilfsmittel“, post hoc zwei Analysen zur Berechnung des adjustierten RR mittels logistischer Regression mit dem Dossier eingereicht. Bei UE erscheint es inhaltlich nicht plausibel das RR adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT zu berechnen. Daher werden die RR adjustiert nach „Alter zu Baseline“, „Behandlungsgruppe“ und dem Stratifizierungsfaktor „Glukokortikoid-Behandlungsregime“ („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“) für das Overall Safety Analysis Set dargestellt.

2.5 Verzerrungspotential auf Studien- und Endpunktebene

Tabelle 9: Verzerrungspotential der Studie EPIDYS

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung Patientin/Patient	Verblindung Behandlungsperson	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential auf Studienebene
EPIDYS	Ja	Ja	Unklar ¹⁾	Unklar ¹⁾	Nein	Unklar ²⁾³⁾	Unklar

¹⁾ Es besteht ein Risiko einer potentiellen Entblindung des Patienten und des Studienpersonals, zuständig für die Überwachung der Sicherheit, aufgrund häufigeren Studienvisiten. Bei UE-bedingter Dosisreduktion war eine engmaschige wöchentliche Überwachung für 8 Wochen von bestimmten Laborparametern (Thrombozyten, Leukozyten, Hämoglobin) erforderlich. Im Givinstat-Arm traten bei 16 % der Patienten UE des PT „Thrombozytopenie“ und bei 18 % des PT „Verminderte Thrombozytenwerte“ auf, nicht aber im Placebo-Arm, die für jeweils bei 14 % der Patienten eine Dosisreduktion zur Folge hatte. Inwieweit die vom pU beschriebene Handlungsanweisung zur Vorbeugung einer potentiellen Entblindung ausreichend waren, ist unklar. Das Verzerrungspotential ist daher unklar.

²⁾ Nach Einschluss von ca. 50 % der Patienten, wurden umfassende Protokolländerungen vorgenommen, die einen Einfluss auf das Verzerrungspotential haben können (u. a. Änderungen der Startdosis, Anpassung von Dosismodifikationen, Anpassung der Ein-/Ausschlusskriterien). Angaben zur Anzahl an in die Studie aufgenommenen Patienten zum Zeitpunkt der Protokolländerung liegen nicht für alle Protokolländerungen vor. Das Verzerrungspotential ist unklar.

³⁾ Für den Stratifikationsfaktor „Glukokortikoid-Behandlungsregime“ („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“) liegen keine Angaben zur Anzahl der Personen pro Strata vor. Zu Baseline wurden ca. 76 % der Patienten mit Deflazacort behandelt. Am häufigsten wurde eine tägliche Applikation angegeben (79–83 %).

Abkürzungen: PT: Preferred Term; pU: pharmazeutischer Unternehmer; UE: Unerwünschte Ereignisse.

Das Verzerrungspotential auf Studienebene wird als unklar eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 10 dargestellt.

Tabelle 10: Verzerrungspotential der in der Nutzenbewertung berücksichtigten und ergänzend dargestellten Endpunkte der Studie EPIDYS

Endpunkt	Verblindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Todesfälle	Ja	Ja	Nein	Nein	Niedrig ¹⁾
Treppensteigen von 4 Stufen (4SC) ²⁾	Unklar ³⁾	Nein ⁴⁾	Nein	Ja ⁵⁾⁶⁾	Hoch
Aufstehen aus der Rückenlage (RFTF)	Unklar ³⁾	Nein ⁴⁾	Nein	Ja ⁵⁾⁶⁾	Hoch
10-Meter-Geh-/Lauftest (10-MWT)	Unklar ³⁾	Nein ⁴⁾	Nein	Ja ⁵⁾⁶⁾	Hoch
6-Minuten-Gehtest (6-MWT)	Unklar ³⁾	Nein ⁴⁾	Nein	Ja ⁵⁾⁶⁾	Hoch
Bestätigter Verlust der Aufstehfähigkeit	Unklar ³⁾	Nein ⁷⁾	Nein	Nein	Hoch
Verlust der Gehfähigkeit ohne Hilfsmittel	Unklar ³⁾	Nein ⁸⁾	Nein	Nein	Hoch
Unerwünschte Ereignisse	Unklar ³⁾	Ja	Nein	Nein	Unklar

¹⁾ Es wird von einem geringen Verzerrungspotential durch die im Studienverlauf vorgenommenen Protokolländerungen ausgegangen.

²⁾ Primärer Endpunkt.

³⁾ Bei UE-bedingter Dosisreduktion war eine engmaschige wöchentliche Überwachung für 8 Wochen von bestimmten Laborparametern (Thrombozyten, Leukozyten, Hämoglobin) erforderlich. Im Givinostat-Arm traten bei 16 % der Patienten UE des PT „Thrombozytopenie“ und bei 18 % des PT „Verminderte Thrombozytenwerte“ auf, nicht aber im Placebo-Arm, die für jeweils bei 14 % der Patienten eine Dosisreduktion zur Folge hatte. Inwieweit die vom pU beschriebene Handlungsanweisung zur Vorbeugung einer potentialen Entblindung ausreichend waren, ist unklar. Das Verzerrungspotential ist daher unklar.

⁴⁾ Es fehlen Angaben zur Anzahl fehlender Werte differenziert nach Gründen bzw. nach Klasse 1 und Klasse 2 (je nach Endpunkt Anteil an fehlenden Werten ca. 20–30%). Das Verzerrungspotential wird als hoch angesehen.

⁵⁾ Die primäre Analyse mittels ANCOVA mit imputierten Daten und die post hoc durchgeführten MMRM-Analysen anhand beobachteter Werte zeigten keine konsistenten Ergebnisse. Für beide Analysen wurde ein statistisch signifikanter Unterschied zu Woche 60 nur für den primären Endpunkt „Treppensteigen von 4 Stufen“ festgestellt. Bei der Veränderung zu Woche 60 (RFTF zu Woche 48) waren für den Placebo-Arm, anders als im Givinostat-Arm, die Werte der ANCOVA höher und das Konfidenzintervall breiter (gegenüber MMRM). Die Ergebnisse können zugunsten der Intervention verzerrt sein. Es ist unklar, weshalb eine höhere Variabilität in der Placebo-Gruppe zu beobachten ist, obwohl die Anzahl an fehlenden Werten zwischen den Behandlungsgruppen ähnlich ist.

⁶⁾ Die Ergebnisse werden zu Woche 60 (RFTF zu Woche 48, aufgrund Rücklaufquote < 70% zu Woche 60), aber nicht zum Studienende (Visite zu Woche 72 / EOS) dargestellt. Es ist unklar, inwiefern Studienabbrüche in die Analyse eingingen und inwiefern Patienten, die eine Erhebung auch außerhalb des präspezifizierten Visitenfensters (alle Studienvisiten: ± 7 Tage) bzw. erweiterten Visitenfensters (Woche 72 ± 2 Monate) berücksichtigt wurden.

⁷⁾ Post hoc durchgeführte Überlebenszeitanalysen anhand imputierter Daten der Overall-Population. Es fehlen Angaben fehlender Werte und differenziert nach Gründen (ca. 20–30 % in beiden Behandlungsarmen). Die gewählte Imputationsstrategie weist Limitationen auf (siehe Kapitel 2.4). Es ist unklar, ob fehlende Werte nach Studienabbruch imputiert wurden, und inwiefern ein Abbruch der Studie ein nicht informatives Ereignis darstellt, welches i. d. R. in einer Überlebenszeitanalyse zensiert werden würde. Das Verzerrungspotential wird als hoch angesehen.

⁸⁾ Post hoc durchgeführte Analysen anhand imputierter Daten der Overall-Population. Es fehlen Angaben fehlender Werte und differenziert nach Gründen (ca. 20 % je Endpunkt). Es ist unklar, inwiefern imputierte Daten, insbesondere Klasse 1, o. Studienabbrüche bei der Analyse berücksichtigt wurden. Das Verzerrungspotential wird als hoch angesehen.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; ANCOVA: Kovarianzanalyse; EOS: End of Study; ITT: Intention to Treat; MMRM: Mixed Model for Repeated Measures; PT: Preferred Term; pU: pharmazeutischer Unternehmer; RFTF: Rise From The Floor; UE: Unerwünschtes Ereignis.

2.6 Indirekter Vergleich

Für die im Anschluss an die randomisierte kontrollierte Studie (RCT) EPIDYS folgende einarmige Verlängerungsstudie 51 (Interimsdatenschnitt: 31.12.2021) war ein indirekter Vergleich mittels „Propensity Score Matching“ der Wirksamkeitsdaten mit den historischen Daten der Studien „ImagingDMD“ und „CINRG DNHS“ präspezifiziert. Der pU zieht den indirekten Vergleich in Modul 4 nicht für die Ableitung des Zusatznutzens heran. Im „European Public Assessment Report“ (EPAR) wird dieser nur supportiv beschrieben [4].

Es wurde ein SAP V.2 (07.11.2022) für die integrierte Wirksamkeitsanalyse mit dem Dossier eingereicht. In diesem SAP wird für die Sicherheitsdaten auf einen separaten SAP verwiesen. Für die Ergebnisse liegen zwei Output-Dateien vor. Eine Auswertung wurde am 30.11.2022 durchgeführt. Auf Anfrage der FDA wurden weitere Analysen durchgeführt (28.08.2023), wobei für diese Analyse kein separater SAP vorgelegt wurde.

Für den indirekten Vergleich wurden für die mit Givinostat behandelten Patienten Daten aus der Studie EPIDYS und der darauffolgenden einarmigen Verlängerungsstudie 51 herangezogen (N = 148). Es wurden zum einen Patienten, die in der Studie EPIDYS in den Interventionsarm randomisiert wurden (N = 118), und zum anderen Patienten (N = 30), die erstmalig Givinostat ab Einschluss in die einarmige Verlängerungsstudie (Studie 51) erhalten haben, für den indirekten Vergleich berücksichtigt. Es handelt sich hierbei um 30 von 36 Patienten, welche die Einschlusskriterien für die Off-Target-Population der Studie EPIDYS erfüllten, aber aufgrund des Rekrutierungsstopps der Off-Target-Population nicht mehr randomisiert wurden. Diese wurden nach Prüfung der Ein-/Ausschlusskriterien in die Verlängerungsstudie (Studie 51) aufgenommen. Patienten aus dem ehemaligen Placebo-Arm der Studie EPIDYS (N = 61) wurden hingegen für den indirekten Vergleich ausgeschlossen, obwohl diese nach Einschluss in die Verlängerungsstudie ebenfalls erstmalig mit Givinostat behandelt wurden. Eine Rationale hierzu konnte in den Studienunterlagen nicht identifiziert werden.

Angaben zu den beiden Studien „ImagingDMD“ und „CINRG DNHS“, die zusammen als externe Kontrolle ausgewertet wurden, liegen nur im SAP V.2 vor. Separate Studienunterlagen wurden mit dem Dossier nicht eingereicht.

Bei „ImagingDMD“ (iDMD) handelt es sich um eine multizentrische prospektive Beobachtungsstudie aus den USA. Zwischen 2010 und 2018 wurden 180 Personen mit DMD im Alter von 4 bis 18 Jahren eingeschlossen. Die mediane Beobachtungsdauer lag bei 4 Jahren (min; max: 0; 8). Die „Cooperative International Neuromuscular Research Group – Duchenne Natural History Study“ (CINRG DNHS) ist eine internationale Beobachtungsstudie (20 Studienzentren, 9 Länder), in die Patienten (N = k. A.) mit gesicherter DMD-Diagnose im Alter von 2 bis 28 Jahren eingeschlossen wurden. Die primäre Rekrutierung erfolgte zwischen 2006 und 2009, im Zeitraum zwischen 2012 und 2016 wurden zusätzliche Patienten mit DMD zwischen 4 und 8 Jahren eingeschlossen. Angaben zur medianen Beobachtungsdauer liegen nicht vor. Gemäß SAP V.2 wurden Funktionstests erhoben, wobei eine genaue Auflistung aller erhobenen Endpunkte fehlt.

In Orientierung an die randomisiert-kontrollierte Studie EPIDYS und die Nachbeobachtungsstudie (Studie 51), die für die mit Givinostat behandelten Patienten herangezogen wurden, dienten als Wirksamkeitsendpunkte das Alter beim permanenten Eintritt folgender Outcomes: „Verlust der Gehfähigkeit ohne Hilfsmittel“, „Durchführung des RFTF > 10 s“, „10-MWT > 10 s“, „Verlust der Aufstehfähigkeit“ und „Fähigkeit zum Treppensteigen, erhoben mittels 4SC“.

Für die externe Kontrolle (Daten aus „ImagingDMD“ und „CINRG DNHS“) fehlen genaue Angaben zu den Endpunkten und deren Operationalisierung. Inwieweit die Operationalisierung der Endpunkte vergleichbar ist, kann nicht beurteilt werden. Gemäß SAP V2.0 der integrierten Datenanalyse

erfolgte für die Studie „CINRG DNHS“, anders als in den Interventionsstudien EPIDYS und 51, keine Beurteilung der Durchführung des 10-MWT mittels einer Skala (0–6 Punkte).

Für die mit Givinostat behandelten Patienten gingen imputierte Werte in die Analyse ein. In der Studie EPIDYS und der einarmigen Verlängerungsstudie 51 wurden fehlende Werte in Klasse 1 und Klasse 2 eingeteilt und mittels singulärer Imputation eines Wertes ergänzt. Dieses Verfahren weist Limitationen im Umgang mit fehlenden Werten und der möglichen Auswirkungen auf die Varianzen auf (Limitationen siehe Kapitel 2.4). Für die Verlaufsstudien wurden keine Imputationen durchgeführt.

Als externe Kontrolle dienten Patienten der Verlaufsstudien „ImagingDMD“ und „CINRG DNHS“ (N = k. A.), für die Patienten nach den zentralen Ein-/Ausschlusskriterien der Studie EPIDYS ausgewählt (N = 197) wurden. Die mediane Beobachtungsdauer bis zum Interimsdatenschnitt (31.12.2021) betrug für den naiven Vergleich, vor dem Propensity Score (PS-)Matching, 25,8 Monate (min; max: 2,9; 50,1) in der Givinostat-Gruppe (N = 148) bzw. 36,3 Monate (min; max: 2,5; 110,6) in der externen Kontrolle (N = 197). Für den naiven Vergleich war die mediane Behandlungsdauer somit nur um ca. 7 Monate länger, als die für die Studie EPIDYS vorgesehene Behandlungsdauer von 18 Monaten.

Es sollte im Anschluss ein PS-Matching durchgeführt werden. Als Kovariable wurden „Alter zu Baseline“, „4SC zu Baseline“, „RFTF zu Baseline“, „Glukokortikoid zu Baseline (Deflazacort; Andere)“ angegeben. Nach dem PS-Matching reduzierte sich die Anzahl an Personen des Interventionsarms auf 126 (85,1 %) bzw. der externen Kontrolle auf 126 (64,0 %). Die präspezifizierte Auswertung der Endpunkte erfolgte anhand nicht-gematchter Daten (Givinostat-Patienten: N = 148 vs. externe Kontrolle: N = 197).

In der auf Anfrage der FDA post hoc durchgeführten Analyse (28.08.2023) werden Auswertungen für 142 mit Givinostat behandelte Patienten gegenüber 142 Patienten der externen Kontrolle verglichen. Für das aktualisierte PS-Matching wurden als Kovariablen jeweils der Baseline-Wert des 4SC, RFTF und 10-MWT sowie das Glukokortikoid-Regime (Deflazacort; Andere) genannt. Es liegen jedoch keine weiteren Angaben zur Methodik vor.

Es ist unklar, ob weitere Unterlagen mit zusätzlichen Informationen existieren, bspw. zur Identifikation und Auswahl der Confounder oder zu systematischen Recherchen nach externen Kontrollen.

Insgesamt werden die verfügbaren Informationen als unzureichend erachtet, um die Validität des indirekten Vergleichs abschließend zu bewerten. Ein nicht-adjustierter „naiver“ Vergleich der Interventionsgruppe mit der externen Kontrolle wird als nicht adäquat angesehen. Für den post hoc auf Anfrage der FDA durchgeführten indirekten Vergleich mittels PS-Matching, liegen keine Angaben zur Methodik vor. Der indirekte Vergleich wird aufgrund der genannten Limitationen nicht für die Nutzenbewertung herangezogen.

3 Ergebnisse der eingeschlossenen Studie

Dargestellt werden die Ergebnisse für den finalen Datenschnitt vom 22.02.2022 der Overall-Population der Studie EPIDYS.

3.1 Studiencharakteristika und Studienmedikation

Insgesamt wurden 359 Patienten gescreent, 179 Patienten (49,9 %) wurden randomisiert. Als häufigster Grund für den Ausschluss während des Screenings wurde Nicht-Erfüllen des Einschlusskriteriums „Muskelfettanteil im M. vastus lateralis > 10 bis < 30 % erhoben mittels funktionellen Algorithmus“ angegeben, das bis Protokollversion 4.0 gültig war.

36 Patienten (10,0 %) der Off-Target-Population haben das Screening abgeschlossen und erfüllten die Ein-/Ausschlusskriterien, wurden aber nicht randomisiert aufgrund eines Rekrutierungsstopps in die Off-Target-Population. 30 dieser Patienten wurden, nach Überprüfung der Ein-/Ausschlusskriterien, als Givinostat-naive Patienten in die einarmige Verlängerungsstudie (Studie 51) eingeschlossen.

