

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**

Wirkstoff: Givinostat

Dossierbewertung vom 15. Oktober 2025

Datum des Amendments: 12. Dezember 2025

Inhaltsverzeichnis

Tabellenverzeichnis	3
Abkürzungsverzeichnis	4
1 Hintergrund.....	5
2 Fragestellung.....	6
3 Liste der verwendeten Quellen.....	6
3.1 Tabellarische Übersicht der nachgereichten Daten des pU.....	6
4 Methodische Aspekte.....	7
4.1 Morbidität	7
4.2 Sicherheit	8
4.3 Statistische Auswertung	9
4.4 Verzerrungspotential auf Studien- und Endpunktebene	12
5 Ergebnisse	14
5.1 Morbidität	14
5.2 Sicherheit	19
6 Zusammenfassung.....	22
Referenzen	25

Tabellenverzeichnis

Tabelle 1: Übersicht der nachgereichten und bereits in der bisherigen Nutzenbewertung berücksichtigten Daten	6
Tabelle 2: Verzerrungspotential der im Amendment berücksichtigten und ergänzend dargestellten Endpunkte der Studie EPIDYS.....	12
Tabelle 3: 4SC (Zeit in Sekunden), MMRM-Analysen (ohne Imputation) zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	14
Tabelle 4: RFTF (Zeit in Sekunden), MMRM-Analysen (ohne Imputation) zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	15
Tabelle 5: 10-MWT (Zeit in Sekunden), MMRM-Analysen (ohne Imputation) zu Woche 72/EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	16
Tabelle 6: 6-MWT (Distanz in Metern), MMRM-Analysen (ohne Imputation) zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	17
Tabelle 7: NSAA-Gesamtscore (Ordinalskala: 0–34 Punkte), MMRM-Analysen zu Woche 72/EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	18
Tabelle 8: Körpergröße (z-Scores), MMRM-Analysen zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	19
Tabelle 9: Körpergewicht (z-Scores), MMRM-Analysen zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022).....	20

Abkürzungsverzeichnis

4SC	4-Stair Climb
6-MWT	6-Minute Walk Test
10-MWT	10-Meter Walk/Run Test
ANCOVA	Kovarianzanalyse (Analysis of Covariance)
DMD	Duchenne-Muskeldystrophie
EOS	End of Study
G-BA	Gemeinsamer Bundesausschuss
KI	Konfidenzintervall
LS	Least Squares
MAR	Missing at Random
MMRM	Mixed Model for Repeated Measures
N	Anzahl
NSAA	North Star Ambulatory Assessment
RFTF	Rise From The Floor
pU	pharmazeutischer Unternehmer
SD	Standardabweichung
SGB	Sozialgesetzbuch
VerfO	Verfahrensordnung des G-BA

1 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.

In die Bewertung werden gemäß Zulassungsstatus Patienten in folgender Indikation eingeschlossen: Behandlung der Duchenne-Muskeldystrophie (DMD) bei gehfähigen Patienten ab 6 Jahren zusammen mit einer Corticosteroid-Behandlung.

Der pharmazeutische Unternehmer (pU) hatte am 14. Juli 2025 ein Dossier zur Nutzenbewertung gemäß § 35a SGB V beim Gemeinsamen Bundesausschuss (G-BA) eingereicht. Als bewertungsrelevante Studie wurde die pivotale Studie EPIDYS (Studie 48) für die Nutzenbewertung herangezogen, 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 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 Nutzenbewertung wurde am 15. Oktober 2025 zusammen mit der Bewertung der Therapiekosten und Patientenzahlen des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) auf der Internetseite des G-BA (<http://www.g-ba.de>) veröffentlicht und damit das schriftliche Stellungnahmeverfahren eingeleitet. Eine mündliche Anhörung fand am 24. November 2025 statt, in der offene Fragen zu den in den schriftlichen Stellungnahmen dargelegten Daten und Argumentationen erörtert wurden.

Im Rahmen der Nutzenbewertung zu Givinostat wurden für die Endpunkte „Treppensteigen von 4 Stufen“ (mittels 4SC), „Aufstehen aus der Rückenlage“ (mittels RFTF), „10-Meter-Geh-/Lauftest“ (mittels 10-MWT) und „6-Minuten-Gehtest“ (mittels 6-MWT) die post hoc durchgeföhrten MMRM-Analysen ohne Imputation zu Woche 60 aufgrund Limitationen in der statistischen Auswertung ergänzend dargestellt.

Die Endpunkte „Körpergröße“ und „Körpergewicht“ wurden in der Nutzenbewertung aufgrund unklarer Validität (keine Berechnung von alters- und (geschlechts-)adjustierten z-Scores) nicht berücksichtigt. Der Endpunkt „Funktionelle Leistungsfähigkeit“ (mittels NSAA) wurde aufgrund unklarer Validität und fehlender Daten zur Änderungssensitivität der Ordinalskala nicht in der Nutzenbewertung dargestellt.

Der pU hat mit der schriftlichen Stellungnahme Informationen zur Operationalisierung der Studienvisiten, Daten bzw. Analysen für den primären Endpunkt „Treppensteigen von 4 Stufen“ zu fehlenden Werten differenziert nach Klasse 1 und 2, Sensitivitätsanalysen mittels MMRM mit dem prädefinierten Analysezeitfenster (± 7 Tage) nachgereicht. Für die Endpunkte „Körpergröße“ und „Körpergewicht“ wurden MMRM-Analysen für z-Scores und für den Endpunkt „Funktionelle Leistungsfähigkeit“ Publikationen zur Reliabilität, Validität, Änderungssensitivität und MMRM-Analysen anhand der Rasch-transformierten linearisierten Skala eingereicht [9].

Zusätzlich reichte der pU mit der Stellungnahme nach der mündlichen Anhörung die entsprechenden Daten und Analysen analog zum Endpunkt „Treppensteigen von 4 Stufen“ für die Endpunkte „Aufstehen aus der Rückenlage“, „10-Meter-Geh-/Lauftest“ und „6-Minuten-Gehtest“ nach [10].

Vor diesem Hintergrund bat der Unterausschuss Arzneimittel um eine zusammenfassende Darstellung und Bewertung der nachgereichten Daten zu den Endpunkten „Treppensteigen von 4 Stufen“, „Aufstehen aus der Rückenlage“, „10-Meter-Geh-/Lauftest“, „6-Minuten-Gehtest“, „Funktionelle Leistungsfähigkeit“, „Körpergröße (z-Scores)“ und „Körpergewicht (z-Scores)“.

2 Fragestellung

Mit diesem Amendment werden die am 05.11.2025 durch den pU im Rahmen des schriftlichen Stellungnahmeverfahrens und am 27.11.2025 im Nachgang zur mündlichen Anhörung vom 24.11.2025 eingereichten Analysen sowie die mit dem Herstellerdossier eingereichten Analysen der Endpunkte „Treppensteigen von 4 Stufen“, „Aufstehen aus der Rückenlage“, „10-Meter-Geh-/Lauftest“, „6-Minuten-Gehtest“, „Funktionelle Leistungsfähigkeit“, „Körpergröße (z-Scores)“ und „Körpergewicht (z-Scores)“ bewertet und dargestellt.

3 Liste der verwendeten Quellen

Zur Bestimmung des Ausmaßes des Zusatznutzens für Givinostat wurden folgende Quellen für dieses Amendment herangezogen:

- Herstellerdossier [8]
- Nutzenbewertung zu Givinostat [3]
- Studienunterlagen Studie EPIDYS (Studie 48) [4,5,6,7]
- Schriftliche Stellungnahme des pU vom 05.11.2025 [9] und vom 27.11.2025 [10]

3.1 Tabellarische Übersicht der nachgereichten Daten des pU

Tabelle 1: Übersicht der nachgereichten und bereits in der bisherigen Nutzenbewertung berücksichtigten Daten

Endpunktategorie Messinstrument/ Endpunkt	Nachgereichte Daten aus dem Stellungnahmeverfahren vom 05.11.2025 und 27.11.2025	Daten aus dem Herstellerdossier
Morbidität 4SC, RFTF, 10-MWT, 6-MWT	<ul style="list-style-type: none"> • Angaben fehlender Werte differenziert nach Klasse 1 und 2. • Sensitivitätsanalysen mittels MMRM ohne Imputation mit dem prädefinierten Analysezeitfenster (± 7 Tage). 	MMRM-Analysen ohne Imputation zu Woche 72 / EOS.
Morbidität NSAA (Gesamtscore)	Sensitivitätsanalysen mittels MMRM mit dem prädefinierten Analysezeitfenster (± 7 Tage) der Ordinalskala (0–34 Punkte).	MMRM-Analysen zu Woche 72 / EOS der Ordinalskala (0–34 Punkte).
Sicherheit Körpergröße (z-Scores), Körpergewicht (z-Scores)	<ul style="list-style-type: none"> • MMRM-Analysen ohne Imputation zu den z-Scores anhand beobachteter Werte zu Woche 72 / EOS. • Sensitivitätsanalysen mittels MMRM ohne Imputation mit dem prädefinierten Analysezeitfenster (± 7 Tage). 	-

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; MMRM: Mixed Model for Repeated Measures; NSAA: North Star Ambulatory Assessment; RFTF: Rise From The Floor.