Tabelle 11: Allgemeine Angaben; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS Allgemeine Angaben	Givinostat N = 118¹⁾	Placebo N = 61¹⁾
Overall-Population ²⁾ , n (%)	118 (100)	61 (100)
Overall Safety Analysis Set, n (%)	118 (100)	61 (100)
Patienten, die die Studie abgeschlossen haben	111 (94,1)	59 (96,7)
Vorzeitiger Abbruch der Studie, n (%)	7 (5,9)	2 (3,3)
Aufgrund von:		
Entzug der Einverständniserklärung	4 (3,4)	2 (3,3)
UE	3 (2,5)	0 (0)
Vorzeitiger Abbruch der Studienmedikation	4 (3,4)	0 (0)
Aufgrund von	k. A.	k. A.
Mediane Beobachtungsdauer	k. A.	k. A.
Mediane Behandlungsdauer, Tage (min; max) ³⁾	506 (24; 624)	507 (232; 596)

¹⁾ Alle randomisierten Patienten.

²⁾ Definition, siehe Kapitel 2.4.

³⁾ Berechnet als: (Zeitpunkt der letzten Dosis der Studienmedikation - Zeitpunkt der ersten Dosis der Studienmedikation) + 1.
Abkürzungen: k. A.: keine Angabe; UE: Unerwünschtes Ereignis.

Die Patientencharakteristika sind in Tabelle 12 dargestellt.

Tabelle 12: Charakterisierung der Studienpopulation; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS Charakterisierung der Studienpopulation	Givinostat N = 118	Placebo N = 61
<i>Alter (Jahre)</i> MW (SD) Median (min; max)	9,78 (2,02) 9,80 (6,30; 15,9)	9,97 (2,08) 9,90 (6,10; 14,4)
<i>Abstammung (genetisch), n (%)</i> schwarz/afroamerikanisch Indigen Amerika/Alaska asiatisch Indigen Hawaii/Pazifik weiß/kaukasisch andere	3 (2,5) 0 (0) 4 (3,4) 0 (0) 106 (89,8) 5 (4,2)	0 (0) 0 (0) 2 (3,3) 0 (0) 57 (93,4) 2 (3,3)
<i>Ethnie, n (%)</i> hispanisch/Latino nicht hispanisch/Latino	9 (7,6) 109 (92,4)	3 (4,9) 58 (95,1)
<i>Region¹⁾, n (%)</i> Europa und Israel USA und Kanada	85 (72,0) 33 (28,0)	36 (59,0) 25 (41,0)
<i>Vorherige Diagnosen/Prozeduren (SOC ≥ 10 %), n (%)</i> Patienten mit vorherigen Diagnosen/Prozeduren Psychiatrische Erkrankungen Skelettmuskulatur-, Bindegewebs- u. Knochenerkrankungen Erkrankungen des Gastrointestinaltrakts Erkrankungen des Nervensystems Untersuchungen	71 (60,2) 23 (19,5) 20 (16,9) 15 (12,7) 11 (9,3) 12 (10,2)	35 (57,4) 11 (18,0) 11 (18,0) 7 (11,5) 11 (18,0) 4 (6,6)
<i>Zeit seit Diagnose bei Einwilligung</i> MW (SD) Median (min; max)	5,45 (2,60) 5,40 (0,1; 11,4)	5,62 (2,72) 5,30 (0,5; 12,3)
<i>DMD-Mutation, n (%)</i> Deletion Duplikation Punktmutation	83 (70,3) 17 (14,4) 18 (15,3)	40 (65,6) 14 (23,0) 7 (11,5)
<i>Körpergröße (cm), n (%)</i> MW (SD) Median (min; max)	125,1 (7,92) 125 (109; 151)	126,7 (9,82) 127 (112; 163)
<i>Gewicht (kg), n (%)</i> MW (SD) Median (min; max)	31,2 (8,89) 29,7 (17,9; 64,1)	32,6 (10,8) 30,0 (23,6; 66,0)
<i>Steroid-Typ, n (%)</i> Prednison Deflazacort Andere	20 (16,9) 91 (77,1) 7 (5,9)	15 (24,6) 45 (73,8) 1 (1,6)
<i>Zeit seit Beginn der Corticosteroid-Therapie (Jahre)</i> MW (SD) Median (min; max)	3,62 (2,10) 3,40 (0,4; 8,9)	3,97 (2,13) 4,00 (0,5; 9,2)

Studie EPIDYS Charakterisierung der Studienpopulation	Givinostat N = 118	Placebo N = 61
<i>Einnahmeschema Corticosteroid-Therapie, n (%)</i>		
täglich	99 (83,9)	48 (78,7)
intermittierend	19 (16,1)	13 (21,3)
<i>Tägliche Corticosteroid-Therapie, n (%)</i>		
täglich	98 (83,1)	48 (78,7)
andere	1 (0,8)	0 (0)
<i>Intermittierende Corticosteroid-Therapie, n (%)</i>		
Jeden zweiten Tag	1 (0,8)	4 (6,6)
10 Tage ON – 5 Tage OFF	1 (0,8)	1 (1,6)
10 Tage ON – 10 Tage OFF	14 (11,9)	4 (6,6)
Am Wochenende	3 (2,5)	4 (6,6)
<i>Corticosteroid-Therapie, n (%)</i>		
Täglich	98 (83,1)	48 (78,7)
Täglich andere	1 (0,8)	0 (0)
Jeden zweiten Tag	1 (0,8)	4 (6,6)
10 Tage ON – 5 Tage OFF	1 (0,8)	1 (1,6)
10 Tage ON – 10 Tage OFF	14 (11,9)	4 (6,6)
Am Wochenende	3 (2,5)	4 (6,6)
<i>Stratifizierungsfaktor: Glukokortikoid-Regime, n (%)</i>		
Deflazacort tgl.		
Deflazacort intermittierend	k. A.	k. A.
Andere Corticosteroide tgl.		
Andere Corticosteroide intermittierend		

¹⁾ Eigene Berechnung.

Abkürzungen: DMD: Duchenne-Muskeldystrophie; k. A.: keine Angabe; MW: Mittelwert; SD: Standardabweichung; SOC: Systemorganklasse.

Exposition mit der Studienmedikation

Tabelle 13: Angaben zur Dosismodifikationen mit der Studienmedikation; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS Dosismodifikationen mit der Studienmedikation	Givinostat N = 118	Placebo N = 61
Patienten mit Modifikation der Dosis, n (%)	86 (72,9)	28 (45,9)
<i>Gründe für Modifikation, n (%)</i>		
UE	54 (45,8)	5 (8,2)
Andere ¹⁾	56 (47,5)	27 (44,4)
<i>Art der Modifikation der Studienmedikation, n (%)</i>		
Erhöhung	6 (5,1)	6 (9,8)
Reduktion ²⁾	57 (48,3)	7 (11,5)
Ausgelassene Dosis	47 (39,8)	24 (39,3)
Permanent unterbrochen	10 (8,5)	1 (1,6)
Vorübergehend unterbrochen	19 (16,1)	6 (9,8)

¹⁾ Weitere Angaben liegen nicht vor.

²⁾ Mediane Zeit bis zur ersten Dosisreduktion in Tagen (min; max): Givinostat (n = 57): 79 (16; 441); Placebo (n = 7): 164 (28; 497).

Abkürzungen: UE: Unerwünschtes Ereignis.

Protokollverletzungen und -abweichungen

Anhand der in der Studie EPIDYS dokumentierten Protokollverletzungen ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

Begleitmedikation

Anhand der in der Studie EPIDYS dokumentierten Begleitmedikation ergeben sich keine verzerrenden Aspekte für die Bewertung des Zusatznutzens.

3.2 Mortalität

In der Studie EPIDYS traten im Studienverlauf keine Todesfälle auf.

3.3 Morbidität

Treppensteigen von 4 Stufen (4SC) – ergänzend dargestellt

Tabelle 14: 4SC (Zeit in Sekunden), MMRM-Analysen anhand beobachteter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS 4SC (Zeit in Sekunden) ¹⁾ (ergänzend dargestellt)	Givinostat N = 118	Placebo N = 61
Baseline		
Verfügbare Werte, n (%)	117 (99)	61 (100)
Fehlende Werte, n (%)	1 (0,9)	0 (0)
MW (SD)	3,58 (1,26)	3,60 (1,27)
Woche 60²⁾		
Verfügbare Werte, n (%) ³⁾	97 (82,2)	48 (78,7)
Fehlende Werte, n (%)	21 (17,8)	13 (21,3)
MW (SD)	k. A.	k. A.
Veränderung zu Woche 60		
Verfügbare Werte, n (%) ³⁾	97 (82,2)	48 (78,7)
LS Mean [95%-KI] ⁴⁾⁵⁾	0,85 [0,50; 1,19]	1,75 [1,27; 2,24]
Veränderung zu Woche 60		
LS-Mean-Differenz [95%-KI] ⁴⁾⁵⁾		-0,91 [-1,50; -0,31]
p-Wert ⁴⁾⁵⁾⁶⁾		0,003

¹⁾ Primärer Endpunkt.

²⁾ Ergebnisse zum Studienende (Visite zu Woche 72 / EOS) werden nicht dargestellt. Es ist unklar, inwiefern Studienabbrüche in die Analyse eingingen und inwiefern Patienten mit einer Erhebung außerhalb des erweiterten Visitenfensters (Woche 72 ± 2 Monate) berücksichtigt wurden.

³⁾ Es liegen zu Woche 60 keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor. Fehlende Werte sind nicht allein durch Studienabbrüche (bis Studienende zu Woche 72 / EOS: Givinostat: n = 7 (5,9 %), Placebo: n = 2 (3,3 %)) erklärbar.

⁴⁾ MMRM anhand beobachteter Werte, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime.

⁵⁾ Es wurden jeweils die präspezifizierten Visitenzeitpunkte berücksichtigt.

⁶⁾ Aufgrund der ergänzenden Darstellung erfolgte keine klinische Relevanzeinschätzung

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; RFTF: Rise From The Floor; SD: Standardabweichung.

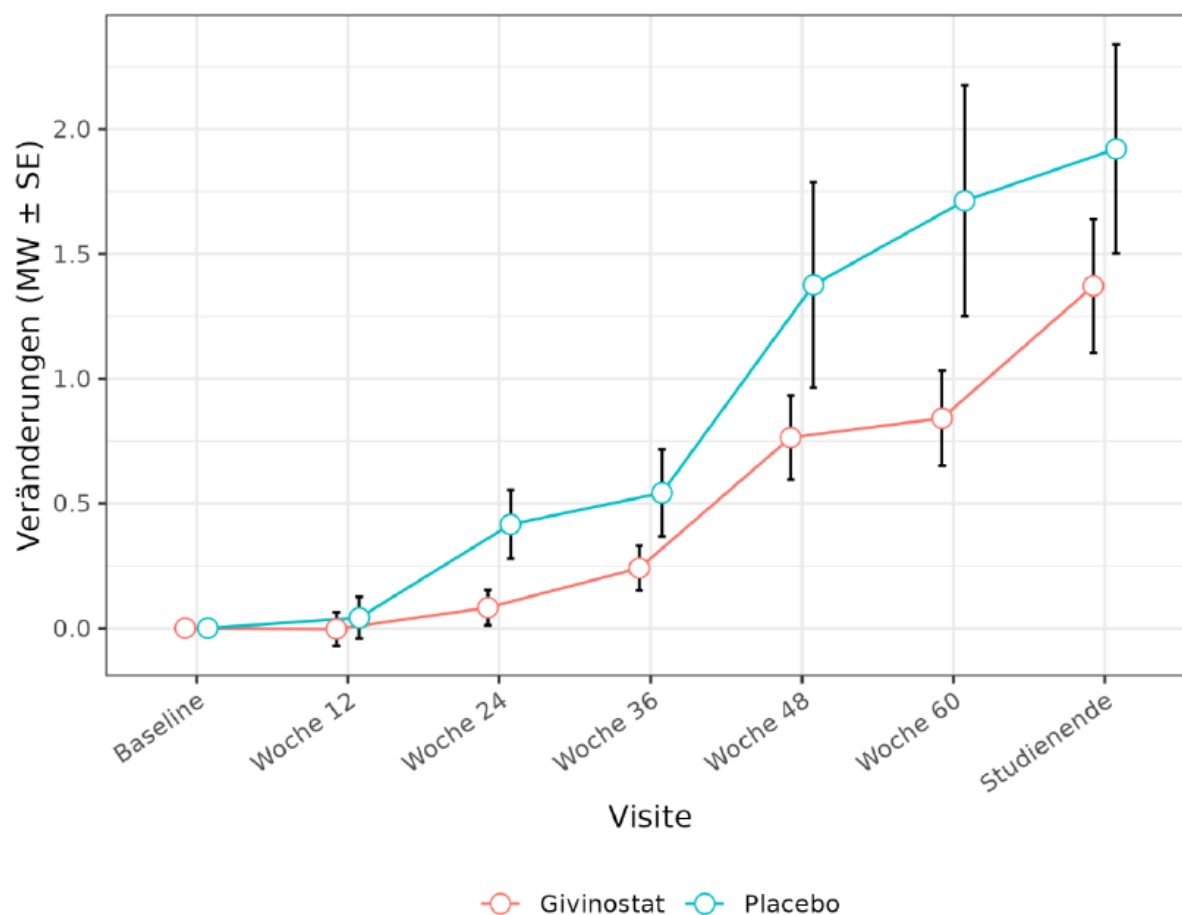


Abbildung 1: 4SC (Zeit in Sekunden), MMRM-Analyse anhand beobachteter Werte bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]

Aufstehen aus der Rückenlage (RFTF) – ergänzend dargestellt

Es werden die Ergebnisse des Endpunkts „Aufstehen aus der Rückenlage“ zu Woche 48 der Overall-Population mittels MMRM-Analyse anhand beobachteter Werte ergänzend dargestellt. Aufgrund des geringen Rücklaufs (< 70 %) werden die Ergebnisse zu Woche 60 nicht dargestellt.

Tabelle 15: RFTF (Zeit in Sekunden), MMRM-Analysen anhand beobachteter Werte zu Woche 48; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS RFTF (Zeit in Sekunden) (ergänzend dargestellt)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	117 (99)	61 (100)
Fehlende Werte, n (%) ¹⁾	1 (0,9)	0 (0)
MW (SD)	5,98 (2,42)	5,88 (2,25)
<i>Woche 48²⁾</i>		
Verfügbare Werte, n (%) ³⁾	84 (71,2)	43 (70,5)
Fehlende Werte, n (%) ³⁾	34 (28,8)	18 (29,5)
MW (SD)	k. A.	k. A.
<i>Veränderung zu Woche 48</i>		
Verfügbare Werte, n (%) ³⁾	84 (71,2)	43 (70,5)
LS Mean [95%-KI] ⁴⁾⁵⁾	1,36 [0,78; 1,95]	1,94 [1,12; 2,76]
<i>Veränderung zu Woche 48</i>		
LS-Mean-Differenz [95%-KI] ⁴⁾⁵⁾		-0,57 [-1,59; 0,44]
p-Wert ⁴⁾⁵⁾		0,27

¹⁾ Keine Angabe, weshalb für eine Person der Interventionsgruppe kein Wert vorliegt.

²⁾ Ergebnisse zum Studienende (Visite zu Woche 72 / EOS) werden nicht dargestellt. Es ist unklar, inwiefern Studienabbrüche in die Analyse eingingen und inwiefern Patienten mit Erhebung außerhalb der prädefinierten Visitenfensters bzw. auch außerhalb des erweiterten Visitenfensters (Woche 72 ± 2 Monate) berücksichtigt wurden.

³⁾ Es liegen zu Woche 48 keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor. Fehlende Werte sind nicht allein durch Studienabbrüche (bis Studienende zu Woche 72 / EOS: Givinostat: n = 7 (5,9 %), Placebo: n = 2 (3,3 %)) erklärbar.

⁴⁾ MMRM anhand beobachteter Werte, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime.

⁵⁾ Es wurden die präspezifizierten Visitenzeitpunkte berücksichtigt.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; RFTF: Rise From The Floor; SD: Standardabweichung.