4 Methodische Aspekte

In der Nutzenbewertung vom 15. Oktober 2025 wurden die post hoc durchgeführten MMRM-Analysen für die Endpunkte „Treppensteigen von 4 Stufen“ (erhoben mittels 4SC), „Aufstehen aus der Rückenlage“ (erhoben mittels RFTF), „10-Meter-Geh-/Lauftest“ (erhoben mittels 10-MWT) und „6-Minuten-Gehtest“ (erhoben mittels 6-MWT) aufgrund von Limitationen in der statistischen Auswertung nur ergänzend dargestellt. Es lagen keine Angaben zur Anzahl fehlender Werte der Klasse 1 und Klasse 2 vor. Zudem bestanden Unklarheiten in der Operationalisierung der Visitenzeitfenster zur letzten Studienvisite zu Woche 72 / EOS. Hinzu kamen Ergebnisunsicherheiten aufgrund der nicht-konsistenten Ergebnisse zwischen den im Dossier eingereichten Analysen (ANCOVA anhand imputierter Daten vs. MMRM-Analysen ohne Imputation).

Aufgrund limitierter Angaben zu Reliabilität, Validität und Änderungssensitivität (Ordinalskala) wurde der Endpunkt „Funktionelle Leistungsfähigkeit“ (erhoben mittels NSAA) in der Nutzenbewertung nicht dargestellt.

Für die Endpunkte „Körpergröße“ und „Körpergewicht“ wurden keine alters- (und geschlechts-) adjustierten z-Scores berechnet, weshalb diese nicht in der Nutzenbewertung berücksichtigt wurden.

4.1 Morbidität

Funktionelle Leistungsfähigkeit (NSAA)

Operationalisierung

Beschreibung

Der „North Star Ambulatory Assessment“ (NSAA) ist ein Instrument zur Beurteilung der Gehfähigkeit von Patienten mit DMD. Mithilfe einer 3-Punkte-Skala (0–2 Punkte) werden 17 Items bewertet und daraus ein Gesamtscore gebildet (0–34 Punkte). Zur ausführlichen Beschreibung und Bewertung der Operationalisierung wird auf die Nutzenbewertung von Givinostat verwiesen [3].

Validität

Eckdaten zu psychometrischen Eigenschaften wurden in verschiedenen Studien zur Reliabilität (Scott et al. (2012) [20], Mazzone et al. (2009) [13], Mayhew et al. (2011) [11], Ricotti et al. (2015) [19], Muntoni et al. (2019) [14], Muntoni et al. (2022) [15]) ermittelt.

In einer Studie aus England mit 191 Patienten mit DMD mit stabiler täglicher oder intermittierender Glukokortikoid-Behandlung wurden die psychometrischen Eigenschaften des NSAA mittels Rasch-Analyse untersucht. Die interne Konsistenz des Gesamtscores wurde als gut befunden, mit Ausnahme des Items „Kopfanheben in Rückenlage“. Bei diesem Item konnte keine Übereinstimmung zwischen dem Antwortverhalten und dem vorhergesagten Rasch-Modell festgestellt werden.

Für den Gesamtscore zeigt sich eine gute Reliabilität [11,17,20], wohingegen bei einzelnen Skalen („Vom Boden Aufstehen“, „Eine Stufe hochsteigen (rechts)“, „Hüpfen“ und „10 Meter laufen“) eine Heterogenität der Test-Retest-Reliabilität vorlag [20]. Weitere Studien zeigen, dass eine standardisierte und mehrfache Schulung der Bewerter/innen für eine einheitliche Erhebung notwendig ist.

Untersuchungen zur Konstruktvalidität zeigen, dass der NSAA bzgl. der abnehmenden Mobilität mit zunehmendem Alter und unterschiedlichen Glukokortikoid-Regimes unterscheiden kann (Zambon et al. (2022) [21], Ayyar Gupta et al. (2023) [1], Pane et al. (2014) [18]).

Es liegen Untersuchungen zur Änderungssensitivität anhand Longitudinaldaten für die Rasch-transformierte Skala (0–100 Punkte) [12,14,15,19] bzw. die Ordinalskala (0–34 Punkte) [1,18,19] vor. Weitere Publikationen untersuchten den kumulativen Anteil an Personen mit einem Funktionsverlust von 2 oder 1 Punkten auf 0 Punkte anhand Longitudinaldaten zum natürlichen Krankheits-

verlauf [14,15]. Die Studien zeigen für beiden Skalen des NSAA-Gesamtscores eine Änderungssensitivität an. Für die Ordinalskala (0–34 Punkte) wurde bei abnehmender Mobilität (gemessen mittels 6-MWT), zunehmendem Alter und unterschiedlichem Glukokortikoid-Regime eine von diesen Parametern abhängige Änderungssensitivität gezeigt (Pan et al (2014) [18]). Ähnliche Ergebnisse wurden für die linearisierte Skala festgestellt [14,15,19].

Zwar kann sich durch die linearisierte Skala eine einheitliche Änderung um eine bestimmte Punktzahl im Zeitverlauf gleichermaßen auf der gesamten Skala widerspiegeln, es werden in den Validierungsstudien jedoch Ergebnisse unter Ausschluss eines Items („Kopfanheben in Rückenlage“) dargestellt. Da die Auswertung der Ordinalskala präspezifiziert war, wird diese dargestellt.

Insgesamt wird der Endpunkt als weitestgehend valide angesehen. Einschränkend ist die für einzelne Items bestehende Heterogenität in der Test-Retest-Reliabilität. Der RFTF (Item 12) und der 10-MWT (Item 17) wurden im Rahmen des NSAA erhoben. Im Rahmen des NSAA wird die Art der Durchführung beurteilt, während bei den separat dargestellten Endpunkten „Aufstehen aus der Rückenlage“ und „10-Meter-Geh-/Lauftest“ die benötigte Zeit erfasst und ausgewertet wurde. Der NSAA liefert somit Informationen zur grundsätzlichen Fähigkeit und zum Grad der Selbstständigkeit der Patienten.

4.2 Sicherheit

Körpergröße (z-Scores) und Körpergewicht (z-Scores)

Die Endpunkte „Körpergröße (z-Scores)“ und „Körpergewicht (z-Scores)“ werden in der Nutzenbewertung berücksichtigt.

Mit der Stellungnahme reicht der pU Informationen zur Erhebung des Körpergewichts nach. Zusätzlich wurden post hoc durchgeführte Analysen für beide Endpunkte (z-Scores) nachgereicht.

Im Rahmen der körperlichen Untersuchung wurde das Körpergewicht erhoben. Das Wiegen sollte ohne Schuhe mit einer kalibrierten Waage erfolgen, die Erhebung musste für jede Studienvisite gleich erfolgen. Falls die Patienten im Studienverlauf ihre Gehfähigkeit verloren, konnte das Gewicht mithilfe einer Rollstuhlwaage, einem Hebelift oder einer Wiegeplattform gemessen werden.

Für die Berechnung der z-Scores wurden die Referenztabellen der „Centers for Disease Control and Prevention“ (CDC) herangezogen [2].

Die Operationalisierung wird als weitestgehend geeignet angesehen. Es bleibt unklar, ob pro Erhebungszeitpunkt lediglich eine Messung erfolgte oder ob mehrere Messungen stattfanden, die anschließend gemittelt wurden. Dennoch ist eine Erhebung von Körpergröße und Körpergewicht in der Regel weitestgehend objektiv möglich.

Die Endpunkte werden als valide eingeschätzt.

In der Patientenpopulation (gehfähige DMD-Patienten im Alter zwischen 6 und 16 Jahren) sind wesentliche Abweichungen der Körpergröße oder des Körpergewichts nicht unmittelbar eine Ausprägung bzw. ein Symptom der Erkrankung. Etwaige Abweichungen können jedoch auftreten, z. B. in Folge von Skelettdeformitäten oder aufgrund von Nebenwirkungen der begleitenden Glukokortikoid-Therapie.

4.3 Statistische Auswertung

Funktionelle Leistungsfähigkeit (NSAA)

Der pU reicht post hoc durchgeführte MMRM-Analysen des Gesamtscores für alle Visiten ein (Ordinalskala, linearisierte Skala). Es wurden Responderanalysen zu Woche 72 für eine Verbesserung bzw. Verschlechterung um mindestens 15 % der Ordinalskala (Responseseschwelle: 5,1 Punkte) sowie der linearisierten Skala (Responseseschwelle: 15 Punkte) eingereicht, adjustiert nach Alter zu Studienbeginn und den Stratifizierungsfaktoren. Für den Gesamtscore wurden fehlende Werte als Non-Responder imputiert.

Umgang mit fehlenden Werten auf Item-Ebene

Zur Bildung des NSAA-Gesamtscores (0–34 Punkte) wurden mehrere Vorgehen zur Imputation beschrieben.

Definition fehlender Werte als „Klasse 1“:

- Frage, ob der NSAA versucht wurde: „Nein“, mit der Begründung: „Durchführung des Tests abgelehnt“ oder „Andere“:
 - Imputation jedes der 17 Items erfolgte ähnlich wie bei kontinuierlichen Werten der Klasse 1 (Bildung des Modalwerts der verfügbaren Werte des jeweiligen Items nach Visite, Behandlungsgruppe, der zugehörigen Subpopulation (= On-/Off-Target) und der Randomisierungsstrata (Glukokortikoid-Behandlungsregime: „Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“)).
 - Sollten mehrere Modalwerte vorliegen, wurde der niedrigste gewählt.
 - Der Gesamtscore wird aus den 17 imputierten Items gebildet.
- Frage, ob der NSAA versucht wurde: „Ja“, aber einzelne Item-Scores (je 0–2 Punkte) fehlen:
 - Verfügbare Werte von ≥ 9 Items: Die fehlenden ≤ 8 Items werden jeweils durch Bildung des Modalwerts der verfügbaren Werte des Patienten zu der Visite imputiert. Sollte ein Modalwert nicht eindeutig bestimmbar sein (mehrere Modalwerte), wurde der niedrigste gewählt.
 - Verfügbare Werte von ≤ 8 Items: Die fehlenden ≥ 9 Items werden mit jeweils „0“ imputiert.
 - Der Gesamtscore wird in beiden Szenarien aus den imputierten und verfügbaren Werten gebildet.