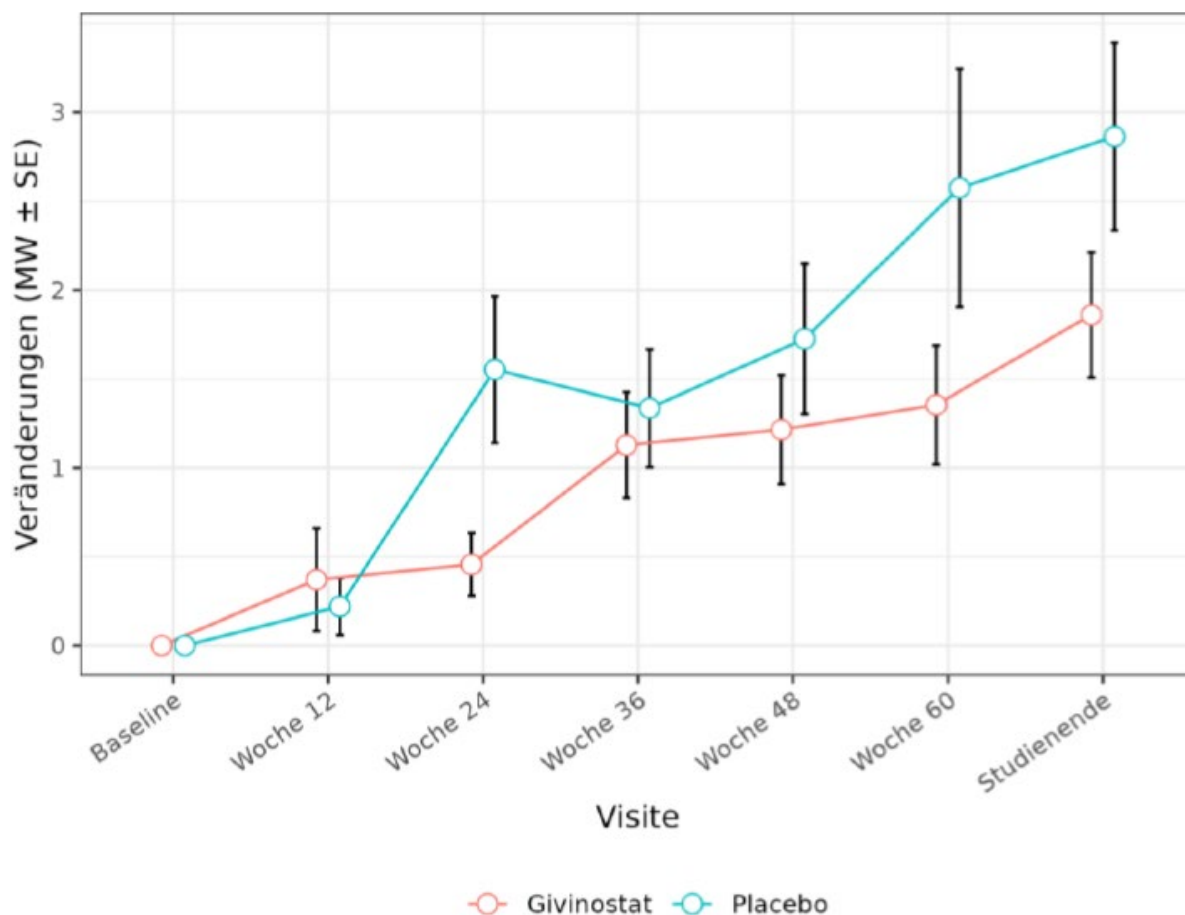


Abbildung 2: RFTF (Zeit in Sekunden), MMRM-Analyse anhand beobachteter Werte bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]

10-Meter-Geh-/Laufstest (10-MWT) – ergänzend dargestellt

Tabelle 16: 10-MWT (Zeit in Sekunden), MMRM-Analysen anhand beobachteter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS 10-MWT (Zeit in Sekunden) (ergänzend dargestellt)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	117 (99)	61 (100)
Fehlende Werte, n (%) ¹⁾	1 (0,9)	0 (0)
MW (SD)	5,54 (1,33)	5,30 (1,03)
<i>Woche 60²⁾</i>		
Verfügbare Werte, n (%) ³⁾	98 (83,1)	48 (78,7)
Fehlende Werte, n (%)	20 (16,9)	13 (21,3)
MW (SD)	k. A.	k. A.
<i>Veränderung zu Woche 60</i>		
Verfügbare Werte, n (%) ³⁾	98 (83,1)	48 (78,7)
LS Mean [95%-KI] ⁴⁾⁵⁾	0,93 [0,66; 1,21]	0,89 [0,50; 1,28]
<i>Veränderung zu Woche 60</i>		
LS-Mean-Differenz [95%-KI] ⁴⁾⁵⁾		0,04 [-0,43; 0,52]
p-Wert ⁴⁾⁵⁾		0,86

¹⁾ Keine Angabe, weshalb für eine Person der Interventionsgruppe kein Wert vorliegt.

²⁾ Ergebnisse zum Studienende (Visite zu Woche 72 / EOS) werden nicht dargestellt. Es ist unklar, inwiefern Studienabbrüche in die Analyse eingingen und inwiefern Patienten mit Erhebung auch außerhalb des erweiterten Visitenfensters (Woche 72 ± 2 Monate) berücksichtigt wurden.

³⁾ Es liegen zu Woche 60 keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor. Fehlende Werte sind nicht allein durch Studienabbrüche (bis Studienende zu Woche 72 / EOS: Givinostat: n = 7 (5,9 %), Placebo: n = 2 (3,3 %)) erklärbar.

⁴⁾ MMRM anhand beobachteter Werte, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime.

⁵⁾ Es wurden die präspezifizierten Visitenzeitpunkte berücksichtigt.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; RFTF: Rise From The Floor; SD: Standardabweichung.

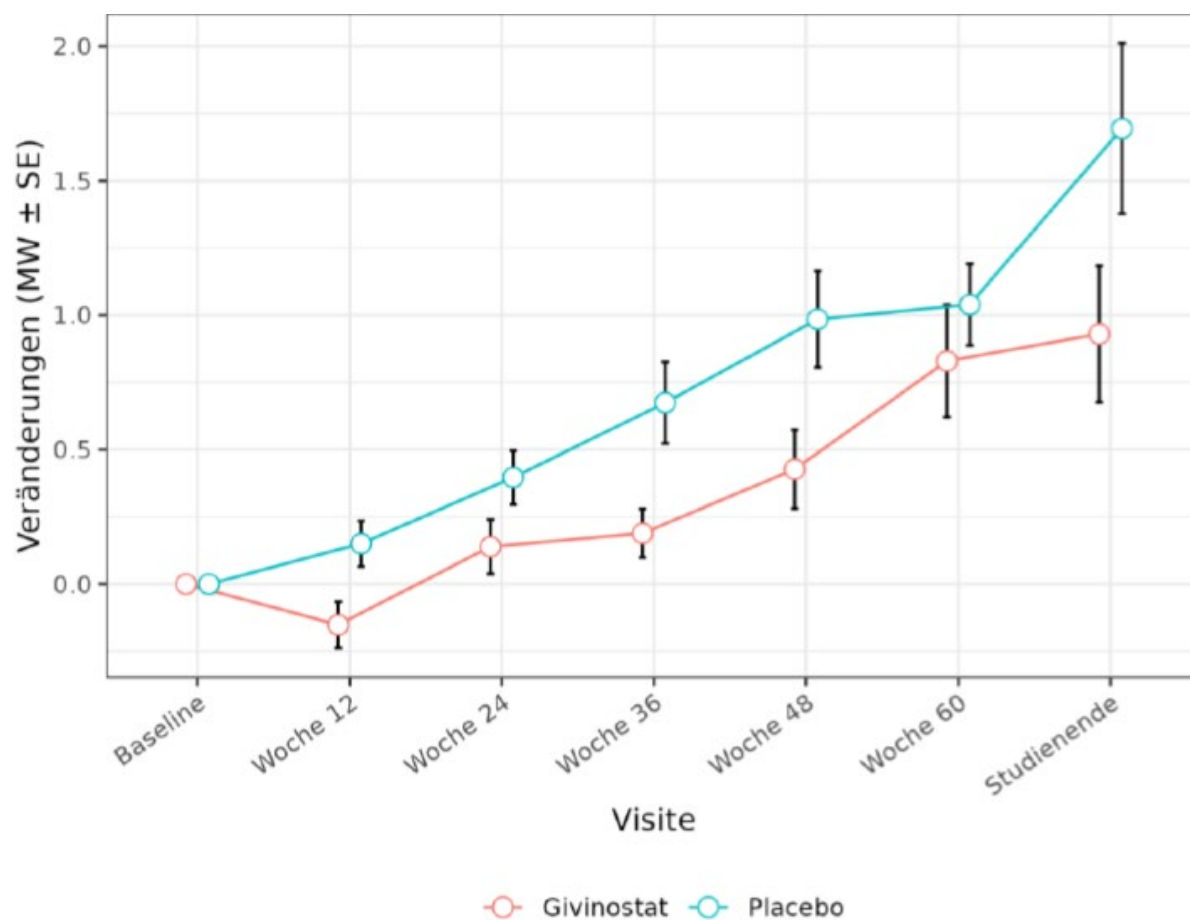


Abbildung 3: 10-MWT (Zeit in Sekunden), MMRM-Analyse anhand beobachteter Werte bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]

6-Minuten-Gehtest (6-MWT) – ergänzend dargestellt

Tabelle 17: 6-MWT (Distanz in Metern), MMRM-Analysen anhand beobachteter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS 6-MWT (Distanz in Metern) (ergänzend dargestellt)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	117 (99,2)	61 (100)
Fehlende Werte, n (%) ¹⁾	1 (0,9)	0 (0)
MW (SD)	398,9 (70,9)	393,7 (61,4)
<i>Woche 60²⁾</i>		
Verfügbare Werte, n (%) ³⁾	97 (82,2)	48 (78,7)
Fehlende Werte, n (%)	21 (17,8)	13 (21,3)
MW (SD)	k. A.	k. A.
<i>Veränderung zu Woche 60</i>		
Verfügbare Werte, n (%) ³⁾	97 (82,2)	48 (78,7)
LS Mean [95%-KI] ⁴⁾⁵⁾	-42,6 [-51,7; -33,5]	-47,0 [-59,8; -34,1]
<i>Veränderung zu Woche 60</i>		
LS-Mean-Differenz [95%-KI] ⁴⁾⁵⁾		4,39 [-11,5; 20,2]
p-Wert ⁴⁾⁵⁾		0,59

¹⁾ Keine Angabe, weshalb für eine Person der Interventionsgruppe kein Wert vorliegt.

²⁾ Ergebnisse zum Studienende (Visite zu Woche 72 / EOS) werden nicht dargestellt. Es ist unklar, inwiefern Studienabbrüche in die Analyse eingingen und inwiefern Patienten mit Erhebung auch außerhalb des erweiterten Visitenfensters (Woche 72 ± 2 Monate) berücksichtigt wurden.

³⁾ Es liegen zu Woche 60 keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor. Fehlende Werte sind nicht allein durch Studienabbrüche (bis Studienende zu Woche 72 / EOS: Givinostat: n = 7 (5,9 %), Placebo: n = 2 (3,3 %)) erklärbar.

⁴⁾ MMRM anhand beobachteter Werte, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime.

⁵⁾ Es wurden jeweils die präspezifizierten Visitenzeitpunkte berücksichtigt.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; RFTF: Rise From The Floor; SD: Standardabweichung.

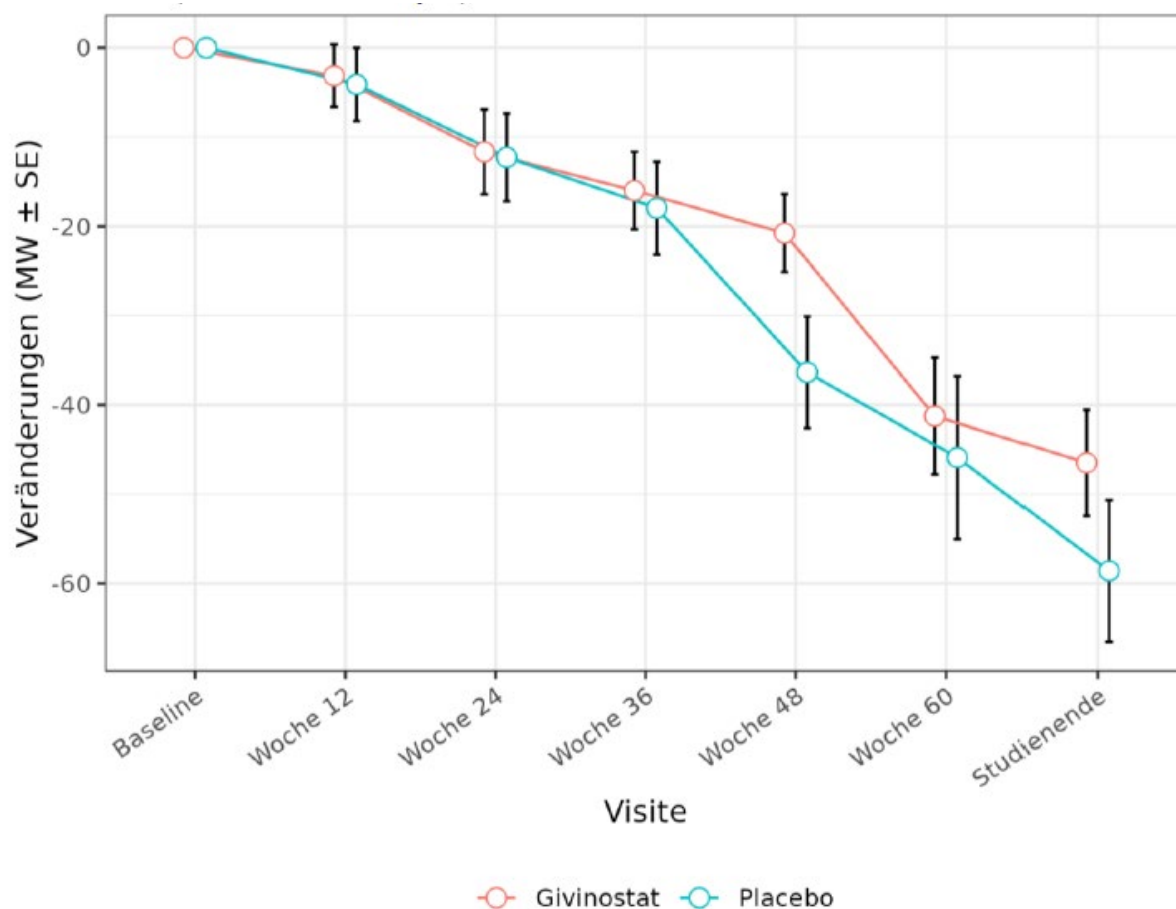


Abbildung 4: 6-MWT (Distanz in Metern), MMRM-Analyse anhand beobachteter Werte bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]

Bestätigter Verlust der Aufstehfähigkeit – ergänzend dargestellt

Tabelle 18: Bestätigter Verlust der Aufstehfähigkeit bis Studienende (Woche 72 / EOS)¹⁾, Überlebenszeitanalyse; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS Zeit bis zum bestätigten Verlust der Aufstehfähigkeit²⁾ (ergänzend dargestellt)	Givinostat N = 118	Placebo N = 61
Ereignisse, n (%) ³⁾ Zensierungen, n% ³⁾⁴⁾	1 (0,9) 117 (99,2)	2 (3,3) 59 (96,7)
Dauer in Monaten ⁵⁾	[11,0; 20,5]	[11,1; 19,6]
Mediane Zeit bis zum Ereignis, Monate [95%-KI]	n. a.	n. a.
Hazard Ratio [95%-KI] ⁶⁾ p-Wert ⁶⁾	0,14 [0,01; 2,26] 0,17	

¹⁾ Es ist unklar, inwiefern zur letzten Studienvsiste (Woche 72 / EOS) auch Studienabbrüche mitumfasst sind und inwiefern Patienten mit Erhebung außerhalb des erweiterten Visitenfensters (Woche 72 ± 2 Monate) berücksichtigt wurden.

²⁾ Verlust der Aufstehfähigkeit war definiert als RFTF-Grad 1 („Nicht in der Lage aus der Rückenlage aufzustehen, auch nicht mit Hilfe eines Stuhls“) und NSAA-Score 0b („Nicht in der Lage den Test durchzuführen“) und sollte an 2 konsekutiven Visiten bestätigt werden.

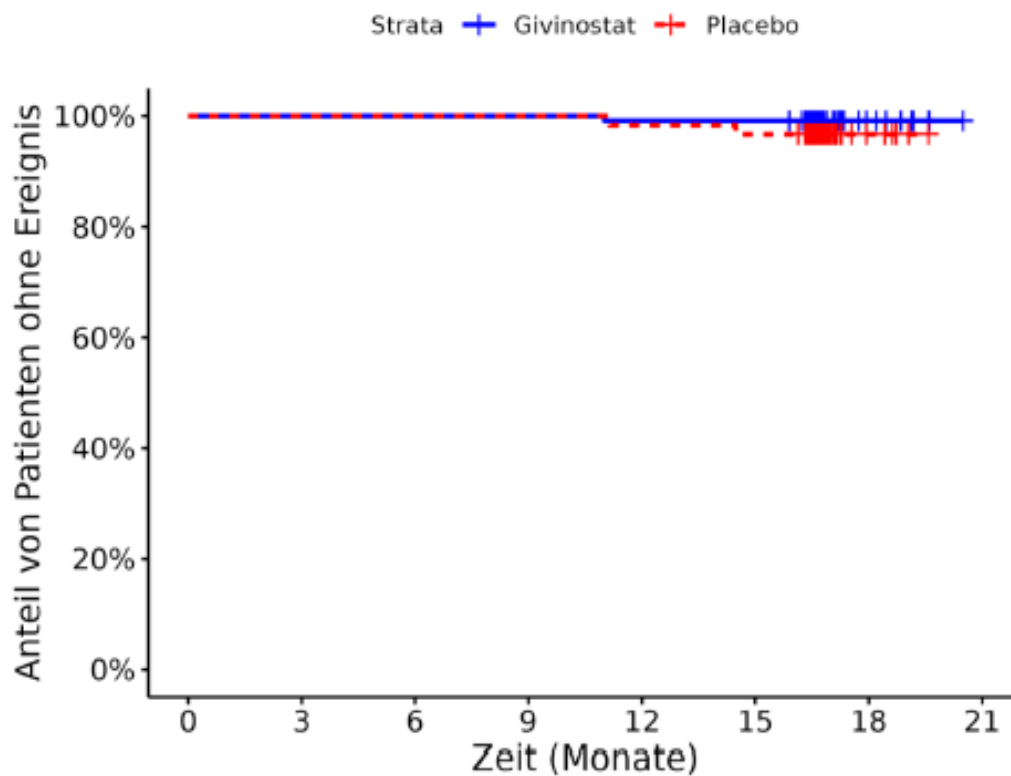
³⁾ Analyse anhand imputierter Daten (siehe Kapitel 2.4).

⁴⁾ Wenn eine Person die Kriterien bis zum Studienende nicht erfüllte, wurde die Zeit bis zum Ereignis mit dem Datum der letzten Erhebung zensiert (inklusive imputierter Daten), d. h. Woche 72. Es ist unklar, ob fehlende Werte nach Studienabbruch ebenfalls imputiert wurden.

⁵⁾ Unklar, inwiefern mit der Dauer der Zeitraum eines bestätigten Verlusts der Aufstehfähigkeit gemeint ist.

⁶⁾ Cox Proportional Hazards adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT, 6-MWT, Behandlungsgruppe, Alter sowie dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; KI: Konfidenzintervall; n. a.: nicht anwendbar; NSAA: North Star Ambulatory Assessment; RFTF: Rise From The Floor.



Patienten unter Risiko

Givinostat	118	118	118	118	117	117	8	0
Placebo	61	61	61	61	60	59	5	0

Abbildung 5: Bestätigter Verlust der Aufstehfähigkeit bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]

Verlust der Gehfähigkeit ohne Hilfsmittel – ergänzend dargestellt

Tabelle 19: Verlust der Gehfähigkeit ohne Hilfsmittel zu Woche 72 / EOS¹⁾; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS Verlust der Gehfähigkeit ohne Hilfsmittel ²⁾ (ergänzend dargestellt)	Givinostat N = 118	Placebo N = 61	RR [95%-KI] ³⁾ ; p-Wert ³⁾
Verlust der Gehfähigkeit (10-MWT und 6-MWT) Verfügbare Werte, n (%) ⁴⁾ Ereignisse, n (%)	117 (99,2) 1 (0,8)	59 (96,7) 2 (3,3)	0,26 [0,02; 2,79]; 0,27

¹⁾ Angaben, inwiefern zur letzten Studienvisite Studienabbrüche mitumfasst sind, liegen nicht vor.

²⁾ Ein Verlust der Gehfähigkeit ohne Hilfsmittel lag vor, wenn zu einer Visite zu einem beliebigen Studienzeitpunkt vor oder zu Studienende (letzte Studienvisite zu Woche 72 / EOS-Visite) folgende Kriterien erfüllt waren: Patient war aufgrund körperlicher Einschränkungen nicht in der Lage den 6-MWT durchzuführen und den 10-MWT in ≤ 30 Sekunden ohne Unterstützung oder Hilfsmittel zu absolvieren ($10\text{-MWT} \leq \text{Grad } 2$). Die Kriterien mussten bei allen darauffolgenden Visiten bis Studienende zu Woche 72 / EOS-Visite vorliegen.

³⁾ RR mit zugehörigem 95%-KI und p-Wert basieren auf logistischen Regressionen adjustiert nach „Alter zu Baseline“ und „Behandlungsgruppe“ sowie dem Stratifizierungsfaktor „Glukokortikoid-Behandlungsregime“. Wenn das Modell nicht konvergiert, wurde die Wald-Methode verwendet.

⁴⁾ Imputierte Daten (siehe Kapitel 2.4).

Abkürzungen: 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; RR: Relatives Risiko; KI: Konfidenzintervall;

3.4 Lebensqualität

Es wurden keine geeigneten Daten zur Lebensqualität erhoben.

3.5 Sicherheit

Es werden die Ergebnisse zu den UE bis Studienende (Woche 72 / EOS) der Studie EPIDYS des Overall Safety Analysis Set dargestellt. Die mediane Behandlungsdauer war in beiden Behandlungsarmen mit ca. 506 Tagen gleich.

Tabelle 20: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE; Studie EPIDYS, Overall Safety Analysis Set (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS Zusammenfassung der UE Personen mit mindestens einem...	Givinostat N = 118 n (%)	Placebo N = 61 n (%)	RR [95%-KI] ¹⁾ ; p-Wert ¹⁾
UE (ergänzend dargestellt)	112 (94,9)	57 (93,4)	-
UE CTCAE-Grad ≥ 3	5 (4,2)	1 (1,6)	2,58 [0,31; 21,6]; 0,39
SUE	8 (6,8)	2 (3,3)	2,07 [0,45; 9,44]; 0,35
UE, das zum Abbruch der Studienmedikation führte	4 (3,4)	0 (0)	4,69 [0,26; 85,70]; 0,30

¹⁾ RR mit zugehörigem 95%-KI und p-Wert basieren auf logistischen Regressionen adjustiert nach „Alter zu Baseline“ und „Behandlungsgruppe“ sowie dem Stratifizierungsfaktor „Glukokortikoid -Behandlungsregime“. Wenn das Modell nicht konvergiert, wurde die Wald-Methode verwendet.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; KI: Konfidenzintervall; RR: Relatives Risiko; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis; - : nicht berechenbar.

Unerwünschte Ereignisse

Tabelle 21: UE mit Inzidenz $\geq 10\%$ oder bei mind. 10 Personen in einem Behandlungsarm; Studie EPIDYS, Overall Safety Analysis Set (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS UE <i>MedDRA-Systemorganklasse</i> <i>Preferred Term</i>	Givinostat N = 118 n (%)	Placebo N = 61 n (%)	RR [95%-KI] ¹⁾ ; p-Wert ¹⁾
Infektionen und parasitäre Erkrankungen	62 (52,5)	40 (65,6)	0,80 [0,62; 1,05]; 0,11
Nasopharyngitis	31 (26,3)	19 (31,1)	0,84 [0,52; 1,39]; 0,48
Rhinitis	6 (5,1)	7 (11,5)	0,45 [0,15; 1,34]; 0,15
Infektion der oberen Atemwege	7 (5,9)	8 (13,1)	0,45 [0,17; 1,19]; 0,11
Erkrankungen des Blutes und des Lymphsystems	20 (16,9)	0 (0)	21,4 [1,31; 347]; 0,03
Thrombozytopenie ²⁾	19 (16,1)	0 (0)	20,3 [1,25; 331]; 0,03
Stoffwechsel- und Ernährungsstörungen	21 (17,8)	3 (4,9)	3,62 [1,12; 11,7]; 0,03
Hypertriglyceridämie ³⁾	14 (11,9)	1 (1,6)	7,24 [0,98; 53,8]; 0,053
Psychiatrische Erkrankungen	11 (9,3)	4 (6,6)	1,52 [0,53; 5,27]; 0,46
Erkrankungen des Nervensystems	31 (26,3)	17 (27,9)	0,91 [0,55; 1,56]; 0,73
Kopfschmerzen	28 (23,7)	14 (23,0)	1,01 [0,58; 1,82]; 0,98
Erkrankungen der Atemwege, des Brustraums und Mediastinums	32 (27,1)	15 (24,6)	1,16 [0,69; 2,03]; 0,58
Husten	13 (11,0)	9 (14,8)	0,77 [0,35; 1,79]; 0,53
Erkrankungen des Gastrointestinaltrakts	79 (66,9)	29 (47,5)	1,43 [1,09; 1,89]; 0,01
Abdominalschmerz ⁴⁾	25 (21,2)	9 (14,8)	1,47 [0,75; 3,10]; 0,29
Schmerzen Oberbauch	17 (14,4)	7 (11,5)	1,33 [0,60; 3,28]; 0,51
Diarrhö ⁴⁾	43 (36,4)	11 (18,0)	2,05 [1,20; 3,75]; 0,01
Erbrechen	34 (28,8)	8 (13,1)	2,29 [1,20; 4,82]; 0,02
Erkrankungen der Haut und des Unterhautgewebes	26 (22,0)	10 (16,4)	1,33 [0,70; 2,70]; 0,40
Ausschlag	11 (9,3)	1 (1,6)	6,03 [1,17; 109]; 0,09
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	34 (28,8)	21 (34,4)	0,83 [0,53; 1,34]; 0,44
Arthralgie	10 (8,47)	1 (1,6)	5,17 [0,68; 39,5]; 0,11
Rückenschmerzen	6 (5,1)	8 (13,1)	0,40 [0,14; 1,10]; 0,08
Myalgie	11 (9,3)	2 (3,3)	3,02 [0,81; 19,6]; 0,15
Schmerz in einer Extremität	8 (6,8)	7 (11,5)	0,59 [0,23; 1,55]; 0,29
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	31 (26,3)	12 (19,7)	1,42 [0,80; 2,64]; 0,25
Fieber	15 (12,7)	5 (8,2)	1,69 [0,69; 4,88]; 0,28
Untersuchungen	51 (43)	9 (14,8)	2,91 [1,68; 5,48]; 0,001
Triglyzeride im Blut erhöht ⁵⁾	14 (11,9)	3 (4,9)	2,51 [0,85; 10,26]; 0,14
Thrombozytenzahl vermindert ⁶⁾	21 (17,8)	0 (0)	22,4 [1,38; 364]; 0,03

Studie EPIDYS UE <i>MedDRA-Systemorganklasse Preferred Term</i>	Givinostat N = 118 n (%)	Placebo N = 61 n (%)	RR [95%-KI] ¹⁾ ; p-Wert ¹⁾
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	37 (31,4)	19 (31,1)	1,02 [0,65; 1,65]; 0,93
Kontusion	11 (9,3)	3 (4,9)	1,90 [0,55; 6,54]; 0,33
Sturz	15 (12,7)	13 (21,3)	0,60 [0,30; 1,22]; 0,15

¹⁾ RR mit zugehörigem 95%-KI und p-Wert basieren auf logistischen Regressionen adjustiert nach „Alter zu Baseline“ und „Behandlungsgruppe“ sowie dem Stratifizierungsfaktor „Corticosteroid-Behandlungsregime“. Wenn das Modell nicht konvergiert, wurde die Wald-Methode für die Effektschätzer, KI und p-Werte verwendet.

²⁾ Dosisreduktion nur im Givinostat-Arm berichtet (n = 17 (14,4 %)); Abbruch der Studienmedikation nur bei 1 Patienten (0,8 %) im Givinostat-Arm. Für 1 Patienten im Givinostat-Arm konnten keine Angaben zu Dosismodifikationen identifiziert werden.

³⁾ Dosisreduktion nur im Givinostat-Arm berichtet (n = 5 (4,2 %)); vorübergehende Dosisunterbrechung bei 7 Patienten (5,9 %) im Givinostat-Arm und 1 (1,6 %) im Placebo-Arm.

⁴⁾ Dosisreduktion nur im Givinostat-Arm berichtet (n = 1 (0,8 %)); Abbruch der Studienmedikation bei 1 Patienten (0,8 %) im Givinostat-Arm.

⁵⁾ Vorübergehende Dosisunterbrechung bei 5 Patienten (4,2 %) im Givinostat-Arm und 2 (3,3 %) im Placebo-Arm.

⁶⁾ Dosisreduktion nur im Givinostat-Arm berichtet (n = 16 (13,6 %)); Abbruch der Studienmedikation nur bei 1 Patienten (0,8 %) im Givinostat-Arm.

Abkürzungen: MedDRA: Medical Dictionary for Regulatory Activities; KI: Konfidenzintervall; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

Der pU legte keine zusätzlichen Auswertungen unter Nichtberücksichtigung von erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung vor. Es ist nicht auszuschließen, dass Ereignisse der Grunderkrankung mit in die Erfassung der UE eingingen.

Schwere UE (CTCAE-Grad ≥ 3)

Während der Studie EPIDYS traten bis zum finalen Datenschnitt (22.02.2022) keine schweren UE (CTCAE-Grad ≥ 3) nach Systemorganklasse und Preferred Term bei ≥ 5 % der Patienten in mindestens einem Studienarm auf.

Schwerwiegende UE

Während der Studie EPIDYS traten bis zum finalen Datenschnitt (22.02.2022) keine SUE nach Systemorganklasse und Preferred Term bei ≥ 5 % der Patienten in mindestens einem Studienarm auf.

UE von besonderem Interesse

In der Studie EPIDYS wurden keine UE von besonderem Interesse erhoben.

4 Diskussion der Methodik und Ergebnisse

Es werden die Ergebnisse der Zulassungsstudie EPIDYS der Overall-Population für den finalen Datenschnitt (22.02.2022) dargestellt.

Die Studie EPIDYS ist eine randomisierte, doppelblinde, Placebo-kontrollierte, multizentrische Phase-III-Studie zur Bewertung der Wirksamkeit und Sicherheit von Givinostat bei gehfähigen männlichen Patienten mit DMD im Alter von ≥ 6 Jahren. Die Randomisierung erfolgte im Verhältnis 2:1 in die Behandlungsarme Givinostat (N = 118) und Placebo (N = 61), stratifiziert nach dem eingesetzten Glukokortikoid-Behandlungsregime („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“). Die Dosierung von Givinostat war abhängig vom Körpergewicht.

Die Studiendauer betrug 19 Monate und umfasste eine 4-wöchige Screening- und eine 18-monatige Behandlungsphase. Die Behandlungszeit bis Studienende war in beiden Behandlungsarmen mit ca. 506 Tagen ähnlich. Im Anschluss konnten die Patienten an der einarmigen Verlängerungsstudie (Studie 51) teilnehmen. Die Studie EPIDYS wurde zwischen 06.07.2017 und 22.02.2022 in 50 Studienzentren in 11 Ländern durchgeführt. Die Studie ist abgeschlossen. Es liegt der finale Datenschnitt (22.02.2022) vor.

Als relevante Patientenpopulation für die vorliegende Nutzenbewertung wird die Overall-Population (Givinostat (N = 118) und Placebo (N = 61)) herangezogen. Diese setzt sich aus zwei Teilpopulation zusammen, der On-Target-Population (Muskelfettanteil im M. vastus lateralis $> 5\%$ bis $\leq 30\%$; Givinostat: N = 81, Placebo: N = 39) und der Off-Target-Population (Muskelfettanteil im M. vastus lateralis $\leq 5\%$ und $> 30\%$; Givinostat: N = 37, Placebo: N = 22).

Der primäre Endpunkt wurde zum Studienende (Woche 72 / EOS) anhand der On-Target-Population untersucht und war definiert als die Zeit (in Sekunden), die zum Steigen von 4 Treppenstufen (4SC) benötigt wird. Gemäß SAP war supportiv für die Wirksamkeitsendpunkte eine Auswertung anhand der Overall-Population möglich, während für UE eine deskriptive Auswertung des Overall Safety Analysis Set (entspricht der Overall-Population) beschrieben war. In den Studienunterlagen lagen nur Analysen der Wirksamkeitsendpunkte für die On-Target-Population vor. Für die Zulassung wurden von der Zulassungsbehörde EMA post hoc durchgeführte Wirksamkeitsanalysen anhand der Overall-Population angefordert [4]. Mit dem Dossier wurden die Ergebnisse zu Wirksamkeit und Sicherheit der Overall-Population eingereicht.

Nach Einschluss von ca. 50 % der Patienten, wurden umfassende Protokolländerungen während der Studie vorgenommen (u. a. Änderung der Startdosis und Dosismodifikationen, Anpassung der Ein-/Ausschlusskriterien). Angaben zu der Anzahl eingeschlossener Patienten zum Zeitpunkt aller weiteren Protokolländerungen liegen nicht vor. Das Verzerrungspotential ist unklar.

Die zu Beginn der Studie aus der On-Target-Population bestehende Studienpopulation wurde auf die Off-Target-Population erweitert und die Untergrenze des Einschlusskriteriums „Muskelfettanteil im M. vastus lateralis“ von $> 10\%$ auf $> 5\%$ der On-Target-Population reduziert. Anstelle des zunächst gewählten Algorithmus zur Vorhersage des Muskelfettanteils im M. vastus lateralis wurde eine Erhebung mittels MRT bereits zum Screening eingeführt, da das Nicht-Erfüllen der Kriterien der On-Target-Population mit 79 von 144 der häufigste Ausschlussgrund war. Ein erneutes Screening der Patienten war ab Protokollversion 6 ebenfalls möglich, wobei unklar ist, wie viele Patienten erneut an Screening-Visiten teilnahmen.

Da aufgrund der COVID-19-Pandemie eine persönliche Teilnahme an der Studienvsiste durch Schließung der Studienzentren limitiert war, wurde zunächst in der Protokolländerung zur COVID-19-Pandemie (12.03.2020) das Zeitfenster der Visiten (± 7 Tage) um zusätzlich + 14 Tage erweitert. Dieses Zeitfenster findet sich nicht mehr in der finalen Version des Studienprotokolls (Version 8

vom 08.04.2020). Die letzte Studienvisite zu Woche 72 konnte gemäß Protokolländerung zur COVID-19-Pandemie in einem erweiterten Zeitfenster von + 2 Monaten stattfinden (im finalen SAP: ± 2 Monate). Es bestehen somit Unklarheiten bezüglich der Angaben zu den Visitenfenstern, die im Laufe der Studie aufgrund der COVID-19-Pandemie geändert wurden. Angaben zur Anzahl der Personen, die während und außerhalb der beschriebenen Zeitfenster zu allen Visiten (ab Monat 3: ± 7 Tage zuzüglich + 14 Tage; anstelle von vorher ± 7 Tage) und des erweiterten Zeitfensters zur letzten Studienvisite zu Woche 72 (± 2 Monate; anstelle von vorher ± 7 Tage) erschienen, liegen für die Overall-Population nicht vor.

Eine Erhebung der Endpunkte zu den im Studienprotokoll definierten Zeitpunkten (in regelmäßigen Abständen von 3 Monaten) war für einen Teil der Studienpopulation gefährdet. Es ergeben sich zusätzliche Unsicherheiten, ob die erweiterten Zeitfenster der Visiten (ab Monat 3: ± 7 Tage + 14 Tage; Woche 72: ± 2 Monate) bei den Analysen zu den Wirksamkeitsendpunkten berücksichtigt wurden.

Studienpopulation

Von 359 gescreenten Patienten wurden 179 Patienten (Givinostat: N = 118; Placebo: N = 61) in die Studie EPIDYS eingeschlossen.

Im Studienverlauf wurden die Ein-/Ausschlusskriterien angepasst und die Studienpopulation, die zunächst aus der On-Target-Population (Muskelfettanteil im M. vastus lateralis $> 5\%$ (vorher $> 10\%$) bis $\leq 30\%$;) bestand, auf die Off-Target-Population (Muskelfettanteil im M. vastus lateralis $\leq 5\%$ und $> 30\%$) erweitert. Die definierten Kriterien des Muskelfettanteils im M. vastus lateralis der Off-Target-Population können zu einer Heterogenität in der Overall-Population beitragen. Angaben zu den Patientencharakteristika der beiden Subpopulationen (On- bzw. Off-Target) liegen nicht vor, sodass unklar bleibt, ob sich die beiden Subpopulationen in den Baseline-Charakteristika möglicherweise unterscheiden.