Definition fehlender Werte als „Klasse 2“:

- Frage, ob der NSAA versucht wurde: „Nein“, mit der Begründung: „Nicht in der Lage den Test durchzuführen aufgrund körperlicher Einschränkungen“.
 - Imputation aller Items und des Gesamtscores mit „0“.

Einschätzung der statistischen Auswertungen

Es wurden für die Overall-Population post hoc durchgeführte MMRM-Analysen unter der Annahme „Missing at Random“ (MAR) für die Messinstrumente 4SC, RFTF, 10-MWT, 6-MWT und NSAA-Gesamtscore (Ordinalskala: 0–34 Punkte) sowie die z-Scores der Endpunkte „Körpergröße“ und „Körpergewicht“ mit dem Dossier bzw. der schriftlichen Stellungnahme des pU eingereicht.

Die MMRM wurden adjustiert nach Behandlung, Visite, Alter zum Studieneinschluss und dem Stratifizierungsfaktor „Glukokortikoid-Behandlungsregime“ („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“). Für die Messinstrumente 4SC, RFTF, 10-MWT, 6-MWT und NSAA-Gesamtscore wurde zusätzlich nach den Baseline-Werten aller „Timed Function Tests“ (4SC, RFTF, 10-MWT) und des 6-MWT adjustiert. Der Interaktionsterm Visite*Behandlung wurde für alle Analysen berücksichtigt.

Visite zu Woche 72

In der Nutzenbewertung lagen Unsicherheiten zu der Operationalisierung der Studienvisite zu Woche 72 / EOS vor (Berücksichtigung vorzeitiger Studienabbrüche, Erweiterung des Visitenzeitfensters von ursprünglich \pm 7 Tage auf zusätzlich \pm 2 Monate während der COVID-19-Pandemie). Zu Woche 60 wurde das Visitenzeitfenster von \pm 7 Tage auf zusätzliche 14 Tage erweitert.

Gemäß der nachgereichten Information des pU wurden Personen, die vorzeitig die Studie abgebrochen haben, in den Analysen zu den Wirksamkeits- und Sicherheitsendpunkten zu Woche 72 nicht berücksichtigt.

Innerhalb des prädefinierten Analysefensters (\pm 7 Tage) hatten 24 % der Patienten im Givinostat- und 33 % der Patienten im Placebo-Arm zu Woche 72 keine Visite. Im Median erfolgte die Visite zu Woche 72 nach Ablauf des Analysefensters von \pm 7 Tage nach ca. 15 Tagen in beiden Armen:

- Givinostat (N = 118), n \approx 28,3 (24 %): Median (min; max): 15 (3; 113)
- Placebo (N = 61) n \approx 20,1 (33 %): Median (min; max): 14,5 (2; 85)

Der Anteil an Personen, die nach dem erweiterten Visitenzeitfenster von 2 Monaten an der Studienvisite zu Woche 72 teilnahmen, war in beiden Behandlungsgruppen gering und in vergleichbarem Ausmaß (Givinostat: n = 6 (5,1 %); Placebo: n = 4 (6,6 %)). Die meisten Patienten, in beiden Armen insgesamt ca. 85 %, kamen innerhalb des Zeitfensters Woche 72 \pm 7 Tage + 14,5/15 Tage (= ca. 21 Tage) ins Studienzentrum, ähnlich wie zur Visite zu Woche 60 (\pm 7 Tage zuzüglich + 14 Tage = 21 Tage).

Der pU hat im Stellungnahmeverfahren bzw. im Nachgang zur mündlichen Anhörung Sensitivitätsanalysen zu den MMRM-Analysen für die Messinstrumente 4SC, RFTF, 10-MWT, 6-MWT und NSAA sowie für die Endpunkte „Körpergröße (z-Scores)“ und „Körpergewicht (z-Scores)“ eingereicht, bei der lediglich Patienten berücksichtigt wurden, die innerhalb des engeren Visitenzeitfensters \pm 7 Tage zu Woche 72 erschienen waren.

Die MMRM-Analysen der Overall-Population für alle bewertungsrelevanten Endpunkte zu der Studienvisite Woche 72 / EOS werden als geeignet angesehen.

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

Mit dem Dossier wurden keine Angaben zu fehlenden Werten differenziert nach Gründen vorgelegt. Diese Angaben wurden im Rahmen der schriftlichen Stellungnahme bzw. im Nachgang zur mündlichen Anhörung für die Erhebungsinstrumente 4SC, RFTF, 10-MWT, 6-MWT und NSAA nachgereicht. Für diese liegen Angaben zu fehlenden Werten der Klasse 1 (Andere Gründe; COVID-19) und der Klasse 2 (Gehunfähigkeit; Körperliche Einschränkungen) vor.

Für 4SC, 10-MWT und 6-MWT wurden mehrheitlich fehlende Werte der Klasse 1 (Givinostat: 11–14 %; Placebo: 7–8 %) angegeben; der Anteil fehlender Werte der Klasse 2 war mit \leq 5 % gering. Eine Differenzierung der Gründe für fehlende Werte bei der Klasse-1-Kategorie „Andere Gründe“ liegt nicht vor. Es ist unklar, inwiefern fehlende Werte, die unter diese Kategorie fallen, zufällig fehlen. Eine abschließende Beurteilung der MAR-Annahme des MMRM ist nicht möglich.

Für den RFTF betrug der Gesamtanteil fehlender Werten in beiden Behandlungsgruppen ca. 28 %, wobei der überwiegende Anteil aus fehlenden Werten der Klasse 2 bestand (Givinostat: 16 %; Placebo: 23 %). Die für die MMRM-Analyse getroffene MAR-Annahme, erscheint bei fehlenden Werten der Klasse 2 nicht zutreffend.

NSAA

Die Rationale zum unterschiedlichen Vorgehen im Umgang fehlender Werte nach Klasse 1 (u. a. Weigerung des Patienten zur Durchführung des Tests) und Klasse 2 (u. a. körperliche Einschränkungen) sowie in Abhängigkeit der Anzahl verfügbarer Items (≥ 9 vs. ≤ 8 Items) zur Bildung des Gesamtscores konnte nicht nachvollzogen werden.

Bildung des Gesamtscores

In Abhängigkeit der verfügbaren Anzahl an Items (≥ 9 vs. ≤ 8 Items) wurden zwei unterschiedliche Imputationsstrategien beschrieben. Der Gesamtscore wurde dann aus den verfügbaren und imputierten Werten gebildet. Eine Rationale, weshalb nur bei einem Anteil verfügbarer Werte von ≤ 8 Items und nicht bei einer anderen Anzahl an verfügbaren Items, die fehlenden Items mit „0“ imputiert wurden, konnte nicht identifiziert werden.

Ein fehlender Wert der Klasse 2 wurde als Vorliegen körperlicher Einschränkungen definiert. Sofern der NSAA (alle Items) aus diesem Grund nicht durchgeführt wurde, erfolgte die Imputation des Gesamtscores mit „0“. Wenn der Patient selbst die Durchführung des Tests verweigerte, erfolgte die Imputationen aller 17 Items gemäß Klasse-1-Imputation. Hierzu wurde der Modalwert der verfügbaren Werte zu der Visite der zugehörigen Behandlungsgruppe, Subpopulation (On-/Off-Target) und Randomisierungsstrata (Glukokortikoid-Behandlungsregime) gebildet. Der Gesamtscore wurde aus den 17 imputierten Items gebildet.

Eine Weigerung des Patienten zur Durchführung des NSAA oder die fehlende Durchführung eines Tests (bzw. der Items) schließt nicht aus, dass der Patient aufgrund körperlicher Einschränkungen die Durchführung verweigerte.

Im Manual werden keine vergleichbaren Angaben zum Umgang fehlender Werte beschrieben. Angaben zur Anzahl an Patienten, bei denen der Gesamtscore vollständig oder teilweise imputiert wurde, liegen nicht vor.

Eingereichte Analysen

Bei DMD handelt es sich um eine progrediente Erkrankung, bei der eine Heilung nicht möglich ist; es kann lediglich der Progress der Erkrankung verzögert werden. Daher werden Responderanalysen für eine Verschlechterung präferiert. Die post hoc durchgeföhrten Responderanalysen für eine Verschlechterung um $\geq 15\%$ zu Woche 72, bei der fehlende Werte als Non-Responder gewertet wurden, werden jedoch als nicht geeignet angesehen. Für die Responderanalysen zur Verschlechterung um $\geq 15\%$ wird die Wertung von fehlenden Werten als Non-Responder (und somit als „keine Verschlechterung“) als nicht adäquat betrachtet.

Aus diesem Grund werden die MMRM-Analysen zum Gesamtscore der Ordinalskala (0–34 Punkte) präferiert. Der Gesamtscore wurde aus beobachteten bzw. imputierten Werten gebildet. Aufgrund der Limitationen im Umgang mit fehlenden Werten zur Bildung des Gesamtscores und den Responderanalysen, werden die post hoc durchgeföhrten MMRM-Analysen zu Woche 72 / EOS lediglich ergänzend dargestellt. Gemäß den nachgereichten Daten des pU beträgt der Anteil fehlender Werte 8–13 % (ausschließlich Klasse 1). Es ist zudem unklar, wie diese nach Bildung des Gesamtscores zustande kamen und ob diese bspw. durch frühzeitigen Studienabbruch und Nicht-Erscheinen im Studienzentrum bedingt sind.