Insgesamt waren die Patientencharakteristika beider Behandlungsarme in der Overall-Population ähnlich. Die Patienten waren in beiden Behandlungsarmen im Median ca. 9,8 Jahre alt (min; max: 6; 16). Die Zeit seit Diagnose bis zum Studieneinschluss war in beiden Behandlungsarmen mit im Median 5 Jahren ähnlich. Der Anteil an Personen aus Europa, die an der Studie teilnahmen, war im Givinostat-Arm (n = 85 (72 %)) um ca. 13 % höher als im Placebo-Arm (n = 36 (59 %)).

Der Anteil an Studienabbrüchen ist in beiden Behandlungsarmen ähnlich. In der Overall-Population haben 6 % der Patienten des Givinostat-Arms (aufgrund des Entzugs der Einverständniserklärung (3 %) und UE (3 %)) sowie im Placebo-Arm 3 % der Patienten (aufgrund des Entzugs der Einverständniserklärung) die Studie abgebrochen.

Studienmedikation

Givinostat wurde abhängig vom Körpergewicht dosiert/verabreicht.

Bei einem verblindeten Review von hämatologischen Sicherheitsdaten vom 18.06.2018, bei dem Daten von Patienten nach 8-wöchiger Behandlung untersucht wurden, benötigten ca. 50–60 % der bis dahin eingeschlossenen Patienten eine Dosisreduktion/-unterbrechung aufgrund einer Thrombozytopenie. Dies hatte zur Folge, dass ab Protokollversion 6 (19.10.2018) eine Reduktion der Anfangsdosis um 1/3 und entsprechende Anpassungen der Dosismodifikationen der Prüfmedikation vorgenommen wurden. Eine engmaschige Überwachung von Laborparametern, insbesondere Thrombozytenzahl, Leukozytenzahl, Hämoglobin bei UE-bedingter Dosisreduktion, sollte durchgeführt werden (vollständiges Blutbild 1 x wöchentlich über 8 Wochen). Bei UE-bedingten Dosisunterbrechungen war eine wöchentliche Überprüfung der Laborparameter (bei Triglyceriden alle 2 Wochen) bis zur Normalisierung der Werte beschrieben.

Im Givinostat-Arm wurde eine Thrombozytopenie bei 16 % (und in der zugehörigen SOC „Erkrankungen des Blutes und des Lymphsystems“ bei 17 %) sowie beim PT „Thrombozytenzahl vermindert“ bei 18 % der Patienten beobachtet, die bei jeweils 14 % eine Dosisreduktion zur Folge hatte. Im Placebo-Arm traten diese UE bei keinem Patienten auf. Es wurden Handlungsanweisungen an das Studienpersonal formuliert, um das Risiko einer potentiellen Entblindung entgegenzuwirken. Die Erhebung der Wirksamkeitsendpunkten und Überwachung der Sicherheit erfolgte durch unterschiedliches Personal im Studienzentrum. Es ist unklar, inwiefern, trotz beschriebener Maßnahmen, eine Kenntnisnahme der Behandlungszuweisung der Patienten und des Studienpersonals, zuständig für die Überwachung der Sicherheit, aufgrund häufigeren Laboruntersuchungen, möglich war. Das Verzerrungspotential ist unklar.

Zudem ist unklar, inwiefern die Anpassung der Startdosis und Dosismodifikationen anhand der Protokolländerungen während des Studienverlaufs einen Einfluss auf das Verzerrungspotential der Studie hat. Patienten wurden nicht einheitlich mit der gleichen Dosierung jeweils innerhalb des Givinostat-/Placebo-Arms behandelt. Dies kann sich in einer Heterogenität bei den Ergebnissen der Wirksamkeits- und Sicherheitsendpunkte widerspiegeln. Eine Dosismodifikation war mit einem Anteil von 73 % im Givinostat-Arm höher als im Placebo-Arm (46 %), am häufigsten wurde eine Dosisreduktion vorgenommen (Givinostat: 48 %; Placebo: 12 %) aufgrund von UE (Givinostat: 46 %; Placebo: 8 %). Angaben zu den weiteren Gründen (je 44–47 % in beiden Behandlungsarmen) liegen nicht vor.

Alle Patienten erhielten begleitend ein Glukokortikoid, wobei mit ca. 77 % am häufigsten Deflazacort verabreicht wurde. Das häufigste Applikationsschema war mit 84 % (Givinostat) vs. 77 % (Placebo) eine tägliche Gabe. Für den Stratifikationsfaktor Glukokortikoid-Behandlungsregime („Deflazacort tgl.“, „Deflazacort intermittierend“, „Andere Corticosteroide tgl.“, „Andere Corticosteroide intermittierend“) liegen keine Angaben zur Anzahl an Personen pro Strata vor. Die mediane Behandlungszeit bis Studienende (Woche 72 / EOS) war in beiden Behandlungsarmen mit ca. 506 Tagen ähnlich.

Bewertung der Validität und Aussagekraft der Studie und des Verzerrungspotentials

Während der Studie wurden umfassende Protokolländerungen (bis inklusive Protokollversion 4.0: ca. 50 % der Teilnehmenden eingeschlossen) vorgenommen, die einen Einfluss auf das Verzerrungspotential haben können (u. a. Änderungen der Startdosis und Anpassung von Dosismodifikationen, Anpassung der Ein-/Ausschlusskriterien, inklusive Erweiterung der Studienpopulation zu On-/Off-Target). Neben den Änderungen zur Definition und Identifikation der On-/Off-Target-Population, wurde ein erneutes Screening der Patienten erlaubt. Patienten mit einer Triglycerid-Konzentration > 300 mg/dl (3,42 mmol/l) wurden ausgeschlossen, da im Studienverlauf vermehrt erhöhte Werte beobachtet wurden. Zusätzlich wurde eine Erweiterung der Überwachung von Laborparametern aufgrund von UE beschrieben. Aufgrund dieser Änderungen und der damit bedingten heterogenen Studienpopulation ist ein gewisser verzerrender Einfluss auf Studienebene möglich (unklares Verzerrungspotential).

Das Risiko einer potentiellen Entblindung der Patienten / des Studienpersonals, das für die Überwachung der Sicherheit zuständig war, bestand aufgrund einer engmaschigen wöchentlichen Überwachung von Laborparametern (Thrombozyten, Leukozyten, Hämoglobin) über 8 Wochen im Falle einer UE-bedingten Dosisreduktion. Bei Dosisunterbrechungen sollte bis zur Normalisierung eine wöchentliche Messung (Triglyceride alle 2 Wochen) erfolgen. Nur im Givinostat-Arm wurde eine Dosisreduktion bei Thrombozytopenie bei 16 % (und in der zugehörigen SOC „Erkrankungen des Blutes und des Lymphsystems“ bei 17 %) sowie PT „Thrombozytenzahl vermindert“ bei 18 % der Patienten beobachtet, wobei diese UE zu einer Dosisreduktion bei jeweils 14 % der Patienten führte. Im Placebo-Arm traten diese UE bei keinem Patienten auf.

Es bestehen Unklarheiten bezüglich der Angaben zu den Visitenfenstern, die im Laufe der Studie aufgrund der COVID-19-Pandemie geändert wurden (Studienvisiten ab Monat 3: ± 7 Tage zuzüglich + 14 Tage, anstelle ± 7 Tage; Studienvisite zu Woche 72: + 2 Monate bzw. ± 2 Monate). Angaben zur Anzahl an Personen, die außerhalb der definierten Zeitfenster zu den Studienvisiten erschienen, konnten nicht identifiziert werden. Eine einheitliche Erhebung der Endpunkte zu den im Studienprotokoll definierten Zeitpunkten war für einen Teil der Studienpopulation gefährdet, was Auswirkungen auf die Aussagesicherheit der Ergebnisse haben kann. Es ergeben sich Unsicherheiten, inwiefern Werte, die außerhalb der präspezifizierten Visiten-Zeitfenster erhoben wurden, für die Analyse der Wirksamkeitsendpunkte berücksichtigt wurden.

Die Patientencharakteristika waren in beiden Behandlungsgruppen vergleichbar. Für die 4 Strata des Stratifikationsfaktors „Glukokortikoid-Behandlungsregime“ liegen keine Angaben zur Anzahl an Personen vor. Am häufigsten wurde mit ca. 77 % in beiden Behandlungsgruppen Deflazacort verabreicht. Das häufigste Applikationsschema war mit 84 % bzw. 77 % eine tägliche Gabe.

Das Verzerrungspotential wird auf Studienebene als unklar bewertet.

Indirekter Vergleich

Für die im Anschluss an die randomisiert-kontrollierte Studie EPIDYS (Studie 48) folgende einarmige Verlängerungsstudie 51 (Interimsdatenschnitt: 31.12.2021) war ein indirekter Vergleich mittels PS-Matching der Wirksamkeitsdaten mit den historischen Daten der Studien „ImagingDMD“ und „CINRG DNHS“ präspezifiziert. Dieser wird aufgrund der in Kapitel 2.6 genannten Limitationen nicht für die Nutzenbewertung herangezogen. Maßgeblich sind die fehlenden Angaben zur Identifikation und Auswahl der Confounder und zu systematischen Recherchen nach externen Kontrollen. Studienunterlagen der externen Kontrollen liegen nicht vor, weshalb u. a. die Vergleichbarkeit der Operationalisierung der herangezogenen Endpunkte („Verlust der Gehfähigkeit ohne Hilfsmittel“, „Bestätigter Verlust der Aufstehfähigkeit“, „10-Meter-Geh-/ Laufstest“, „Aufstehen auf der Rückenlage“, „Treppensteigen von 4 Stufen“) unklar ist. Zudem konnte das geplante präspezifizierte PS-Matching nicht durchgeführt werden, da dies einen Ausschluss von 22 Personen unter der Givinostat-Behandlung bedeutet hätte. Für das auf Anfrage der FDA durchgeführte PS-Matching liegen keine weiteren Angaben zur Methodik vor. Für den naiven Vergleich war die mediane Behandlungsdauer bis zum Interimsdatenschnitt (31.12.2021) um ca. 7 Monate länger, als die für die Studie EPIDYS vorgesehene Behandlungsdauer von 18 Monaten.

Insgesamt werden die verfügbaren Informationen als unzureichend erachtet, um die Validität des indirekten Vergleichs abschließend zu bewerten. Der präspezifizierte nicht-adjustierte „naive“ Vergleich der Interventionsgruppe mit der externen Kontrolle wird als nicht adäquat angesehen. Auch der pU zieht den indirekten Vergleich in Modul 4 nicht zur Ableitung des Zusatznutzens heran. Im EPAR erfolgt eine supportive Darstellung der Ergebnisse [4].

4.1 Zulassungsstatus und Zulassungspopulation

Zulassungsstatus und Anwendungsgebiet

Givinostat wird zur Behandlung der Duchenne-Muskeldystrophie (DMD) bei gehfähigen Patienten ab 6 Jahren verabreicht. Laut EPAR handelt es sich hierbei um einen neuen Wirkstoff, der im vorliegenden Anwendungsgebiet eine bedingte Zulassung erhielt [4].

Gemäß Fachinformation wird die zugelassene Dosis vom Körpergewicht abhängig berechnet und zweimal täglich peroral eingenommen; das Mindestgewicht beträgt 15 kg. Ab einem Körpergewicht von 60 kg war eine Dosisanpassung nicht mehr notwendig. Bei Auftreten von unerwünschten Ereignissen ist eine zweimalige Dosisreduktion möglich, die ebenfalls Körpergewicht-abhängig berechnet wird. Die Entscheidung, ob die Behandlung bei Patienten, die ihre

Gehfähigkeit verlieren, fortgesetzt werden soll, wird nach Ermessen des ärztlichen Prüfungspersonals nach Beurteilung der gesamten Vorteile und Risiken getroffen [21].

In der Studie EPIDYS wurde im Studienverlauf die Dosierung (sowohl Startdosis als auch Dosismodifikationen) aufgrund von UE angepasst. Im Unterschied zur Fachinformation mit 4 Dosisstufen, wurde die Givinostat-Dosis in der Studie in 9 Dosierungsstufen abhängig vom Körpergewicht eingeteilt. Für die Zulassung wurde das Mindestgewicht von 10 kg (wie im Studienprotokoll vorgesehen) auf 15 kg angehoben, da in die Studie Patienten mit einem Mindestgewicht von 15 kg eingeschlossen wurden. Die in der Fachinformation beschriebene Dosierung beruht auf pharmakokinetischen Modellierungen, die der pU für die Zulassung eingereicht hat [4].

Das Anwendungsgebiet ist breiter definiert, als die in der Studie EPIDYS untersuchte Population. An der Studie nahmen Patienten ab 6 Jahren teil, die gehfähig waren. Gemäß Zulassung war keine Obergrenze für das Alter vorgeben. Das mittlere Alter in der Studie EPIDYS betrug ca. 10 Jahre. Es wurden Patienten bis 16 Jahre eingeschlossen. Daten für ältere Patienten liegen nicht vor. Jedoch verlieren aufgrund des progredienten Krankheitsverlauf die Patienten bei Krankheitsfortschritt die Gehfähigkeit. Im Studienverlauf trat dies nur bei 3 Patienten bis Studienende (Woche 72 / EOS) auf.

Übertragbarkeit auf deutschen Versorgungskontext

In der Studie EPIDYS lag der Anteil an Patienten aus dem europäischen Raum bei ca. 68 % und aus dem nordamerikanischen Raum bei 32 %. Ca. 90 % der Studienteilnehmenden waren kaukasischer Abstammung. Die Daten werden hinsichtlich der genannten Parameter als auf den deutschen Versorgungskontext übertragbar angesehen.

4.2 Mortalität

Es traten keine Todesfälle in der Studie EPIDYS bis zum Studienende (Woche 72 / EOS) auf.

4.3 Morbidität

Für die Endpunktkategorie „Morbidität“ legte der pU Daten zu den Endpunkten „Treppensteigen von 4 Stufen“, „Aufstehen aus der Rückenlage“, „Bestätigter Verlust der Aufstehfähigkeit“, „10-Meter-Geh-/Laufstest“, „6-Minuten-Gehtest“, „Verlust der Gehfähigkeit ohne Hilfsmittel“, „Funktionelle Leistungsfähigkeit“, „Körperliche Funktionsfähigkeit“ und „Muskelfettanteil“ vor.

4SC, RFTF, 10-MWT und 6-MWT

Für die Endpunkte „Treppensteigen von 4 Stufen“ (erhoben mittels 4SC), „Aufstehen aus der Rückenlage“ (erhoben mittels RFTF), „10-Meter-Geh-/Laufstest“ (erhoben mittels 10-MWT) und „6-Minuten-Gehtest“ (erhoben mittels 6-MWT) werden die post hoc durchgeführten MMRM-Analysen anhand beobachteter Werte aufgrund von Limitationen nur ergänzend in der Nutzenbewertung dargestellt.

Es wird angenommen, dass die MMRM-Analysen unter MAR-Annahme durchgeführt wurden. Die Annahme, dass fehlende Werte zufällig fehlen, erscheint bei fehlenden Werten der Klasse 2 nicht zutreffend. Der Anteil an fehlenden Werten war in beiden Behandlungsgruppen mit 20–30 % je nach Endpunkt ähnlich. Angaben zur Anzahl fehlender Werte differenziert nach Gründen (Anteil an Personen fehlender Werte der Klasse 1 und der Klasse 2) liegen nicht vor. Eine abschließende Beurteilung der MMRM-Analysen ist daher nicht möglich.

Im Gegensatz zum ANCOVA-Modell (primäre Analyse, siehe folgender Absatz) wird eine Varianzreduktion sowie eine Veränderung der Verteilung der Daten nicht angenommen.

Als primäre Analyse wurde eine ANCOVA anhand imputierter Daten durchgeführt. Das beschriebene Vorgehen zum Umgang mit fehlenden Werten wird als kritisch gesehen: Wie bereits die FDA und die EMA [4,29] diskutiert haben, handelt es sich um singuläre Imputationen, die zu einer Reduktion des Standardfehlers führen können. Wenn die Variabilität der Daten unterschätzt wird, kann das zu einer Verringerung der p-Werte und einer Verkleinerung der Konfidenzintervalle führen. Die Imputation mit dem doppelten Maximalwert ist außerdem abhängig von den beobachteten Daten und kann nicht grundsätzlich als eine Art Non-Responder-Imputation gewertet werden. Sie kann außerdem bei großen Werten zu einer Veränderung der Verteilung führen, für die die ANCOVA ggf. nicht mehr geeignet sein könnte. Die im Anschluss an die Imputationen nach Klasse 1 bzw. Klasse 2 folgende LOCF-Imputation führt außerdem dazu, dass sich Patienten im Zeitverlauf nicht weiter verschlechtern können, was bei einer progredienten Erkrankung nicht realistisch erscheint. Es konnte darüber hinaus kein Hinweis gefunden werden, ob im ANCOVA-Modell mögliche heterogene Varianzen aufgrund der 2:1-Randomisierung oder auch durch die Imputationen berücksichtigt wurden.

Die MMRM-Analyse wird daher gegenüber der Auswertung mit ANCOVA bevorzugt.

Ergebnisse der MMRM-Analysen wurden zu allen Visitenzeitpunkten und zum Studienende im Dossier dargestellt, wobei unklar ist, wie das Studienende definiert wurde. Gemäß SAP war eine separate Auswertung zur letzten Studienvisite in Woche 72 bzw. der EOS-Visite bei vorzeitigem Studienabbruch vorgesehen. Für die letzte Studienvisite zu Woche 72 wurde das Visitenfenster aufgrund durch COVID-19 bedingte Restriktionen von ± 7 Tage auf ± 2 Monate erweitert. Aufgrund der beschriebenen Limitationen der letzten Studienvisite (Woche 72 / EOS) werden die Ergebnisse zu Woche 60 bevorzugt. Aufgrund der COVID-19-Pandemie bestand die Möglichkeit die Studienvisite (Woche 60 ± 7 Tage) auf zuzüglich + 14 Tage zu erweitern. Es ist unklar, inwiefern Patienten nur in die Analyse eingingen, wenn diese innerhalb des Zeitfensters ± 7 Tage zu Woche 60 an der Studienvisite teilnahmen.

Zusammenfassend kann die vom pU für die MMRM-Analysen angegebene Anzahl an verfügbaren Werten zu den Studienvisiten zu Woche 60 und Studienende (Woche 72 / EOS) nicht nachvollzogen werden.

Es wurde für die Overall-Population nur für den primären Endpunkt „Treppensteigen von 4 Stufen“ ein statistisch signifikanter Vorteil von Givinostat gegenüber Placebo zu Woche 60 beobachtet (LS Mean: -0,91 (95%-KI: [-1,50; -0,31]); $p = 0,003$). Da der Endpunkt aufgrund von Limitationen in der statistischen Auswertung ergänzend dargestellt wird, erfolgte keine Einschätzung der klinischen Relevanz.

Die primären Analysen, mittels ANCOVA mit imputierten Daten sowie post hoc durchgeführte MMRM-Analysen anhand beobachteter Werte, zeigten keine konsistenten Ergebnisse. Für beide Analysen wurde ein statistisch signifikanter Unterschied zu Woche 60 nur für den primären Endpunkt festgestellt. Bei der Veränderung zu Woche 60 (RFTF zu Woche 48) waren für den Placebo-Arm, anders als im Givinostat-Arm, die Werte der ANCOVA höher und das Konfidenzintervall breiter (gegenüber MMRM). Die Ergebnisse können zugunsten der Intervention verzerrt sein. Es ist unklar, weshalb eine höhere Variabilität in der Placebo-Gruppe zu beobachten ist, obwohl die Anzahl an fehlenden Werten zwischen den Behandlungsgruppen ähnlich ist.

„Bestätigter Verlust der Aufstehfähigkeit“ und „Verlust der Gehfähigkeit ohne Hilfsmittel“

Die beiden Endpunkte „Bestätigter Verlust der Aufstehfähigkeit“ und „Verlust der Gehfähigkeit ohne Hilfsmittel“ werden aufgrund von Limitationen in der Auswertung ergänzend dargestellt.

Es wurden post hoc für die Overall-Population durchgeführte Überlebenszeitanalysen des Endpunkts „Bestätigter Verlust der Aufstehfähigkeit“ anhand imputierter Daten bis Woche 72 / EOS eingereicht. Angaben zu fehlenden Werten differenziert nach Gründen für beide Behandlungsarme wurden nicht eingereicht (ca. 20–30 % in beiden Behandlungsarmen). Die Imputation fehlender Werte der Klasse 1 sollte analog zum Umgang mit fehlenden Werten des NSAA, durch Bilden eines Modalwerts, erfolgen. Dies wird, ähnlich wie beim Umgang fehlender kontinuierlicher Werte der Klasse 1, kritisch gesehen (siehe Ausführungen Kapitel 2.4). Die Rationale zur Wahl des niedrigsten Modalwerts bei mehreren verfügbaren Werten ist unklar.

Im SAP wird angegeben, dass bei Überlebenszeitanalysen, sofern das Ereignis nicht eintraf, zum letzten Erhebungszeitpunkt zensiert wurde, d. h. zur Visite in Woche 72. Es ist unklar, ob fehlende Werte nach vorzeitigem Abbruch der Studie ergänzt wurden und inwiefern ein Abbruch der Studie ein nicht informatives Ereignis darstellt, welches i. d. R. in einer Überlebenszeitanalyse zensiert werden würde.

Es liegen nur Ergebnisse zum Studienende (Woche 72 / EOS) vor. Es bestehen Limitationen bezüglich der Operationalisierung der Zeitfenster aller Visiten, insbesondere zu Woche 72 (siehe Ausführungen weiter oben). Es ist unklar, ob die präspezifizierten Visiten-Zeitpunkte oder der Zeitpunkt der tatsächlichen Erhebung in die Analyse eingingen.

Die Kaplan-Meier-Kurven deuten auf das Auftreten von insgesamt 3 Ereignissen vor Woche 72 hin, weshalb eine Darstellung der Ergebnisse bis Studienende (Woche 72 / EOS) trotz dieser Limitationen als weniger verzerrend angesehen wird. Es zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen. Das Verzerrungspotential wird als hoch angesehen.

Mit dem Dossier wurde keine Analyse zur Berechnung eines adjustierten RR eingereicht, obwohl diese aufgrund ähnlicher Behandlungszeiten der Studienarme bei dem vorliegenden chronischen Krankheitsbild als adäquat eingeschätzt werden.

Für den Endpunkt „Verlust der Gehfähigkeit ohne Hilfsmittel“ wurden post hoc durchgeführte Analysen bis Woche 72 / EOS anhand der Overall-Population durchgeführt. Fehlende Werte wurden in Klasse 1 und Klasse 2 eingeteilt und fehlende Werte der Klasse 2 (Gehunfähigkeit, körperliche Einschränkungen) wurden als Ereignis gewertet. Angaben zu fehlenden Werten differenziert nach Gründen (ca. 20 % je Endpunkt) für beide Behandlungsarme wurden nicht eingereicht. Es zeigte sich kein statistisch signifikanter Unterschied zwischen den Behandlungsarmen. Das Verzerrungspotential wird als hoch angesehen.

Die Endpunkte „Funktionelle Leistungsfähigkeit (NSAA)“ und „Körperliche Funktionsfähigkeit (PODCI)“ werden aufgrund unzureichender Operationalisierung und unklarer Validität nicht zur Nutzenbewertung herangezogen. Der Endpunkt „Muskelfettanteil“ wird in der Nutzenbewertung aufgrund nicht gegebener Patientenrelevanz nicht berücksichtigt.

4.4 Lebensqualität

Es liegen keine bewertungsrelevanten Daten vor.

4.5 Sicherheit

Die Endpunkte „Körpergröße“, „Körpergewicht“ und „BMI“ werden in der Nutzenbewertung aufgrund unklarer Validität (keine Berechnung von alters- und (geschlechts-)adjustierten z-Scores) nicht berücksichtigt.

Unerwünschte Ereignisse

UE wurden ab dem Zeitpunkt der Einnahme der Studienmedikation bis zum Studienende (Woche 72 / EOS) erhoben. Falls Teilnehmende nicht an der einarmigen Verlängerungsstudie (Studie 51) teilnahmen, war eine 1-monatige Nachbeobachtungszeit nach der letzten Einnahme der Studienmedikation vorgesehen. Die mediane Behandlungszeit war bis zum Studienende mit ca. 506 Tagen zwischen den Behandlungsarmen vergleichbar. Der pU legte keine zusätzlichen Auswertungen unter Nichtberücksichtigung von erkrankungsbezogenen Ereignissen bzw. Ereignissen der Grunderkrankung vor. Es ist nicht auszuschließen, dass Ereignisse der Grunderkrankung in den beobachteten UE enthalten sind. Anzumerken ist jedoch, dass hier in den Daten kein PT identifiziert werden konnte, welches in einem direkten Zusammenhang gesehen wird.

In der Studie EPIDYS traten bei ca. 95 % der Personen UE auf. Bei den Gesamtraten konnte zwischen den Behandlungsarmen kein statistisch signifikanter Unterschied bei schweren UE, SUE und UE, die zum Abbruch der Studienmedikation führten, festgestellt werden. Der Anteil an Personen mit schweren UE lag bei < 5 % und mit SUE bei < 7 %. UE, die zum Abbruch der Studienmedikation führten, traten nur bei 4 Patienten (3 %) im Givinostat-Arm auf.

Für folgende UE jeglichen Schweregrads auf SOC- und PT-Ebene liegt ein statistisch signifikanter Unterschied zuungunsten von Givinostat gegenüber Placebo vor:

- SOC „Erkrankungen des Blutes u. des Lymphsystems“: RR = 21,4 (95%-KI: [1,31; 347]); p = 0,03
 - PT „Thrombozytopenie“: RR = 20,3 (95%-KI: [1,25; 331]); p = 0,03
- SOC „Stoffwechsel- und Ernährungsstörungen“: RR = 3,62 (95%-KI: [1,12; 11,7]); p = 0,03
- SOC „Erkrankungen des Gastrointestinaltrakts“: RR = 1,43 (95%-KI: [1,09; 1,89]); p = 0,01
 - PT „Diarrhö“: RR = 2,05 (95%-KI: [1,20; 3,75]); p = 0,01
 - PT „Erbrechen“: RR = 2,29 (95%-KI: [1,20; 4,82]); p = 0,02
- SOC „Untersuchungen“: RR = 2,91 (95%-KI: [1,68; 5,48]); p = 0,001
 - PT „Thrombozytenzahl vermindert“: RR = 22,4 (95%-KI: [1,38; 364]); p = 0,03

Für keine UE nach SOC und PT liegt ein statistisch signifikanter Unterschied zugunsten von Givinostat gegenüber Placebo vor.

Bei schweren UE, SUE und UE, die zum Abbruch der Studienmedikation führten, war der Anteil an Patienten < 5 % auf SOC- und PT-Ebene.

Eine engmaschige wöchentliche Überwachung über 8 Wochen von Laborparametern (Thrombozyten, Leukozyten, Hämoglobin) war im Falle einer UE-bedingten Dosisreduktion bzw. bis zum Erreichen der Normwerte bei Dosisunterbrechung notwendig, wodurch häufige Studienvisiten erforderlich waren. Nur im Givinostat-Arm wurde eine Thrombozytopenie bei 16 % (Dosisreduktion bei 14 %) bzw. in der zugehörigen SOC „Erkrankungen des Blutes und des Lymphsystems“ bei 17 % der Patienten beobachtet; im Placebo-Arm traten diese UE bei keinem Patienten auf. Zudem trat mit 18 % das PT „Thrombozytenzahl vermindert“ nur im Givinostat-Arm (Dosisreduktion bei 14 %) auf.

Eine Unterbrechung der Studienmedikation war, gemäß Studienprotokoll, bei moderater bis schwerer Diarrhö beschrieben, der PT „Diarrhö“ unabhängig vom Schweregrad trat in beiden Behandlungsgruppen auf, mit einem um ca. 18 % höheren Anteil im Givinostat-Arm (Givinostat: 36 %; Placebo: 18 %); eine Dosisreduktion war nur bei 1 Person im Givinostat-Arm notwendig.

Der PT „Hypertriglyceridämie“ trat bei 12 % im Givinostat- und bei 1 % im Placebo-Arm auf; jedoch war eine Dosisreduktion nur bei 4,2 % der Patienten im Givinostat-Arm, eine Dosisunterbrechung in beiden Behandlungsarmen bei 1,6 bis 4,2 % der Patienten notwendig.

Die EMA bewertet das Sicherheitsprofil von Givinostat in der vorliegenden Indikation als sicher. Es traten vom Schweregrad überwiegend milde UE auf. Auftretende UE können durch entsprechende Maßnahmen zur Überwachung von UE kontrolliert werden. [4]

Das Verzerrungspotential wird als unklar bewertet. Aufgrund von häufigeren Studienvisiten besteht das Risiko einer potentiellen Entblindung des Patienten und des Studienpersonals, das für die Überwachung der Sicherheit zuständig ist, da bei Auftreten bestimmter UE und damit verbundener Dosisreduktionen eine engmaschige Kontrolle der Laborparameter erforderlich ist.

5 Anforderung an eine qualitätsgesicherte Anwendung

Die Angaben der Fachinformation sind zu berücksichtigen.

6 Zusammenfassung der Nutzenbewertung

Givinostat ist zugelassen für die Behandlung der Duchenne-Muskeldystrophie (DMD) bei gehfähigen Patienten ab 6 Jahren zusammen mit einer Corticosteroid-Behandlung. Die Nutzenbewertung von Givinostat basiert auf der zulassungsbegründenden Studie EPIDYS, einer randomisierten, doppelblinden, Placebo-kontrollierten, multizentrischen Phase-III-Studie zur Bewertung der Wirksamkeit und Sicherheit von Givinostat bei gehfähigen männlichen Patienten mit DMD im Alter von ≥ 6 Jahren. Die Randomisierung erfolgte im Verhältnis 2:1 in die Behandlungsarme Givinostat (N = 118) und Placebo (N = 61), stratifiziert nach den eingesetzten Glukokortikoiden („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“). Die Dosierung von Givinostat war abhängig vom Körpergewicht. Die Studiendauer betrug 19 Monate und umfasste eine 4-wöchige Screening- und eine 18-monatige Behandlungsphase.

Die Ergebnisse der in der Nutzenbewertung berücksichtigten Endpunkte der Studie werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Givinostat ein statistisch signifikant positiver Effekt bei hohem oder unklarem ($\uparrow$) oder niedrigem ($\uparrow\uparrow$) Verzerrungspotential bzw. ein statistisch signifikant negativer Effekt bei hohem oder unklarem ($\downarrow$) oder niedrigem ($\downarrow\downarrow$) Verzerrungspotential bzw. kein Unterschied ($\leftrightarrow$) gezeigt werden konnte.

Weitergehende, über die Angaben in der nachfolgenden zusammenfassenden Tabelle hinausgehende, Erläuterungen zur Methodik und den Ergebnissen der Endpunkte, finden sich in den jeweiligen Kapiteln 2.3 bis 2.5 und 3.2 bis 3.5 der Nutzenbewertung.

Tabelle 22: Zusammenfassende Darstellung der Ergebnisse der Studie EPIDYS der Overall-Population bis Woche 72 / EOS (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS Darstellung der Ergebnisse	Givinostat N = 118		Placebo N = 61		Wirkstoff vs. Komparator	Effekt
Mortalität						
	N ¹⁾	Personen mit Ereignis, n (%)	N ¹⁾	Personen mit Ereignis, n (%)	RR [95%-KI] ²⁾ ; p-Wert ²⁾	
Todesfälle	118	0	61	0	-	↔
Sicherheit ³⁾						
Unerwünschte Ereignisse	N ¹⁾	Personen mit Ereignis, n (%)	N ¹⁾	Personen mit Ereignis, n (%)	RR [95%-KI] ²⁾ ; p-Wert ²⁾	
UE CTCAE-Grad ≥ 3	118	5 (4,2)	61	1 (1,6)	2,58 [0,31; 21,6]; 0,39	↔
SUE	118	8 (6,8)	61	2 (3,3)	2,07 [0,45; 9,44]; 0,35	↔
UE, das zum Abbruch der Studienmedikation führte	118	4 (3,4)	61	0 (0)	4,69 [0,26; 85,7]; 0,30	↔

¹⁾ Anzahl entspricht denjenigen Personen, die für die Berechnung der jeweiligen Maßzahlen herangezogen wurden.

²⁾ RR mit zugehörigem 95%-KI und p-Wert basieren auf logistischen Regressionen adjustiert nach Glukokortikoid-Behandlungsregime und Alter zu Baseline. Wenn das Modell nicht konvergiert, wurde die Wald-Methode verwendet.

³⁾ Weitere Ergebnisse zur Sicherheit (UE nach Organsystemen und Einzelereignisse) sind in Kapitel 3.5 dargestellt.

Abkürzungen: CTCAE: Common Terminology Criteria for Adverse Events; EOS: End of Study; KI: Konfidenzintervall; RR: Relatives Risiko; (S)UE: (Schwerwiegendes) Unerwünschtes Ereignis.

Referenzen

1. **Allen DD, Gorton GE, Oeffinger DJ, Tylkowski C, Tucker CA, Haley SM.** Analysis of the pediatric outcomes data collection instrument in ambulatory children with cerebral palsy using confirmatory factor analysis and item response theory methods. *J Pediatr Orthop* 2008;28(2):192-8.
2. **Arora H, Willcocks RJ, Lott DJ, Harrington AT, Senesac CR, Zilke KL, et al.** Longitudinal timed function tests in Duchenne muscular dystrophy: ImagingDMD cohort natural history. *Muscle Nerve* 2018;58(5):631-638.
3. **Daltroy LH, Liang MH, Fossel AH, Goldberg MJ.** The POSNA pediatric musculoskeletal functional health questionnaire: report on reliability, validity, and sensitivity to change. Pediatric Outcomes Instrument Development Group. *Pediatric Orthopaedic Society of North America. J Pediatr Orthop* 1998;18(5):561-71.
4. **European Medicines Agency (EMA).** Duvyzat (givinostat): European public assessment report EMEA/H/C/006079/0000 [online]. 25.04.2025. Amsterdam (NED): EMA. [Zugriff: 17.07.2025]. URL: https://www.ema.europa.eu/en/documents/assessment-report/duvyzat-epar-public-assessment-report_en.pdf.
5. **European Medicines Agency (EMA).** Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy [online]. 17.12.2015. Amsterdam (NED): EMA. [Zugriff: 10.09.2025]. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-duchenne-and-becker-muscular-dystrophy_en.pdf.
6. **Gemeinsamer Bundesausschuss (G-BA).** Amendment zur 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; Vamorolon; Dossierbewertung vom 15. April 2024 [online]. Berlin (GER): G-BA; 2024. [Zugriff: 04.08.2025]. URL: https://www.g-ba.de/downloads/92-975-7600/2024-07-04_Amendment-G-BA_Vamorolon_D-1037.pdf.
7. **Gemeinsamer Bundesausschuss (G-BA).** Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie: Anlage XII; Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V); Vamorolon (Duchenne-Muskeldystrophie, ≥ 4 Jahre) [online]. Berlin (GER): G-BA; 2024. [Zugriff: 04.08.2025]. URL: https://www.g-ba.de/downloads/39-261-6703/2024-07-04_AM-RL-XII_Vamorolon_D-1037_BAnz.pdf.
8. **Gemeinsamer Bundesausschuss (G-BA).** 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; Vamorolon; Behandlung von Duchenne-Muskeldystrophie bei Patientinnen und Patienten ab 4 Jahren [online]. Berlin (GER): G-BA; 2024. [Zugriff: 04.08.2025]. URL: https://www.g-ba.de/downloads/92-975-7435/2024-01-15_Nutzenbewertung-G-BA_Vamorolon_D-1037.pdf.