Körpergröße (z-Scores) und Körpergewicht (z-Scores)

Für die Endpunkte „Körpergröße (z-Scores)“ und „Körpergewicht (z-Scores)“ wurden keine Angaben zu fehlenden Werten differenziert nach Klasse 1 und 2 eingereicht, da eine Dokumentation der Gründe für fehlende Werte für die Studie nicht vorgesehen war. Jedoch war der Anteil fehlender

Werte gering (beide Behandlungsgruppen: „Körpergewicht“ ca. 5 %; „Körpergröße“ 5–8 %). Die eingereichten MMRM-Analysen werden aufgrund der geringen Anzahl fehlender Werte als geeignet angesehen.

4.4 Verzerrungspotential auf Studien- und Endpunktebene

Das Verzerrungspotential auf Studienebene wird als unklar eingeschätzt (siehe Nutzenbewertung [3]). Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 2 dargestellt.

Tabelle 2: Verzerrungspotential der im Amendment berücksichtigten und ergänzend dargestellten Endpunkte der Studie EPIDYS

Endpunkt	Verbindung adäquat	ITT adäquat	Hinweise auf ergebnisgesteuerte Berichterstattung	Sonstige, das Verzerrungspotential beeinflussende Punkte	Verzerrungspotential
Treppensteigen von 4 Stufen (4SC) ¹⁾	Ja ²⁾	Nein ³⁾	Nein	Ja ⁴⁾	Hoch
Aufstehen aus der Rückenlage (RFTF)	Ja ²⁾	Nein ⁵⁾	Nein	Ja ⁴⁾	Hoch
10-Meter-Geh-/Lauftest (10-MWT)	Ja ²⁾	Nein ³⁾	Nein	Ja ⁴⁾	Hoch
6-Minuten-Gehtest (6-MWT)	Ja ²⁾	Nein ³⁾	Nein	Nein	Hoch
Funktionelle Leistungsfähigkeit (NSAA)	Ja ²⁾	Nein ⁶⁾	Nein	Ja ⁷⁾	Hoch
Körpergröße (z-Scores)	Ja ²⁾	Nein ⁸⁾	Nein	Nein	Niedrig
Körpergewicht (z-Scores)	Ja ²⁾	Nein ⁸⁾	Nein	Nein	Niedrig

¹⁾ Primärer Endpunkt.

²⁾ 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. Eine Dosisreduktion wurde für beide UE für jeweils 14 % der Patienten beschrieben. Im Placebo-Arm traten bei keiner Person die oben genannten UE auf. Die Endpunkte wurden durch Bewerter/innen erhoben, die nicht in die Überwachung der Sicherheitsendpunkte involviert waren, weshalb von einer niedrigen Auswirkung auf das Verzerrungspotential auszugehen ist.

³⁾ Der Gesamtanteil an fehlenden Werten lag je nach Endpunkt bei ca. 10–14 %, wobei der überwiegende Anteil durch fehlende Werte Klasse 1 (Andere Gründe; COVID-19: Givinostat: 11–14 %; Placebo: 7–8 %) bedingt war. Der Anteil fehlender Werte Klasse 2 (Gehunfähigkeit; Körperliche Einschränkungen) war mit ≤ 5 % gering. Eine Differenzierung der Gründe für fehlende Werte der Klasse-1-Kategorie „Andere Gründe“ liegt nicht vor. Es ist unklar, inwiefern fehlende Werte, die unter diese Kategorie fallen, zufällig fehlen. Eine abschließende Beurteilung der MAR-Annahme des MMRM ist somit nicht möglich. Das Verzerrungspotential wird als hoch eingeschätzt.

⁴⁾ Die Sensitivitätsanalyse mittels MMRM unter Berücksichtigung des prädefinierten Zeitfensters von ± 7 Tage war zu Woche 72 / EOS statistisch nicht signifikant, jedoch war der Rücklauf < 70 % (je nach Endpunkt im Givinostat-Arm bei 51–64 % bzw. im Placebo-Arm bei 48–57 %). Die Effektrichtung war mit der MMRM-Analyse anhand der Overall-Population vergleichbar, jedoch zeigten sich geringere Werte im Effektschätzer. Das Verzerrungspotential wird als unklar eingeschätzt.

⁵⁾ Der Gesamtanteil an fehlenden Werten lag in beiden Behandlungsgruppen bei ca. 28 %, wobei der überwiegende Anteil aus fehlenden Werten Klasse 2 (Gehunfähigkeit; Körperliche Einschränkungen) bestand (Givinostat: 16 %; Placebo: 23 %). Die für die MMRM-Analysen getroffene Annahme, dass fehlende Werte zufällig fehlen (MAR), erscheint bei fehlenden Werten der Klasse 2 nicht zutreffend. Es ist von einem hohen Verzerrungspotential auszugehen.

⁶⁾ Beim Anteil fehlender Werte beträgt der Unterschied 5 % zwischen den Behandlungsarmen (Givinostat: 13 % vs. Placebo: 8 %). Die Ergebnisse sind als potentiell verzerrt zu bewerten. Es ist unklar, wie die fehlenden Werte nach

Bildung des Gesamtscores zustande kamen und ob diese bspw. durch frühzeitigen Studienabbruch und Nicht-Erscheinen im Studienzentrum bedingt sind.

- ⁷⁾ Die Rationale zum unterschiedlichen Vorgehen im Umgang fehlender Werte nach Klasse 1 (u. a. Weigerung des Patienten zur Durchführung des Tests) und Klasse 2 (u. a. körperliche Einschränkungen) sowie in Abhängigkeit der Anzahl verfügbarer Items (≥ 9 vs. ≤ 8 Items) zur Bildung des Gesamtscores konnte nicht nachvollzogen werden. Ein fehlender Wert Klasse 2 wurde als Vorliegen von körperlichen Einschränkungen definiert. Eine Weigerung des Patienten zur Durchführung des NSAA bzw. die fehlende Durchführung eines Tests (= Items), schließt körperliche Einschränkungen nicht aus. Das Verzerrungspotential wird als hoch angesehen.
- ⁸⁾ Der Anteil fehlender Werte lag bei 5–8 %, wobei Angaben zu den Gründen nicht erhoben wurden. Aufgrund der geringen Anzahl an fehlenden Werten wird das Verzerrungspotential als niedrig angesehen.

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; EOS: End of Study; ITT: Intention to Treat; MAR: Missing at Random; MMRM: Mixed Model for Repeated Measures; NSAA: North Star Ambulatory Assessment; PT: Preferred Term; RFTF: Rise From The Floor; UE: Unerwünschtes Ereignis.

5 Ergebnisse

5.1 Morbidität

Treppensteigen von 4 Stufen (4SC)

*Tabelle 3: 4SC (Zeit in Sekunden), MMRM-Analysen (ohne Imputation) zu Woche 72 / EOS;
Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)*

Studie EPIDYS 4SC (Zeit in Sekunden) ¹⁾	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	117 (99,2)	61 (100)
Fehlende Werte, n (%)	1 (0,85)	0 (0)
MW (SD)	3,58 (1,26)	3,60 (1,27)
<i>Woche 72 / EOS²⁾</i>		
Verfügbare Werte, n (%)	104 (88,1)	53 (86,9)
Fehlende Werte ³⁾ , n (%)	14 (11,9)	8 (13,1)
Klasse 1	13 (11,0)	5 (8,20)
Klasse 2	1 (0,85)	3 (4,92)
MW (SD)	4,86 (3,37)	5,43 (3,71)
<i>Veränderung zu Woche 72 / EOS²⁾</i>		
Verfügbare Werte, n (%)	104 (88,1)	53 (86,9)
LS Mean [95%-KI] ⁴⁾	1,40 [1,07; 1,73]	2,13 [1,66; 2,59]
<i>Veränderung zu Woche 72 / EOS²⁾</i>		
LS-Mean-Differenz [95%-KI] ⁴⁾		-0,73 [-1,30; -0,15]
p-Wert ⁴⁾		0,013

¹⁾ Primärer Endpunkt.

²⁾ In die Analysen gingen Personen ein, inklusive derer die nach dem erweiterten Visitenzeitfenster (± 7 Tage zzgl. 2 Monate), die an der Visite teilnahmen. Der Anteil an Personen, die außerhalb des prädefinierten Analysezeitfensters ± 7 Tage, an der Studienvisite teilnahmen, lag bei 24 % im Givinostat- und 33 % im Placebo-Arm. Im Median erschienen diese Patienten ca. 14,5 bzw. 15 Tage später zu der Studienvisite.

³⁾ Fehlende Werte Klasse 1: Andere Gründe; COVID-19. Fehlende Werte Klasse 2: Gehunfähigkeit; Körperliche Einschränkungen.

⁴⁾ MMRM ohne Imputation unter MAR-Annahme, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Visite, Behandlung*Visite, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“).

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

Bei der Sensitivitätsanalyse mittels MMRM unter Berücksichtigung der Patienten, die innerhalb des Visitenzeitfensters ± 7 Tage an der Visite teilnahmen, wurde zur Studienvisite zu Woche 72 kein statistisch signifikanter Unterschied beobachtet (LS-Mean-Differenz: -0,17 (95%-KI: [-0,82; 0,47]); p = 0,60). Zu beachten ist der geringe Rücklauf (< 70 %) zu der Visite (Givinostat: n = 76 (64 %); Placebo: n = 33 (54 %)).