9. **Gemeinsamer Bundesausschuss (G-BA).** Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie: Anlage XII; Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V); Cipaglucosidase alfa (Morbus Pompe, Kombination mit Miglustat) vom 1. Februar 2024 [online]. Berlin (GER): G-BA; 2024. [Zugriff: 01.10.2025]. URL: https://www.g-ba.de/downloads/40-268-10198/2024-02-01_AM-RL-XII_Cipaglucosidase-alfa_D-964_TrG.pdf.
10. **Gemeinsamer Bundesausschuss (G-BA).** Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie: Anlage XII; Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a des Fünften Buches Sozialgesetzbuch (SGB V); Vamorolon (Duchenne-Muskeldystrophie, ≥ 4 Jahren) [online]. Berlin (GER): G-BA; 2024. [Zugriff: 04.08.2025]. URL: https://www.g-ba.de/downloads/40-268-10615/2024-07-04_AM-RL-XII_Vamorolon_D-1037_TrG.pdf.
11. **Henricson E, Abresch R, Han JJ, Nicorici A, Goude Keller E, de Bie E, et al.** The 6-minute walk test and person-reported outcomes in boys with Duchenne muscular dystrophy and typically developing controls: longitudinal comparisons and clinically-meaningful changes over one year. PLoS Curr 2013;5.
12. **Italfarmaco.** Integrated summary of efficacy (ISE) for givinostat in Duchenne muscular dystrophy (DMD); statistical analysis plan [unveröffentlicht]. 07.11.2022.
13. **Italfarmaco.** A randomised, double blind, placebo controlled, multicentre study to evaluate the efficacy and safety of givinostat in ambulant patients with Duchenne muscular dystrophy (DMD); DSC/14/2357/48; final clinical study protocol [unveröffentlicht]. 07.05.2016.
14. **Italfarmaco.** Randomised, double blind, placebo controlled, multicentre study to evaluate the efficacy and safety of givinostat in ambulant patients with Duchenne muscular dystrophy (DMD): clinical evaluator manual; DSC/14/2357/48; manual [unveröffentlicht]. 03.02.2021.
15. **Italfarmaco.** Randomised, double blind, placebo controlled, multicentre study to evaluate the efficacy and safety of givinostat in ambulant patients with Duchenne muscular dystrophy (DMD); DSC/14/2357/48; final clinical study report [unveröffentlicht]. 10.08.2022.
16. **Italfarmaco.** Randomised, double blind, placebo controlled, multicentre study to evaluate the efficacy and safety of givinostat in ambulant patients with Duchenne muscular dystrophy (DMD); DSC/14/2357/48; statistical analysis plan [unveröffentlicht]. 13.04.2022.
17. **Italfarmaco.** Tables and figures; DSC/14/2357/48; data [unveröffentlicht]. 2022.
18. **Italfarmaco.** Tables for ISE statistical analysis; DSC/14/2357/48; DSC/14/2357/51; statistical analysis plan tables [unveröffentlicht]. 28.08.2023.
19. **Italfarmaco.** Tables for ISE statistical analysis; DSC/14/2357/48; DSC/14/2357/51; statistical analysis plan tables [unveröffentlicht]. 30.11.2022.
20. **ITF Pharma.** Dossier zur Nutzenbewertung gemäß § 35a SGB V, Modul 4 A: Givinostat (Duvyzat), Behandlung der Duchenne-Muskeldystrophie (DMD) bei gehfähigen Patienten ab 6 Jahren zusammen mit einer Corticosteroid-Behandlung; Medizinischer Nutzen und medizinischer Zusatznutzen, Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen [unveröffentlicht]. 15.07.2025.
21. **ITF Pharma.** Duvyzat 8,86 mg/ml Suspension zum Einnehmen [online]. 07.2025. Frankfurt/Main. [Zugriff: 17.07.2025]. URL: <https://www.fachinfo.de/>.

22. **Mayhew A, Cano S, Scott E, Eagle M, Bushby K, Muntoni F, et al.** Moving towards meaningful measurement: Rasch analysis of the North Star Ambulatory Assessment in Duchenne muscular dystrophy. *Dev Med Child Neurol* 2011;53(6):535-42.
23. **Mazzone ES, Messina S, Vasco G, Main M, Eagle M, D'Amico A, et al.** Reliability of the North Star Ambulatory Assessment in a multicentric setting. *Neuromuscul Disord* 2009;19(7):458-61.
24. **McDonald CM, Henricson EK, Abresch RT, Florence J, Eagle M, Gappmaier E, et al.** The 6-minute walk test and other clinical endpoints in Duchenne muscular dystrophy: reliability, concurrent validity, and minimal clinically important differences from a multicenter study. *Muscle Nerve* 2013;48(3):357-68.
25. **McDonald CM, Henricson EK, Han JJ, Abresch RT, Nicorici A, Elfring GL, et al.** The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. *Muscle Nerve* 2010;41(4):500-10.
26. **Muntoni F, Guglieri M, Mah JK, Wagner KR, Brandsema JF, Butterfield RJ, et al.** Novel approaches to analysis of the North Star Ambulatory Assessment (NSAA) in Duchenne muscular dystrophy (DMD): observations from a phase 2 trial. *PLoS One* 2022;17(8):e0272858.
27. **Murali CN, Cuthbertson D, Slater B, Nguyen D, Turner A, Harris G, et al.** Pediatric outcomes data collection instrument is a useful patient-reported outcome measure for physical function in children with osteogenesis imperfecta. *Genet Med* 2020;22(3):581-589.
28. **Scott E, Eagle M, Mayhew A, Freeman J, Main M, Sheehan J, et al.** Development of a functional assessment scale for ambulatory boys with Duchenne muscular dystrophy. *Physiother Res Int* 2012;17(2):101-9.
29. **U.S. Food and Drug Administration (FDA).** Duvyzat (givinostat): 217865Orig1s000; integrated review [online]. 21.03.2024. Silver Spring (USA): FDA. [Zugriff: 17.07.2025]. URL: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2024/217865Orig1s000IntegratedR.pdf.

Anhang

Treppensteigen von 4 Stufen (4SC)

Die Ergebnisse sind mit denen der MMRM-Analysen ohne Imputation fehlender Werte vergleichbar, wobei für die ANCOVA ein breiteres Konfidenzintervall zu beobachten ist (LS-Mean-Differenz: ANCOVA: -2,66 (95%-KI: [-4,18; -1,13]); MMRM: -0,91 (95%-KI: [-1,50; -0,31])).

Im Placebo-Arm wird für das LS Mean der Veränderung zu Woche 60 geringfügig höhere Werte bei den ANCOVA mittels imputierter Daten gegenüber MMRM anhand beobachteter Werte (LS Mean: ANCOVA: 3,37 (95%-KI: [2,15; 4,61]); MMRM: 1,75 (95%-KI: [1,27; 2,24])) beobachtet. Im Givinostat-Arm sind für beide Analysen die Werte zum LS Mean der Veränderung zu Woche 60 vergleichbar (LS Mean: ANCOVA: 0,72 (95%-KI: [-0,16; 1,60]); MMRM: 0,85 (95%-KI: [0,50; 1,19])).

*Tabelle 23: 4SC (Zeit in Sekunden), ANCOVA anhand imputierter Werte zu Woche 60;
Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)*

Studie EPIDYS 4SC (Zeit in Sekunden) ¹⁾	Givinostat N = 118	Placebo N = 61
Baseline n (%) ²⁾³⁾ MW (SD)	118 (100) 3,58 (1,25)	61 (100) 3,60 (1,27)
Woche 60 ⁴⁾ n (%) ³⁾⁵⁾ MW (SD)	118 (100) k. A.	61 (100) k. A.
Veränderung zu Woche 60 ⁴⁾ n (%) ³⁾⁵⁾ LS Mean [95%-KI] ⁶⁾	118 (100) 0,72 [-0,16; 1,60]	61 (100) 3,37 [2,15; 4,61]
Veränderung zu Woche 60 LS-Mean-Differenz [95%-KI] ⁶⁾ p-Wert ⁶⁾	-2,66 [-4,18; -1,13] 0,0007	

¹⁾ Primärer Endpunkt.

²⁾ Für eine Person im Interventionsarm sind keine Werte verfügbar.

³⁾ Imputierte Daten. Fehlende Werte der Klasse 1 (andere Gründe, COVID) wurden imputiert als Mittelwert aller nicht-fehlenden Werte des Endpunkts des jeweiligen Zeitpunkts aller Patienten des zugehörigen Behandlungsarms, des gleichen Strata (Glukokortikoid-Behandlungsregime) und der Subpopulation (On- oder Off-Target-Population). Fehlende Werte der Klasse 2 (Gehunfähigkeit, körperliche Einschränkungen) wurden als Non-Responder imputiert (entweder als „0“ oder als doppelter Maximalwert der aufgezeichneten nicht-fehlenden Werte aller Patienten).

⁴⁾ Ergebnisse zum Studienende (Visite zu Woche 72 / EOS) werden nicht dargestellt. Es ist unklar, inwiefern Studienabbrüche in die Analyse eingingen und inwiefern Patienten mit Erhebung außerhalb des erweiterten Visitenfensters (Woche 72 ± 2 Monate) berücksichtigt wurden.

⁵⁾ Verfügbare Werte zu Woche 60: Givinostat: n = 97 (82,2 %); Placebo: n = 48 (78,7 %). Es liegen keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor. Fehlende Werte sind nicht allein durch Studienabbrüche (bis Studienende in Woche 72 / EOS: Givinostat: n = 7 (5,9 %); Placebo: n = 2 (3,3 %)) erklärbar.

⁶⁾ ANCOVA anhand imputierter Daten zu Woche 60, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT, 6-MWT und Patientenalter zum Studienstart und dem Stratifizierungsfaktor „Glukokortikoid-Behandlungsregime“.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; ANCOVA: Kovarianzanalyse; EOS: End of Study; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MW: Mittelwert; RFTF: Rise From The Floor; SD: Standardabweichung.

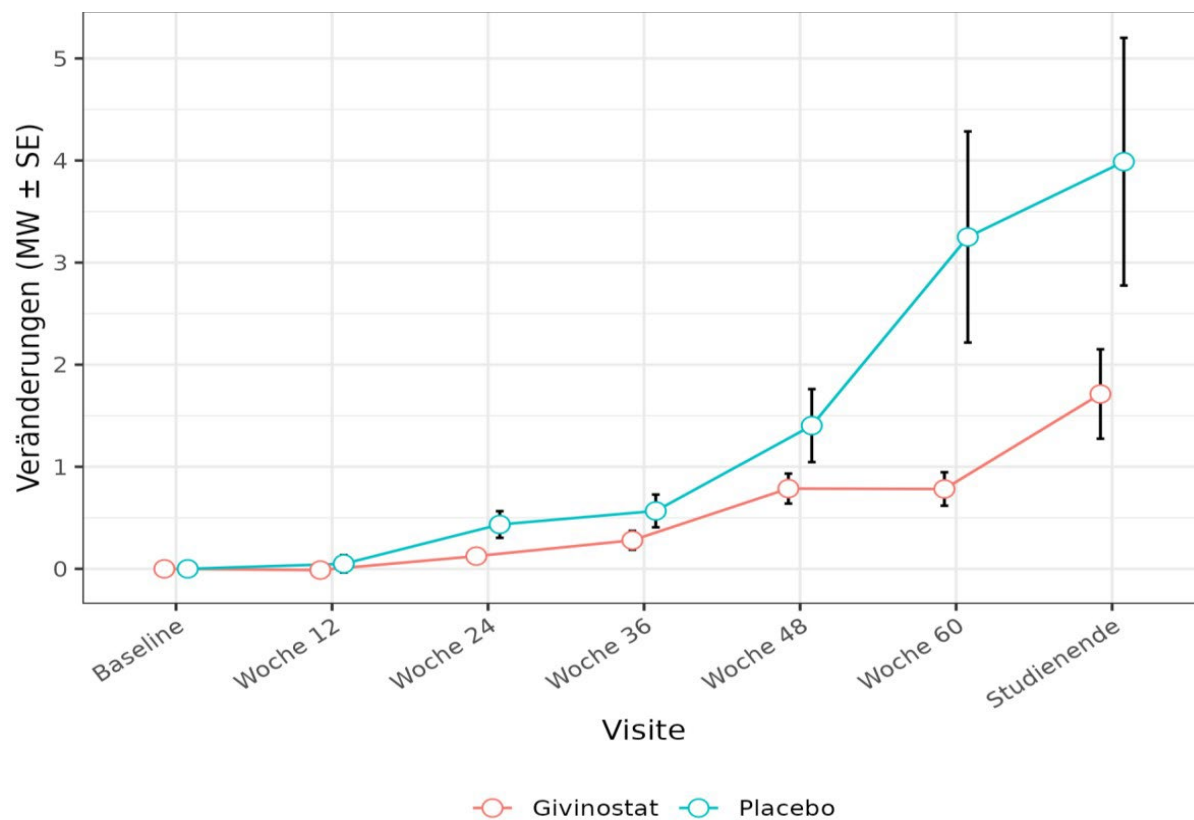


Abbildung 6: 4SC (Zeit in Sekunden) bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]

Aufstehen aus der Rückenlage (RFTF)

Es werden die Ergebnisse zum Endpunkt „Aufstehen aus der Rückenlage“ mittels ANCOVA anhand imputierter Werte zu Woche 48 der Overall-Population dargestellt. Aufgrund eines geringen Rücklaufs (< 70 %) werden die Ergebnisse zu Woche 60 nicht berücksichtigt.

Bei der ANCOVA mittels imputierter Daten und den MMRM-Analysen ohne Imputation fehlender Werte zeigt sich kein statistisch signifikanter Unterschied zu Woche 48 (LS-Mean-Differenz: ANCOVA: -3,05 (95%-KI: [-8,30; 2,20]); MMRM: -0,57 (95%-KI: [-1,59; 0,44])). Bei den ANCOVA kann für den Effektschätzer ein breiteres Konfidenzintervall beobachtet werden. Bei beiden Behandlungsgruppen werden für die ANCOVA, im Vergleich zu den MMRM-Analysen, höhere Werte für die Veränderung zu Woche 48 beobachtet (Givinostat, LS Mean: ANCOVA: 5,94 (95%-KI: [2,92; 8,96]); MMRM: 1,36 (95%-KI: [0,78; 1,95]) // Placebo, LS Mean: ANCOVA: 8,98 (95%-KI: [4,77; 13,2]); MMRM: 1,94 (95%-KI: [1,12; 2,76])).

*Tabelle 24: RFTF (Zeit in Sekunden), ANCOVA anhand imputierter Werte zu Woche 48;
Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)*

Studie EPIDYS RFTF (Zeit in Sekunden)	Givinostat N = 118	Placebo N = 61
Baseline n (%) ¹⁾²⁾ MW (SD)	118 (100) 6,89 (7,43)	61 (100) 6,76 (7,31)
Woche 48 ³⁾ n (%) ¹⁾⁴⁾ MW (SD)	118 (100) k. A.	61 (100) k. A.
Veränderung zu Woche 48 ³⁾ n (%) ¹⁾⁴⁾ LS Mean [95%-KI] ⁵⁾	118 (100) 5,94 [2,92; 8,96]	61 (100) 8,98 [4,77; 13,2]
Veränderung zu Woche 48 LS-Mean-Differenz [95%-KI] ⁵⁾ p-Wert ⁵⁾	-3,05 [-8,30; 2,20] 0,25	

¹⁾ Imputierte Daten. Fehlende Werte der Klasse 1 (andere Gründe, COVID) wurden imputiert als Mittelwert aller nicht-fehlenden Werte des Endpunkts des jeweiligen Zeitpunkts aller Patienten des zugehörigen Behandlungsarms, des gleichen Strata (Glukokortikoid-Behandlungsregime) und der Subpopulation (On- oder Off-Target-Population). Fehlende Werte der Klasse 2 (Gehunfähigkeit, körperliche Einschränkungen) wurden als Non-Responder imputiert (entweder als „0“ oder als doppelter Maximalwert aller aufgezeichneten nicht-fehlenden Werte aller Patienten).

²⁾ Für eine Person im Givinostat-Arm war der Baseline-Wert nicht verfügbar.

³⁾ Ergebnisse zu Woche 60 (Rücklauf < 70 %) und zum Studienende (Visite zu Woche 72 / EOS) werden nicht dargestellt. Zur Visite zum Studienende (Woche 72 / EOS) ist unklar, inwiefern Studienabbrüche in die Analyse eingingen und inwiefern Patienten mit Erhebung außerhalb des erweiterten Visitenfensters (Woche 72 ± 2 Monate) berücksichtigt wurden.

⁴⁾ Verfügbare Werte zu Woche 48: Givinostat: n = 84 (71,2 %); Placebo: n = 43 (70,5 %). Es liegen zu Woche 48 keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor. Fehlende Werte sind nicht allein durch Studienabbrüche (bis Studienende in Woche 72 / EOS: Givinostat: n = 7 (5,9 %); Placebo: n = 2 (3,3 %)) erklärbar.

⁵⁾ ANCOVA anhand imputierter Daten zu Woche 48, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT, 6-MWT und Patientenalter zum Studienstart und dem Stratifizierungsfaktor „Glukokortikoid-Behandlungsregime“.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; ANCOVA: Kovarianzanalyse; EOS: End of Study; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MW: Mittelwert; RFTF: Rise From The Floor; SD: Standardabweichung.

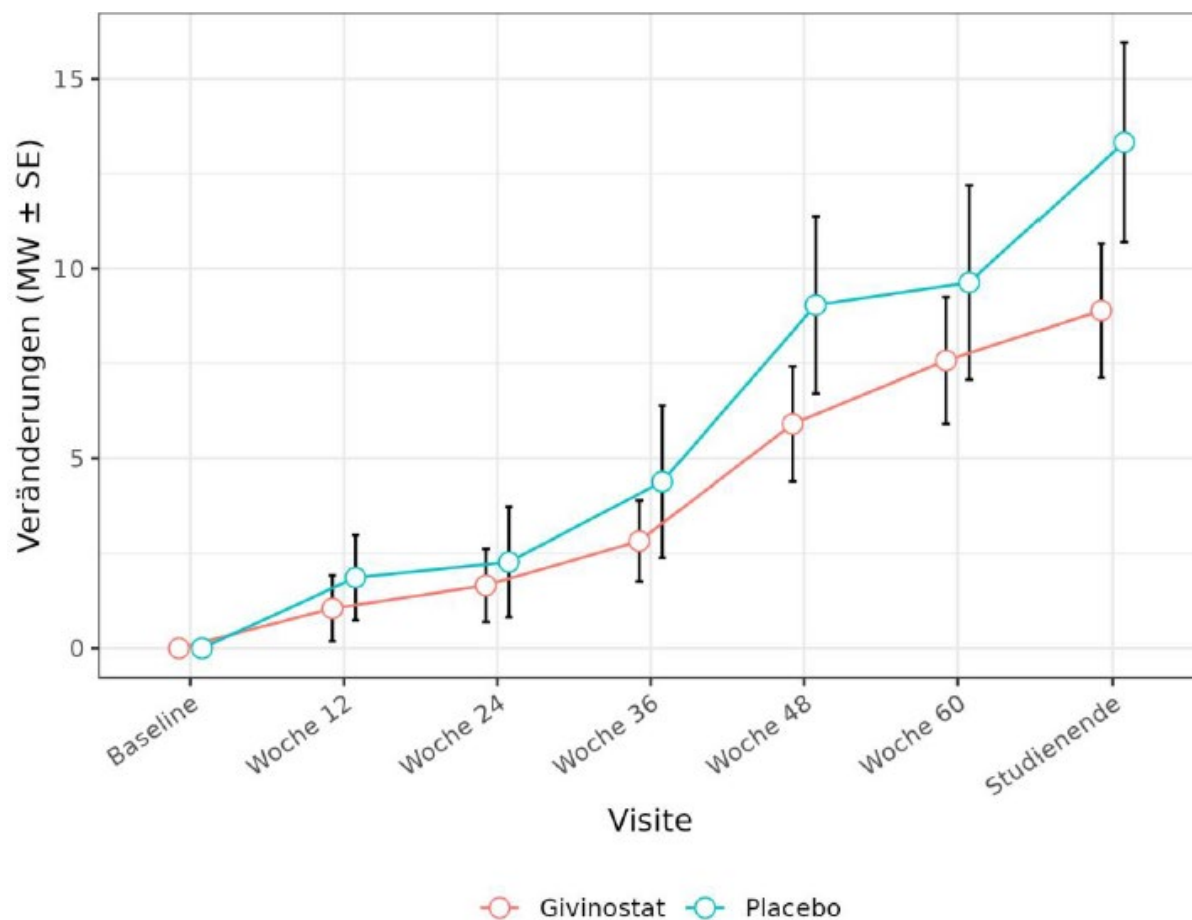


Abbildung 7: RFTF (Zeit in Sekunden) bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]

10-Meter-Geh-/Lauftest (10-MWT)

Anders als bei der Analyse mittels MMRM anhand beobachteter Werte (LS-Mean-Differenz: 0,04 (95%-KI: [-0,43; 0,52])) konnte für den Endpunkt mittels ANCOVA anhand imputierter Daten ein statistisch signifikanter Effekt (LS-Mean-Differenz: -1,82 (95%-KI: [-3,62; -0,03])) beobachtet werden.

Im Placebo-Arm liegen für die Veränderung zu Woche 60 höhere Werte bei den ANCOVA mittels imputierter Daten gegenüber MMRM anhand beobachteter Werte vor (LS Mean: ANCOVA: 2,67 (95%-KI: [1,22; 4,12]); MMRM: 0,89 (95%-KI: [0,50; 1,28])).

Im Givinostat-Arm werden bei beiden Analysen vergleichbare Werte bei der Veränderung zu Woche 60 beobachtet (LS Mean: ANCOVA: 0,84 (95%-KI: [-0,19; 1,88]); MMRM: 0,93 (95%-KI: [0,66; 1,21])).

*Tabelle 25: 10-MWT (Zeit in Sekunden), ANCOVA anhand imputierter Werte zu Woche 60;
Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)*

Studie EPIDYS 10-MWT (Zeit in Sekunden)	Givinostat N = 118	Placebo N = 61
Baseline n (%) ¹⁾² MW (SD)	118 (100) 5,56 (1,34)	61 (100) 5,30 (1,03)
Woche 60 ³⁾ n (%) ¹⁾⁴⁾ MW (SD)	118 (100) k. A.	61 (100) k. A.
Veränderung zu Woche 60 ³⁾ n (%) ¹⁾⁴⁾ LS Mean [95%-KI] ⁵⁾	118 (100) 0,84 [-0,19; 1,88]	61 (100) 2,67 [1,22; 4,12]
Veränderung zu Woche 60 LS-Mean-Differenz [95%-KI] ⁵⁾ p-Wert ⁵⁾	-1,82 [-3,62; -0,03] 0,047	

¹⁾ Imputierte Daten. Fehlende Werte der Klasse 1 (andere Gründe, COVID) wurden imputiert als Mittelwert aller nicht-fehlenden Werte des Endpunkts des jeweiligen Zeitpunkts aller Patienten des zugehörigen Behandlungsarms, des gleichen Strata (Glukokortikoid-Behandlungsregime) und der Subpopulation (On- oder Off-Target-Population). Fehlende Werte der Klasse 2 (Gehunfähigkeit, körperliche Einschränkungen) wurden als Non-Responder imputiert (entweder als „0“ oder als doppelter Maximalwert aller aufgezeichneten nicht-fehlenden Werte aller Patienten).

²⁾ Für eine Person im Givinostat-Arm war der Baseline-Wert nicht verfügbar.

³⁾ Ergebnisse zum Studienende (Visite zu Woche 72 / EOS) werden nicht dargestellt. Es ist unklar, inwiefern Studienabbrüche in die Analyse zum Studienende eingingen und inwiefern Patienten mit Erhebung im erweiterten Visitenfenster (Woche 72 ± 2 Monate) berücksichtigt wurden.

⁴⁾ Verfügbare Werte zu Woche 60: Givinostat: n = 98 (83,1 %); Placebo: n = 48 (78,7 %). Es liegen zu Woche 60 keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor. Fehlende Werte sind nicht allein durch Studienabbrüche (bis Studienende in Woche 72 / EOS: Givinostat: n = 7 (5,9 %); Placebo: n = 2 (3,3 %)) erklärbar.

⁵⁾ ANCOVA anhand imputierter Daten zu Woche 60, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT, 6-MWT und Patientenalter zum Studienstart und dem Stratifizierungsfaktor „Glukokortikoid-Behandlungsregime“.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; ANCOVA: Kovarianzanalyse; EOS: End of Study; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MW: Mittelwert; RFTF: Rise From The Floor; SD: Standardabweichung.

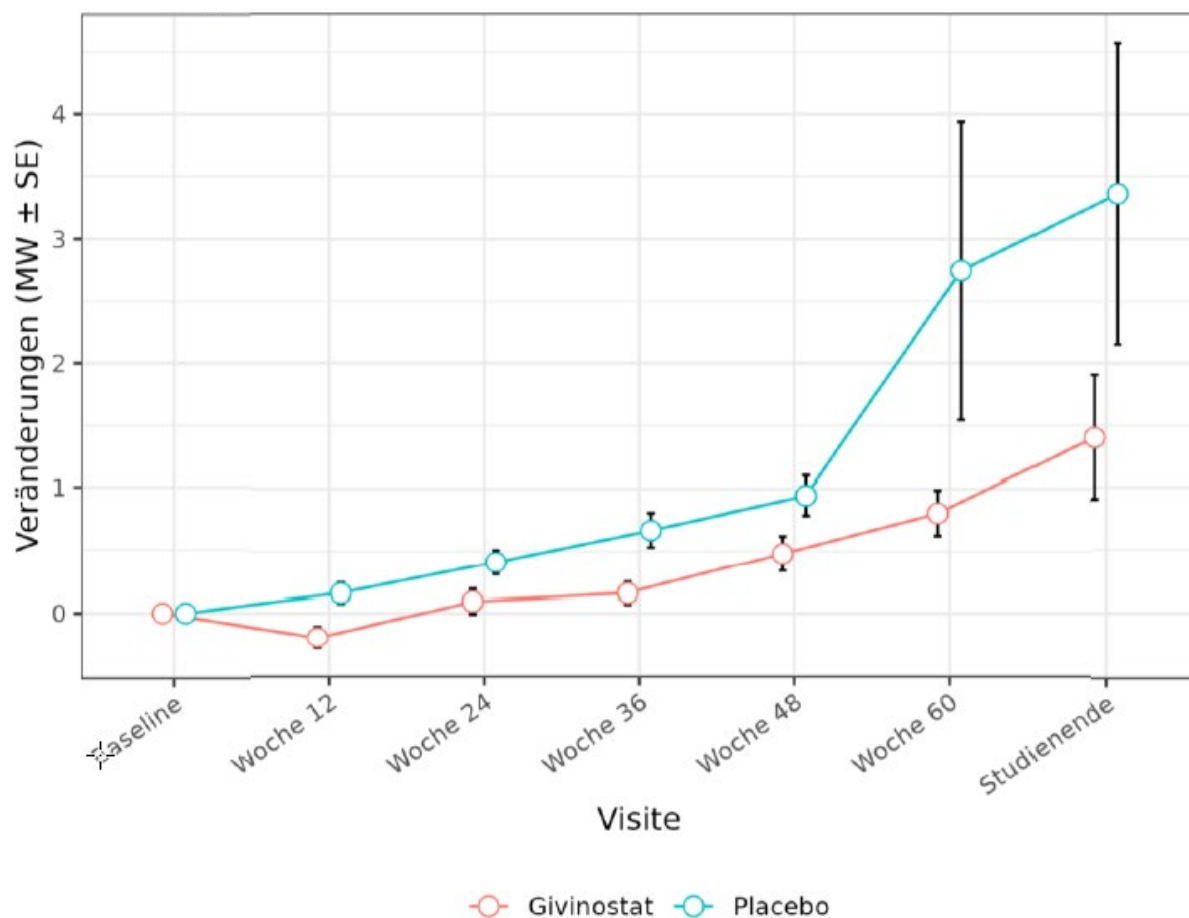


Abbildung 8: 10-MWT (Zeit in Sekunden) bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]

6-Minuten-Gehtest (6-MWT)

Anders als bei der Analyse mittels MMRM anhand beobachteter Werte (LS-Mean-Differenz: 4,39 (95%-KI: [-11,5; 20,2])) konnten für den Endpunkt mittels ANCOVA anhand imputierter Daten ein statistisch signifikanter Effekt (LS-Mean-Differenz: 24,9 (95%-KI: [3,12; 46,7])) beobachtet werden.

Für den Placebo-Arm wird bei der ANCOVA mittels imputierter Werte gegenüber der MMRM-Analyse eine größere Veränderung zu Woche 60 beobachtet (LS Mean: ANCOVA: -61,1 (95%-KI: [-78,7; -43,6]); MMRM: -47,0 (95%-KI: [-59,8; -34,1])). Im Givinostat-Arm werden bei beiden Analysen vergleichbare Werte für die Veränderung zu Woche 60 beobachtet (LS Mean: ANCOVA: -36,2 (95%-KI: [-48,7; -23,7]); MMRM: 42,6 (95%-KI: [-51,70; -33,52])).

Tabelle 26: 6-MWT (Distanz in Metern), ANCOVA anhand imputierter Werte zu Woche 60; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS 6-MWT (Distanz in Metern)	Givinostat N = 118	Placebo N = 61
Baseline n (%) ¹⁾²⁾ MW (SD)	118 (100) 398,3 (70,9)	61 (100) 393,7 (61,4)
Woche 60 ³⁾ n (%) ¹⁾⁴⁾ MW (SD)	118 (100) k. A.	61 (100) k. A.
Veränderung zu Woche 60 ³⁾ n (%) ¹⁾⁴⁾ LS Mean [95%-KI] ⁵⁾	118 (100) -36,2 [-48,7; -23,69]	61 (100) -61,1 [-78,7; -43,]
Veränderung zu Woche 60 LS-Mean-Differenz [95%-KI] ⁵⁾ p-Wert ⁵⁾	24,9 [3,12; 46,7] 0,025	

¹⁾ Imputierte Daten. Fehlende Werte der Klasse 1 (andere Gründe, COVID) wurden imputiert als Mittelwert aller nicht-fehlenden Werte des Endpunkts des jeweiligen Zeitpunkts aller Patienten des zugehörigen Behandlungsarms, des gleichen Strata (Glukokortikoid-Behandlungsregime) und der Subpopulation (On- oder Off-Target-Population). Fehlende Werte der Klasse 2 (Gehunfähigkeit, körperliche Einschränkungen) wurden als Non-Responder imputiert (entweder als „0“ oder als doppelter Maximalwert aller aufgezeichneten nicht-fehlenden Werte aller Patienten).

²⁾ Für eine Person im Givinostat-Arm war der Baseline-Wert nicht verfügbar.

³⁾ Ergebnisse zum Studienende (Visite zu Woche 72 / EOS) werden nicht dargestellt. Es ist unklar, inwiefern Studienabbrüche in die Analyse zum Studienende eingingen und inwiefern Patienten mit Erhebung im erweiterten Visitenfenster (Woche 72 ± 2 Monate) berücksichtigt wurden.

⁴⁾ Verfügbare Werte zu Woche 60: Givinostat: n = 97 (82,2 %); Placebo: n = 48 (78,7 %). Es liegen zu Woche 60 keine Angaben zur Anzahl an fehlenden Werten differenziert nach Gründen vor. Fehlende Werte sind nicht allein durch Studienabbrüche (bis Studienende in Woche 72 / EOS: Givinostat: n = 7 (5,9 %); Placebo: n = 2 (3,3 %)) erklärbar.

⁵⁾ ANCOVA anhand imputierter Daten zu Woche 60, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT, 6-MWT und Patientenalter zum Studienstart und dem Stratifizierungsfaktor „Glukokortikoid-Behandlungsregime“.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; ANCOVA: Kovarianzanalyse; EOS: End of Study; k. A.: keine Angabe; KI: Konfidenzintervall; LS: Least Squares; MW: Mittelwert; RFTF: Rise From The Floor; SD: Standardabweichung.

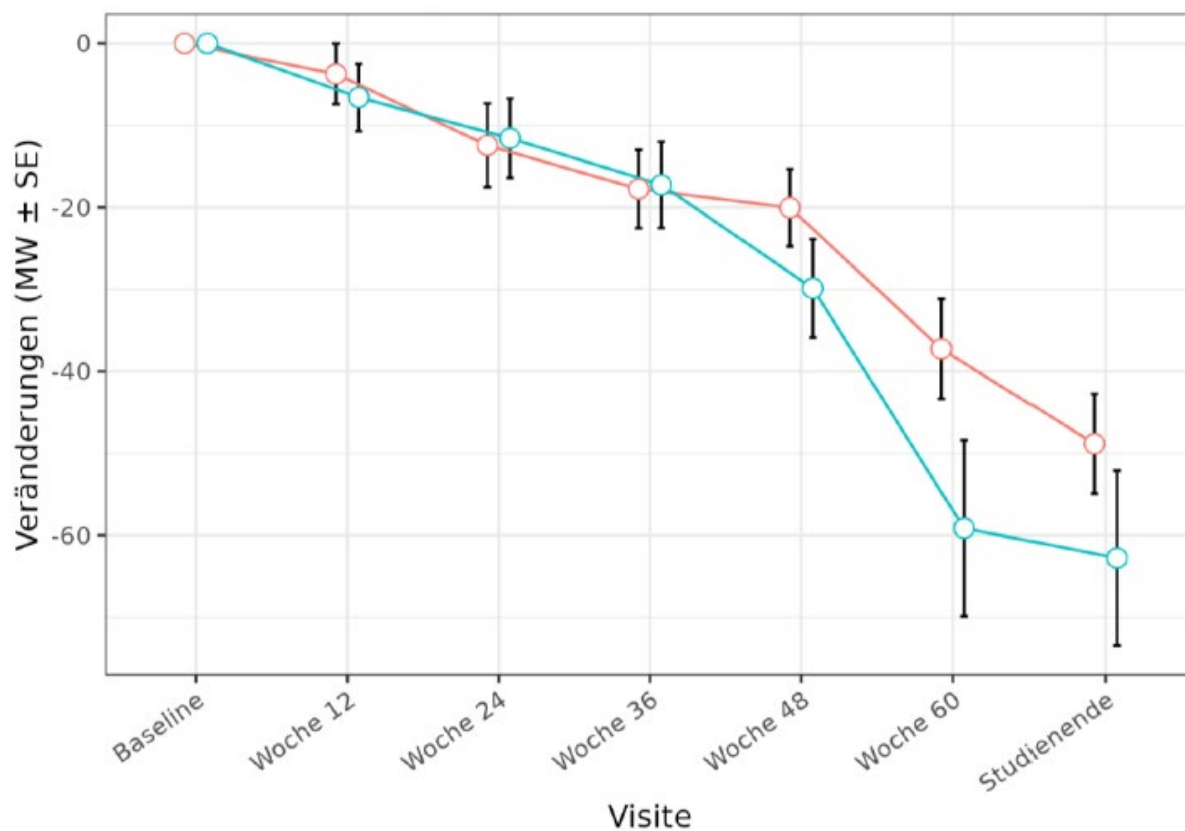


Abbildung 9: 6-MWT (Distanz in Meter) bis Studienende (Woche 72 / EOS); Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022) [20]