In der MMRM-Analyse zu Woche 72 / EOS wird für den Endpunkt „Treppensteigen von 4 Stufen“ ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen zugunsten von Givinostat beobachtet (geringere Verschlechterung der Zeit beim Treppensteigen von 4 Stufen um -0,73 Sekunden im Givinostat-Arm gegenüber dem Placebo-Arm). Das Verzerrungspotential für den Endpunkt ist hoch. Die obere Grenze des 95%-KI erscheint mit -0,15 Sekunden zu gering, um den beobachteten Effekt als klinisch relevant einzustufen. Diese Einschätzung wird weiterhin gestützt durch die unsichere Robustheit der Ergebnisse in der Gesamtschau der (Sensitivitäts-)Analysen und Unsicherheiten aufgrund fehlender Werte (siehe Tabelle 2).

Aufstehen aus der Rückenlage (RFTF)

Tabelle 4: RFTF (Zeit in Sekunden), MMRM-Analysen (ohne Imputation) zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS RFTF (Zeit in Sekunden)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	117 (99,2)	61 (100)
Fehlende Werte, n (%)	1 (0,85)	0 (0)
MW (SD)	5,98 (2,42)	5,88 (2,25)
<i>Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	85 (72,0)	44 (72,1)
Fehlende Werte ²⁾ , n (%)	33 (28,0)	17 (27,9)
Klasse 1	14 (11,9)	3 (4,92)
Klasse 2	19 (16,1)	14 (23,0)
MW (SD)	7,24 (4,41)	8,15 (4,32)
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	85 (72,0)	44 (72,1)
LS Mean [95%-KI] ³⁾	2,28 [1,70; 2,87]	3,45 [2,64; 4,27]
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
LS-Mean-Differenz [95%-KI] ³⁾		-1,17 [-2,18; -0,16]
p-Wert ³⁾		0,023

¹⁾ In die Analysen gingen Personen ein, inklusive derer die nach dem erweiterten Visitenzeitfenster (± 7 Tage zzgl. 2 Monate), die an der Visite teilnahmen. Der Anteil an Personen, die außerhalb des prädefinierten Analysezeitfensters ± 7 Tage, an der Studienvisite teilnahmen, lag bei 24 % im Givinostat- und 33 % im Placebo-Arm. Im Median erschienen diese Patienten ca. 14,5 bzw. 15 Tage später zu der Studienvisite.

²⁾ Fehlende Werte Klasse 1: Andere Gründe; COVID-19. Fehlende Werte Klasse 2: Gehunfähigkeit; Körperliche Einschränkungen.

³⁾ MMRM ohne Imputation unter MAR-Annahme, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Visite, Behandlung*Visite, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“). Die für die MMRM-Analysen getroffene MAR-Annahme erscheint bei fehlenden Werten der Klasse 2 nicht zutreffend.

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

Bei der Sensitivitätsanalyse mittels MMRM (ohne Imputation) unter Berücksichtigung der Patienten, die innerhalb des Visitenzeitfensters ± 7 Tage an der Visite teilnahmen, wurde zur Studienvisite zu Woche 72 kein statistisch signifikanter Unterschied beobachtet (LS Mean: -0,77 (95%-KI: [-1,95; 0,42]; p = 0,20). Zu beachten ist der geringe Rücklauf (< 70 %) zu der Visite (Givinostat: n = 60 (51 %); Placebo: n = 29 (48 %)).

In der MMRM-Analyse zu Woche 72 / EOS wird für den Endpunkt „Aufstehen aus der Rückenlage“ ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen zugunsten von Givinostat beobachtet (geringere Verschlechterung der Zeit beim Aufstehen aus der Rückenlage um -1,17 Sekunden im Givinostat-Arm gegenüber dem Placebo-Arm). Das Verzerrungspotential für den Endpunkt ist hoch. Die obere Grenze des 95%-KI erscheint mit -0,16 Sekunden zu gering, um den beobachteten Effekt als klinisch relevant einzustufen. Diese Einschätzung wird weiterhin gestützt durch die unsichere Robustheit der Gesamtschau der (Sensitivitäts-)Analysen und Unsicherheiten aufgrund fehlender Werte (siehe Tabelle 2).

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

Tabelle 5: 10-MWT (Zeit in Sekunden), MMRM-Analysen (ohne Imputation) zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS 10-MWT (Zeit in Sekunden)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	117 (99,2)	61 (100)
Fehlende Werte, n (%)	1 (0,85)	0 (0)
MW (SD)	5,54 (1,33)	5,30 (1,03)
<i>Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	104 (88,1)	55 (90,2)
Fehlende Werte ²⁾ , n (%)	14 (11,9)	6 (9,84)
Klasse 1	13 (11,0)	4 (6,56)
Klasse 2	1 (0,85)	2 (3,28)
MW (SD)	6,45 (3,21)	6,95 (2,93)
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	104 (88,1)	55 (90,2)
LS Mean [95%-KI] ³⁾	1,08 [0,81; 1,35]	1,63 [1,26; 2,00]
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
LS-Mean-Differenz [95%-KI] ³⁾		-0,55 [-1,01; -0,10]
p-Wert ³⁾		0,018

¹⁾ In die Analysen gingen Personen ein, inklusive derer die nach dem erweiterten Visitenzeitfenster (± 7 Tage zzgl. 2 Monate), die an der Visite teilnahmen. Der Anteil an Personen, die außerhalb des prädefinierten Analysezeitfensters ± 7 Tage, an der Studienvisite teilnahmen, lag bei 24 % im Givinostat- und 33 % im Placebo-Arm. Im Median erschienen diese Patienten ca. 14,5 bzw. 15 Tage später zu der Studienvisite.

²⁾ Fehlende Werte Klasse 1: Andere Gründe; COVID-19. Fehlende Werte Klasse 2: Gehunfähigkeit; Körperliche Einschränkungen.

³⁾ MMRM ohne Imputation unter MAR-Annahme, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Visite, Behandlung*Visite, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“).

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

Bei der Sensitivitätsanalyse mittels MMRM ohne Imputation unter Berücksichtigung der Patienten, die innerhalb des Visitenzeitfensters ± 7 Tage an der Visite teilnahmen, wurde zur Studienvisite zu Woche 72 kein statistisch signifikanter Unterschied beobachtet (LS Mean: -0,26 (95%-KI: [-0,77; 0,25]); p = 0,32). Zu beachten ist der geringe Rücklauf (< 70 %) zu der Visite (Givinostat: n = 75 (64 %); Placebo: n = 35 (57 %)).

In der MMRM-Analyse zu Woche 72 / EOS wird für den Endpunkt „10-Meter-Geh-/Lauftest“ ein statistisch signifikanter Unterschied zwischen den Behandlungsarmen zugunsten von Givinostat beobachtet (geringere Verschlechterung der Zeit beim Gehen von 10 Metern um -0,55 Sekunden im Givinostat-Arm gegenüber dem Placebo-Arm). Das Verzerrungspotential für den Endpunkt ist hoch. Die obere Grenze des 95%-KI erscheint mit -0,1 Sekunden zu gering, um den beobachteten Effekt als klinisch relevant einzustufen. Diese Einschätzung wird weiterhin gestützt durch die unsichere Robustheit der Ergebnisse in der Gesamtschau der (Sensitivitäts-)Analysen und Unsicherheiten aufgrund fehlender Werte (siehe Tabelle 2).

6-Minuten-Gehtest (6-MWT)

Tabelle 6: 6-MWT (Distanz in Metern), MMRM-Analysen (ohne Imputation) zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS 6-MWT (Distanz in Metern)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	117 (99,2)	61 (100)
Fehlende Werte, n (%)	1 (0,85)	0 (0)
MW (SD)	398,9 (70,9)	393,7(61,4)
<i>Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	101 (85,6)	54 (88,5)
Fehlende Werte ²⁾ , n (%)	17 (14,4)	7 (11,5)
Klasse 1	16 (13,6)	5 (8,20)
Klasse 2	1 (0,85)	2 (3,28)
MW (SD)	355,8 (96,2)	340,4 (93,1)
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	101 (85,6)	54 (88,5)
LS Mean [95%-KI] ³⁾	-47,8 [-56,8; -38,9]	-62,4 [-74,7; -50,0]
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
LS-Mean-Differenz [95%-KI] ³⁾		14,5 [-0,82; 29,9]
p-Wert ³⁾		0,064

¹⁾ In die Analysen gingen Personen ein, inklusive derer die nach dem erweiterten Visitenzeitfenster (± 7 Tage zzgl. 2 Monate), die an der Visite teilnahmen. Der Anteil an Personen, die außerhalb des prädefinierten Analysezeitfensters ± 7 Tage, an der Studienvisite teilnahmen, lag bei 24 % im Givinostat- und 33 % im Placebo-Arm. Im Median erschienen diese Patienten ca. 14,5 bzw. 15 Tage später zu der Studienvisite.

²⁾ Fehlende Werte Klasse 1: Andere Gründe; COVID-19. Fehlende Werte Klasse 2: Gehunfähigkeit; Körperliche Einschränkungen.

³⁾ MMRM ohne Imputation unter MAR-Annahme, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Visite, Behandlung*Visite, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“).

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

Bei der Sensitivitätsanalyse mittels MMRM anhand beobachteter Werte unter Berücksichtigung der Patienten, die innerhalb des Visitenzeitfensters ± 7 Tage an der Visite teilnahmen, wurde zur Studienvisite zu Woche 72 / EOS kein statistisch signifikanter Unterschied beobachtet (LS Mean: 13,2 (95%-KI: [-4,07; 30,5]); p = 0,13). Zu beachten ist der geringe Rücklauf (< 70 %) zu der Visite (Givinostat: n = 73 (62 %); Placebo: n = 35 (57 %)).

Funktionelle Leistungsfähigkeit (NSAA) – ergänzende Darstellung

Tabelle 7: NSAA-Gesamtscore (Ordinalskala: 0–34 Punkte), MMRM-Analysen zu Woche 72 / EOS; Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)

Studie EPIDYS NSAA-Gesamtscore (0–34 Punkte)¹⁾ (ergänzend dargestellt)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	117 (99,2)	61 (100)
Fehlende Werte, n (%)	1 (0,85)	0 (0)
MW (SD)	23,9 (5,34)	24,1 (4,85)
<i>Woche 72 / EOS²⁾</i>		
Verfügbare Werte, n (%)	103 (87,3)	56 (91,8)
Fehlende Werte ³⁾⁴⁾ , n (%)	15 (12,7)	5 (8,2) ⁵⁾
Klasse 1	15 (12,7)	5 (8,2)
Klasse 2	0 (0)	0 (0)
MW (SD)	20,7 (7,41)	19,0 (7,19)
<i>Veränderung zu Woche 72 / EOS²⁾</i>		
Verfügbare Werte, n (%)	103 (87,3)	57 (93,4)
LS Mean [95%-KI] ⁶⁾	-3,35 [-4,01; -2,70]	-5,06 [-5,96; -4,16]
<i>Veränderung zu Woche 72 / EOS²⁾</i>		
LS-Mean-Differenz [95%-KI] ⁶⁾		1,71 [0,59; 2,83]
p-Wert ⁶⁾		0,003

¹⁾ Der NSAA-Gesamtscore (0–34 Punkte) besteht aus 17 Items (0–2 Punkte), die die funktionelle Leistungsfähigkeit bei DMD messen. Höhere Werte zeigen eine bessere funktionelle Leistung an.

²⁾ In die Analysen gingen Personen ein, inklusive derer die nach dem erweiterten Visitenzeitfenster (± 7 Tage zzgl. 2 Monate), die an der Visite teilnahmen. Der Anteil an Personen, die außerhalb des prädefinierten Analysezeitfensters ± 7 Tage, an der Studienvisite teilnahmen, lag bei 24 % im Givinostat- und 33 % im Placebo-Arm. Im Median erschienen diese Patienten ca. 14,5 bzw. 15 Tage später zu der Studienvisite.

³⁾ Weigerte sich der Patient den Test durchzuführen, wurden für die 17 Items die fehlenden Werte als Klasse 1 definiert und mit dem verfügbaren Modalwert der jeweiligen Behandlungsgruppe, Visite und der zugehörigen Subpopulation (= On-/Off-Target) und Glukokortikoid-Behandlungsregime ergänzt. Wurde der Test aufgrund körperlicher Einschränkungen nicht durchgeführt, wurden die Items und der Gesamtscore mit „0“ imputiert.

⁴⁾ Der Gesamtscore wurde aus verfügbaren und imputierten Werten gebildet. In Abhängigkeit der Anzahl verfügbarer Werte (≥ 9 vs. ≤ 8 Items) wurden 2 Imputationsstrategien beschrieben (siehe Kapitel 4.3). Angaben zur Anzahl an Personen mit der jeweiligen Imputationsstrategie liegen nicht vor.

⁵⁾ Gemäß der im Rahmen der Stellungnahme eingereichten Analysen liegen Werte für 56 Patienten (91,8 %) im Placebo-Arm vor, während bei den mit dem Dossier eingereichten MMRM-Analysen für 1 Person mehr verfügbare Werte angegeben werden.

⁶⁾ MMRM ohne Imputation unter MAR-Annahme, adjustiert nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT, Behandlungsgruppe, Visite, Behandlung*Visite, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“).

Abkürzungen: 4SC: 4-Stair Climb; 6-MWT: 6-Minute Walk Test; 10-MWT: 10-Meter Walk/Run Test; DMD: Duchenne-Muskeldystrophie; EOS: End Of Study; KI: Konfidenzintervall; LS: Least Squares; MAR: Missing at Random; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; NSAA: North Star Ambulatory Assessment; RFTF: Rise From The Floor; SD: Standardabweichung.

Bei der Sensitivitätsanalyse mittels MMRM unter Berücksichtigung der Patienten, die innerhalb des Visitenzeitfensters ± 7 Tage an der Visite teilnahmen, wurde zur Studienvisite zu Woche 72 ein statistisch signifikanter Vorteil zugunsten von Givinostat beobachtet (LS-Mean-Differenz: 1,50 (95%-KI: [0,26; 2,74]); p = 0,018). Zu beachten ist der geringe Rücklauf von < 70 % zu der Visite (Givinostat: n = 74 (63 %); Placebo: n = 36 (59 %)).

5.2 Sicherheit

Körpergröße (z-Scores)

*Tabelle 8: Körpergröße (z-Scores), MMRM-Analysen zu Woche 72 / EOS;
Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)*

Studie EPIDYS Körpergröße (z-Scores)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	116 (98,3)	59 (96,7)
Fehlende Werte, n (%)	2 (1,69)	2 (3,28)
MW (SD)	-1,84 (1,08)	-1,78 (1,19)
<i>Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	109 (92,4)	58 (95,1)
Fehlende Werte ²⁾ , n (%)	9 (7,63)	3 (4,92)
MW (SD)	-1,23 (1,35)	-1,08 (1,49)
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	109 (92,4)	58 (95,1)
LS Mean [95%-KI] ³⁾	0,63 [0,56; 0,70]	0,71 [0,62; 0,81]
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
LS-Mean-Differenz [95%-KI] ³⁾		-0,08 [-0,20; 0,04]
p-Wert ³⁾		0,19

¹⁾ In die Analysen gingen Personen ein, inklusive derer die nach dem erweiterten Visitenzeitfenster (± 7 Tage zzgl. 2 Monate), die an der Visite teilnahmen. Der Anteil an Personen, die außerhalb des prädefinierten Analysezeitfensters ± 7 Tage, an der Studienvisite teilnahmen, lag bei 24 % im Givinostat- und 33 % im Placebo-Arm. Im Median erschienen diese Patienten ca. 14,5 bzw. 15 Tage später zu der Studienvisite.

²⁾ Fehlende Werte differenziert nach Gründen wurden im Rahmen der Studie für den Endpunkt nicht erhoben.

³⁾ MMRM ohne Imputation unter MAR-Annahme, adjustiert nach Behandlungsgruppe, Behandlungsvisite, Behandlung*Woche, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“).

Abkürzungen: EOS: End of Study; KI: Konfidenzintervall; LS: Least Squares; MAR: Missing at Random; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; SD: Standardabweichung.

Es wurden Sensitivitätsanalysen zur Körpergröße unter Einbeziehung der prädefinierten Analysezeitfenster von ± 7 Tage eingereicht. Es zeigte sich kein statistisch signifikanter Unterschied zu Woche 72 (LS-Mean-Differenz: -0,13, (95%-KI: [-0,27; 0,003]); p = 0,0548). Zu beachten ist die geringe Rücklaufquote von < 70 % (Givinostat: n = 78 (66 %); Placebo: n = 37 (61 %)). Für die Sensitivitätsanalysen wurde für das MMRM eine zusätzliche Adjustierung nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT angegeben.

Körpergewicht (z-Scores)

*Tabelle 9: Körpergewicht (z-Scores), MMRM-Analysen zu Woche 72 / EOS;
Studie EPIDYS, Overall-Population (finaler Datenschnitt: 22.02.2022)*

Studie EPIDYS Körpergewicht (z-Scores)	Givinostat N = 118	Placebo N = 61
<i>Baseline</i>		
Verfügbare Werte, n (%)	118 (100)	61 (100)
Fehlende Werte, n (%)	0 (0)	0 (0)
MW (SD)	-0,37 (1,27)	-0,32 (1,50)
<i>Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	111 (94,1)	58 (95,1)
Fehlende Werte ²⁾ , n (%)	7 (5,93)	3 (4,92)
MW (SD)	0,09 (1,38)	0,39 (1,53)
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
Verfügbare Werte, n (%)	111 (94,1)	58 (95,1)
LS Mean [95%-KI] ³⁾	0,49 [0,43; 0,55]	0,73 [0,64; 0,81]
<i>Veränderung zu Woche 72 / EOS¹⁾</i>		
LS-Mean-Differenz [95%-KI] ³⁾		-0,24 [-0,34; -0,14]
p-Wert ³⁾		< 0,0001

¹⁾ Es gingen alle Patienten in die Analyse ein. Der Anteil an Personen, die außerhalb des prädefinierten Analysezeitfensters ± 7 Tage, an der Studienvisite teilnahmen, lag bei 24 % im Givinostat- und 33 % im Placebo-Arm. Im Median erschienen diese Patienten ca. 14,5 bzw. 15 Tage später zu der Studienvisite.

²⁾ Fehlende Werte differenziert nach Gründen-wurden im Rahmen der Studie für den Endpunkt nicht erhoben.

³⁾ MMRM anhand beobachteter Werte, adjustiert nach Behandlungsgruppe, Behandlungsvisite, Behandlung*Woche, Patientenalter zum Studienstart und dem Stratifizierungsfaktor Glukokortikoid-Behandlungsregime („Deflazacort tgl.“; „Deflazacort intermittierend“; „Andere Corticosteroide tgl.“; „Andere Corticosteroide intermittierend“).

Abkürzungen: EOS: End of Study; KI: Konfidenzintervall; LS: Least Squares; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; SD: Standardabweichung.

Es wurden Sensitivitätsanalysen zum Körpergewicht unter Einbeziehung des prädefinierten Analysezeitfensters von ± 7 Tage eingereicht. Zu Studienende (Woche 72) ergab sich ein signifikanter Unterschied im Körpergewicht gegenüber Baseline zwischen Givinostat und Placebo (LS-Mean-Differenz: -0,25 (95%-KI: [-0,37; -0,14]); p < 0,0001). Zu beachten ist die geringe Rücklaufquote von < 70 % in beiden Behandlungsgruppen (Givinostat: n = 80 (68 %); Placebo: n = 37 (61 %)). Für die Sensitivitätsanalysen wurde für das MMRM eine zusätzliche Adjustierung nach den Baseline-Werten des 4SC, RFTF, 10-MWT und 6-MWT angegeben.

Gesamteinschätzung zu Körpergröße und Körpergewicht (z-Scores)

Zu Baseline lagen keine auffallenden Imbalancen in Körpergröße und Körpergewicht (z-Scores) zwischen den Behandlungsgruppen vor; BMI-z-Werte liegen nicht vor. Es liegen keine adäquaten Angaben zum Vergleich der Glukokortikoid-Dosierung zu Baseline und während des Studienverlaufs vor, sodass etwaige Imbalancen nicht geprüft werden konnten. Aufgrund der Randomisierung und Vorgaben zur stabilen Dosierung während des Studienverlaufs wird von keinen wesentlichen Einflüssen auf das Verzerrungspotential ausgegangen. Das Verzerrungspotential ist niedrig.

Jedoch waren die Körpergrößen (z-Scores) in beiden Armen im Mittel (-1,84 im Givinostat-Arm bzw. -1,78 im Placebo-Arm) deutlich geringer als in der Referenzpopulation. Dies spiegelt sich zu Baseline auch in – unter Berücksichtigung eines medianen Alters von 10 Jahren – vergleichsweise hohen durchschnittlichen BMI-Werten von 21,6 im Givinostat-Arm und 23,1 im Placebo-

Arm wider (Altersspanne: 6–16 Jahre) [8]. Das deutet daraufhin, dass die Teilnehmenden entgegen der Baseline-Charakteristika des Körpergewichts (z-Scores) nicht als untergewichtig eingeschätzt werden sollten und in Anbetracht ihrer Körpergröße und ihres Alters eher übergewichtig sind [16]. Im Vergleich zur Referenzpopulation liegt das isoliert betrachtete Körpergewicht sowohl zu Baseline als auch zu Woche 72 nahe der Referenzpopulation.

Hinsichtlich der Körpergröße zeigte sich kein statistisch signifikanter Unterschied zu Woche 72 (LS-Mean-Differenz: -0,08, (95%-KI: [-0,20; 0,04]); p = 0,19). Jedoch sind hinsichtlich des Wachstums in beiden Armen Aufholeffekte zu beobachten, der im Placebo-Arm (statistisch nicht signifikant) etwas größer war. Beim Körpergewicht ergab sich zu Studienende (Woche 72) gegenüber Baseline ein statistisch signifikanter Unterschied zwischen Givinostat und Placebo (-0,25 SD im Vergleich zur Referenzpopulation weniger Gewichtszunahme im Givinostat-Arm als im Placebo-Arm; LS-Mean-Differenz: -0,24 [95%-KI: [-0,34; -0,14]; p < 0,0001].

Die Körpergröße (z-Scores) zu Baseline ist im Vergleich zur Referenzpopulation deutlich geringer, wodurch zu Studienende in beiden Armen Aufholeffekte zu beobachten sind. Der anfangs erhöhte BMI sinkt in beiden Armen von Baseline zu Woche 72 trotz zunehmenden Alters der pädiatrischen Population. Daher scheint die Gewichtszunahme im Vergleich zur Referenzpopulation in beiden Armen maßgeblich durch die Aufholeffekte im Wachstum bedingt zu sein. Zudem liegt die Veränderung des Körpergewichts (z-Scores) im Placebo-Arm mit einer numerischen Zunahme um 0,73 sehr nahe an der numerischen Zunahme der Körpergröße (z-Scores) von 0,71, während die Veränderung des Körpergewichts (z-Scores) im Givinostat-Arm mit einer numerischen Zunahme um 0,49 etwas geringer als die numerische Zunahme bei der Körpergröße (z-Scores) um 0,63 ausfällt.

Die Daten deuten in der Gesamtschau nicht darauf hin, dass sowohl im Placebo-Arm als auch im Givinostat-Arm im Kontext von Alter und Körpergrößenwachstum relevante Körpergewichtszunahmen zu beobachten sind, die auf das Nebenwirkungsprofil von Glukokortikoiden zurückzuführen sind. Der beobachtete statistisch signifikante Effekt zum Körpergewicht (z-Scores) wird nicht als klinisch relevant eingeschätzt.

6 Zusammenfassung

Es wurden für die Overall-Population post hoc durchgeführte MMRM-Analysen unter MAR-Annahme für die Messinstrumente 4SC, RFTF, 10-MWT, 6-MWT und NSAA-Gesamtscore (linearisierte Skala, Ordinalskala) sowie die z-Scores der Endpunkte „Körpergröße“ und „Körpergewicht“ mit dem Dossier bzw. der schriftlichen Stellungnahme des pU eingereicht.

Im Rahmen des schriftlichen Stellungnahmeverfahrens wurden Daten zu fehlenden Werten differenziert nach Klasse 1 (Andere Gründe; COVID-19) und Klasse 2 (Gehunfähigkeit; Körperliche Einschränkungen) nachgereicht, Informationen zum eingesetzten MMRM sowie Sensitivitätsanalysen der MMRM-Analysen unter Ausschluss von Patienten, die außerhalb des prädefinierten Zeitfensters zur Studienvisite Woche 72 / EOS teilnahmen.

Der Anteil an Personen, die nach dem prädefinierten Visitenzeitfenster von ± 7 Tage an der Studienvisite teilnahmen, lag bei 24 % im Givinostat- bzw. 33 % im Placebo-Arm, wobei die Patienten im Median ca. 15 Tage außerhalb des Visitenzeitfensters von ± 7 Tage an der Studienvisite teilnahmen; vergleichbar mit dem Visitenzeitfenster zu Woche 60 (± 7 Tage zuzüglich 14 Tage). Aufgrund des längeren Erhebungszeitraums werden die Ergebnisse zu Woche 72 / EOS dargestellt.

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

In der Nutzenbewertung wurden die für die Erhebungsinstrumente 4SC, RFTF, 10-MWT und 6-MWT post hoc durchgeführten MMRM-Analysen aufgrund von statistischen Limitationen nur ergänzend dargestellt (fehlende Konsistenz zu den primären Analysen mittels ANCOVA anhand imputierter Daten; fehlende Angabe zu fehlenden Werten differenziert nach Gründen; unklare Operationalisierung der letzten Studienvisite zu Woche 72 / EOS).

Diese Limitationen wurden im Stellungnahmeverfahren vom pU adressiert und konnten teilweise ausgeräumt werden. Die post hoc durchgeführten MMRM-Analysen ohne Imputation zu Woche 72 / EOS werden im Amendment dargestellt.

Für folgende Erhebungsinstrumente wurde für die Overall-Population ein statistisch signifikanter Vorteil von Givinostat gegenüber Placebo zu Woche 72 beobachtet:

- 4SC: LS Mean: -0,73 (95%-KI: [-1,30; -0,15]); p = 0,013
- RFTF: LS Mean: -1,17 (95%-KI: [-2,18; -0,16]); p = 0,023
- 10-MWT: LS Mean: -0,55 (95%-KI: [-1,01; -0,10]); p = 0,018

Für den 6-MWT wurde kein statistisch signifikanter Unterschied zu Woche 72 / EOS zwischen den Behandlungsgruppen der Overall-Population beobachtet.

Für die in der Overall-Population post hoc durchgeführten MMRM-Sensitivitätsanalysen dieser Endpunkte, unter Berücksichtigung des prädefinierten Zeitfensters von ± 7 Tage, war zu Woche 72 / EOS die Rücklaufquote < 70 % (Givinostat: 51–64 %; Placebo: 48–57 %) und die Ergebnisse waren nicht statistisch signifikant. Die Effektrichtung war gleich mit der Hauptanalyse, jedoch zeigten sich geringere Werte im Effektschätzer.

Es ergibt sich ein hohes Verzerrungspotential für die 4 Endpunkte.

Für den 4SC, 10-MWT und 6-MWT lag der Gesamtanteil fehlender Werte bei 10–14 %, wobei der überwiegende Anteil durch Klasse 1 (Givinostat-Arm: 11–14 %; Placebo-Arm: 7–8 %) bedingt war. Inwieweit die MAR-Annahme des MMRM gegeben ist, kann abschließend nicht beurteilt werden, da keine Differenzierung der Klasse-1-Kategorie „Andere Gründe“ vorliegt.

Für den RFTF ergibt sich ein hohes Verzerrungspotential aufgrund des hohen Anteils fehlender Werte in beiden Behandlungsarmen (ca. 28 %). Die für die MMRM-Analysen getroffene MAR-

Annahme, erscheint bei 16–23 % fehlenden Werten der Klasse 2 nicht zutreffend. Daher werden die MMRM-Analysen für das Erhebungsinstrument als nur bedingt geeignet gewertet.

Zudem ergeben sich Unsicherheiten für 4SC, RFTF und 10-MWT aufgrund der unsicheren Robustheit der MMRM-Analysen.

Klinische Relevanzeinschätzung

Für 4SC, RFTF und 10-MWT zeigen die MMRM-Analysen zu Woche 72 / EOS einen statistisch signifikanten Unterschied zwischen den Behandlungsarmen zugunsten von Givinostat (geringere Verschlechterung in der benötigten Zeit zum Absolvieren des Tests der LS-Mean-Differenz um -0,73 beim 4SC, um -1,17 beim RFTF, um -0,55 beim 10-MWT im Givinostat-Arm gegenüber dem Placebo-Arm). Allerdings ist die obere Grenze des 95%-KI mit -0,10 bis -0,16 Sekunden als gering einzustufen. Das Verzerrungspotential für die Endpunkte ist hoch. Der beobachtete Effekt wird nicht als klinisch relevant eingestuft.

NSAA – ergänzend dargestellt

Der NSAA liefert Informationen zur grundsätzlichen Fähigkeit und zum Grad der Selbstständigkeit eines Patienten. In der Nutzenbewertung zu Givinostat wurde der NSAA u. a. aufgrund der limitierten Angaben zu Reliabilität, Validität und Änderungssensitivität der Ordinalskala (Gesamtscore: 0–34 Punkte) nicht dargestellt. Der pU hat im Rahmen des Stellungnahmeverfahrens Literatur zu Reliabilität, Validität und Änderungssensitivität eingereicht. Das Messinstrument wird als weitestgehend valide angesehen und Auswertungen der präspezifizierten Ordinalskala präferiert (siehe Abschnitt 4.1).

Die Ergebnisse zum NSAA-Gesamtscore werden aufgrund von Limitationen im Umgang mit fehlenden Werten ergänzend dargestellt. Die Rationale zum unterschiedlichen präspezifizierten Vorgehen im Umgang mit fehlenden Werten der Klasse 1 (u. a. Weigerung des Patienten zur Durchführung des Tests) und der Klasse 2 (Körperliche Einschränkungen) sowie in Abhängigkeit der Anzahl verfügbarer Werte (≥ 9 vs. ≤ 8 Items) zur Bildung des Gesamtscores, konnte nicht nachvollzogen werden (siehe Abschnitt 4.3) Angaben zur Anzahl an Patienten, bei denen zur Bildung des Gesamtscores eine Imputation verwendet wurde, liegen nicht vor.

Für die Overall-Population wurde ein statistisch signifikanter Vorteil von Givinostat gegenüber Placebo zu Woche 72 / EOS bei der MMRM-Analyse ohne Imputation festgestellt (Ordinalskala: 0–34 Punkte): LS-Mean-Differenz = 1,71 (95%-KI: [0,59; 2,83]); $p = 0,003$.

Das Verzerrungspotential wird aufgrund der Anzahl fehlender Werte und den Limitationen im Umgang mit fehlenden Werten als hoch angesehen. Es ist unklar, wie die fehlenden Werte (Givinostat: 13 %; Placebo: 8 %) nach Bildung des Gesamtscores zustande kamen und inwiefern diese bspw. durch frühzeitigen Studienabbruch und Nicht-Erscheinen im Studienzentrum bedingt sind.

Körpergröße (z-Scores) und Körpergewicht (z-Scores)

Für die Endpunkte „Körpergröße“ und „Körpergewicht“ wurden für die mit dem Dossier eingereichten Analysen keine alters- (und geschlechts-)adjustierten z-Scores berechnet, weshalb die Endpunkte nicht in der Nutzenbewertung berücksichtigt wurden. Diese Analysen wurden mit der schriftlichen Stellungnahme des pU nachgereicht.

Für die Overall-Population wurde nur für den Endpunkt „Körpergewicht (z-Scores)“ ein statistisch signifikanter Vorteil von Givinostat gegenüber Placebo zu Woche 72 beobachtet (LS-Mean-Differenz: 0,24 (95%-KI: [-0,34; -0,14]); $p < 0,0001$).

Es wurden Sensitivitätsanalysen zu „Körpergewicht“ und „Körpergröße“ unter Einbeziehung der prädefinierten Analysezeitfenster eingereicht. Die Ergebnisse waren für beide Endpunkte vergleichbar. Der signifikante Unterschied in der Veränderung des Körpergewichts (z-Score) blieb bestehen, jedoch war der Rücklauf beider Endpunkte < 70 % zu der Visite (Givinostat: 66–68 %; Placebo: 61 %).

Es wird bei beiden Endpunkten von einem niedrigen Verzerrungspotential ausgegangen.

Klinische Relevanz einschätzung

Für das Körpergewicht (z-Scores) zeigte sich zu Woche 72 / EOS ein statistisch signifikanter Unterschied (LS-Mean-Differenz: 0,24 (95%-KI: [-0,34; -0,14]), was auf eine geringere Gewichtszunahme im Givinostat-Arm gegenüber dem Placebo-Arm hindeutet. Hinsichtlich der Körpergröße zeigte sich kein statistisch signifikanter Unterschied zu Woche 72 / EOS.

Zu Baseline ist die Körpergröße (z-Scores) im Vergleich zur Referenzpopulation deutlich geringer (Givinostat: -1,84; Placebo: -1,78), zu Studienende sind in beiden Armen Aufholeffekte zu beobachten (Givinostat: 0,63; Placebo: 0,71).

Im Vergleich zur Körpergröße ist die Veränderung des Körpergewichts (z-Scores) im Placebo-Arm mit einer numerischen Zunahme um 0,73 sehr nahe an der numerischen Zunahme der Körpergröße (z-Scores), anders als die Veränderung im Givinostat-Arm, wo die Zunahme mit 0,49 etwas geringer ausfällt. Zusammenfassend scheint die Gewichtszunahme im Vergleich zur Referenzpopulation in beiden Armen maßgeblich durch die Aufholeffekte im Wachstum bedingt zu sein.

Die Daten deuten in der Gesamtschau nicht darauf hin, dass im Placebo- oder Givinostat-Arm im Kontext von Alter und Körpergrößenwachstum relevante Körpergewichtszunahmen zu beobachten sind, die auf das Nebenwirkungsprofil von Glukokortikoiden zurückzuführen sind. Der beobachtete statistisch signifikante Effekt zum Körpergewicht (z-Scores) wird als nicht klinisch relevant eingeschätzt.

Referenzen

1. Ayyar Gupta V, Pitchforth JM, Domingos J, Ridout D, Iodice M, Rye C, et al. Determining minimal clinically important differences in the North Star Ambulatory Assessment (NSAA) for patients with Duchenne muscular dystrophy. PLoS One 2023;18(4):e0283669.
2. Flegal KM, Cole TJ. Construction of LMS parameters for the centers for disease control and prevention 2000 growth charts. Natl Health Stat Report 2013;Feb 11(63):1-3.
3. 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 Wirkstoff: Givinostat [online]. Berlin (GER): G-BA; 15.10.2025. [Zugriff: 26.11.2025]. URL: https://www.g-ba.de/downloads/92-975-8952/2025-10-15_Nutzenbewertung_G-BA_Givinostat_D-1223.pdf.
4. 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.
5. 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.
6. 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.
7. 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.
8. 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 [online]. Berlin (GER): Gemeinsamer Bundesausschuss (G-BA); 15.07.2025. [Zugriff: 26.11.2025]. URL: https://www.g-ba.de/downloads/92-975-8947/2025_07_15_Modul4A_Givinostat.pdf.
9. ITF Pharma. Schriftliche Stellungnahme zur Nutzenbewertung nach § 35a SGB V und Kosten-Nutzen-Bewertung nach § 35b SGB V: Givinostat (Duvyzat) [unveröffentlicht]. 05.11.2025.
10. ITF Pharma. Schriftliche Stellungnahme zur Nutzenbewertung nach § 35a SGB V und Kosten-Nutzen-Bewertung nach § 35b SGB V: Givinostat (Duvyzat); Nachreichung [unveröffentlicht]. 27.11.2025.
11. 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.
12. Mayhew AG, Cano SJ, Scott E, Eagle M, Bushby K, Manzur A, et al. Detecting meaningful change using the North Star Ambulatory Assessment in Duchenne muscular dystrophy. Dev Med Child Neurol 2013;55(11):1046-52.
13. Mazzzone 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.

14. **Muntoni F, Domingos J, Manzur AY, Mayhew A, Guglieri M, Sajeev G, et al.** Categorising trajectories and individual item changes of the North Star Ambulatory Assessment in patients with Duchenne muscular dystrophy. PLoS One 2019;14(9):e0221097.
15. **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.
16. **Neuhäuser H, Schienkiewitz A, Rosario AS, Dortschy R, Kurth B-M.** Referenzperzentile für anthropometrische Maßzahlen und Blutdruck aus der Studie zur Gesundheit von Kindern und Jugendlichen in Deutschland (KiGGS); 2. erw. Aufl. [online]. Berlin (GER): Robert Koch-Institut (RKI); 2013. [Zugriff: 12.12.2025]. (Beiträge zur Gesundheitsberichterstattung des Bundes). URL: <https://edoc.rki.de/bitstream/handle/176904/3254/28jWMA04ZjppM.pdf>.
17. **Okama LO, Zampieri LM, Ramos CL, Toledo FO, Alves CRJ, Mattiello-Sverzut AC, et al.** Reliability and validity analyses of the North Star Ambulatory Assessment in Brazilian Portuguese. Neuromuscul Disord 2017;27(8):723-729.
18. **Pane M, Mazzone ES, Sivo S, Sormani MP, Messina S, D'Amico A, et al.** Long term natural history data in ambulant boys with Duchenne muscular dystrophy: 36-month changes. PLoS One 2014;9(10):e108205.
19. **Ricotti V, Ridout DA, Pane M, Main M, Mayhew A, Mercuri E, et al.** The NorthStar Ambulatory Assessment in Duchenne muscular dystrophy: considerations for the design of clinical trials. J Neurol Neurosurg Psychiatry 2016;87(2):149-55.
20. **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.
21. **Zambon AA, Ayyar Gupta V, Ridout D, Manzur AY, Baranello G, Trucco F, et al.** Peak functional ability and age at loss of ambulation in Duchenne muscular dystrophy. Dev Med Child Neurol 2022;64(8):979-988.