

Nutzenbewertung



Gemeinsamer
Bundesausschuss

**von Arzneimitteln mit neuen Wirkstoffen nach
§ 35a SGB V**

**Bewertung von Arzneimitteln für seltene Leiden
nach § 35a Absatz 1 Satz 10 i.V.m.
5. Kapitel § 12 Nr. 1 Satz 2 Verfo**

**Wirkstoff: Ataluren (Neubewertung nach
Fristablauf)**

Datum der Veröffentlichung: 1. September 2016

Inhalt

Tabellenverzeichnis.....	3
Abkürzungsverzeichnis.....	4
Hintergrund.....	6
1 Einführung.....	8
2 Nutzenbewertung.....	10
2.1 Fragestellung.....	10
2.2 Liste der verwendeten Quellen.....	10
2.3 Studiencharakteristika.....	10
2.4 Endpunkte und Erhebungsinstrumente.....	16
2.5 Ergebnisse zum Zusatznutzen.....	28
3 Methodische Anmerkungen und Bewertung der Unterlagen.....	40
3.1 Zulassungsstatus und Zulassungspopulation von Ataluren.....	40
3.2 Design und Methodik.....	40
3.3 Wirksamkeit.....	41
3.4 Lebensqualität.....	44
3.5 Sicherheit.....	44
4 Anforderungen an eine qualitätsgesicherte Anwendung.....	46
5 Zusammenfassung der Bewertung.....	47
Referenzen.....	50

Tabellenverzeichnis

Tabelle 1: Charakterisierung der Studie 020

Tabelle 2: Charakteristika der Studienpopulation in der Studie 020

Tabelle 3: Charakteristika der Studienpopulation nach Stratifizierungsfaktoren in der Studie 020

Tabelle 4: Charakterisierung der Intervention in der Studie 020

Tabelle 5: Vom pU im Dossier als patientenrelevant bezeichnete und ergänzend in der Nutzenbewertung berücksichtigte Endpunkte

Tabelle 6: Verzerrungspotential der Endpunkte in der Studie 020

Tabelle 7: Beschreibung der Endpunkte der Kategorie Morbidität

Tabelle 8: Beschreibung der Endpunkte der Kategorie Lebensqualität

Tabelle 9: Beschreibung der Endpunkte der Kategorie Unerwünschte Ereignisse

Tabelle 10: Ergebnisse zum primären Endpunkt in der Studie 020 (ITT-Population)

Tabelle 11: Ergebnisse zu sekundären und explorativen Endpunkten in der Studie 020 (ITT-Population)

Tabelle 12: Ergebnisse zur Lebensqualität in der Studie 020 (ITT-Population)

Tabelle 13: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE in Studie 020

Tabelle 14: UE in der Studie 020 unabhängig von der Kausalität, die bei > 5 % der Patienten auftraten (As-treated-Population)

Tabelle 15: UE Grad 3 in der Studie 020 (As-treated-Population)

Tabelle 16: Renale und hepatische behandlungsassoziierte UE in der Studie 020 (As-treated-Population)

Tabelle 17: Schwerwiegende UE in der Studie 020 (As-treated-Population)

Tabelle 18: Ergebnisse der Metaanalyse der Patienten aus den Studien 007 und 020 in der Subgruppe ≥ 300 bis < 400 Meter (ADP) im 6MWT im Vergleich zu Placebo

Tabelle 19: Zusammenfassende Darstellung der Ergebnisse der Studie 020 (ITT-Population, präspezifizierte Auswertung) sowie der Metaanalyse (Studien 007 und 020, ADP-Population)

Abkürzungsverzeichnis

6MWD	6-Minuten-Gehstrecke (6-minutes walk distance)
6MWT	6-Minuten-Gehtest
ACTH	Adrenokortikotropes Hormon
ADP	ambulatory decline phase
AM-NutzenV	Arzneimittelnutzenbewertungsverordnung
ANCOVA	Analysis of Covariance (Kovarianzanalyse)
BMD	Becker-Muskeldystrophie
CHMP	Committee for Medicinal Products for Human Use
CK	Kreatinkinase
CSR	Clinical Study Report (Studienbericht)
CTCAE	Common Terminology Criteria for Adverse Events
CTD	Common Technical Document
DMC	Data Monitoring Committee
DMD	Duchenne-Muskeldystrophie
EMA	European Medicines Agency
EPAR	European Public Assessment Report
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention-to-Treat
IVR/IWR	Interactive Voice Response (Interactive Web Response)
kg KG	Körpergewicht gemessen in Kilogramm
KI	Konfidenzintervall
LOCF	last observation carried forward
LS	Least Square
m	Meter

MedDRA	Medical dictionary for regulatory activities
MCID	Minimal clinically important difference
MMRM	mixed-model repated measures
MW	Mittelwert
N	Anzahl
nmDBMD	Nonsense-Mutation Duchenne-Becker Muskeldystrophie
NSAA	North Star Ambulatory Assessment
PODCI	Pediatric Outcomes Data Collection Instrument
PT	Preferred Term
pU	Pharmazeutischer Unternehmer
SAP	Statistical Analysis Plan (statistischer Analyseplan)
SD	Standardabweichung
SGB	Sozialgesetzbuch
SOC	System Organ Class (Systemorganklasse)
SUE	Schwerwiegende/s unerwünschte/s Ereignis/se
TFT	Timed Function Test
UE	Unerwünschte/s Ereignis/se
VerfO	Verfahrensordnung

Hintergrund

Ataluren 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 10 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 10 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.

Dementsprechend hat der G-BA in seiner Sitzung am 15. März 2012 den mit Beschluss vom 1. August 2011 erteilten Auftrag an das IQWiG zur Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen gemäß § 35a Absatz 2 SGB V in der Weise abgeändert, dass bei Orphan Drugs eine Beauftragung des IQWiG mit der Durchführung einer Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann erfolgt, wenn der Umsatz des betreffenden Arzneimittels die gesetzliche Grenze von 50 Millionen Euro überschritten hat und damit einer uneingeschränkten Nutzenbewertung unterliegt (vgl. § 35a Absatz 1 Satz 11 SGB V).

Der G-BA bestimmt gemäß 5. Kapitel § 12 Absatz 1 Nummer 1 Satz 2 der Verfahrensordnung (VerfO) des G-BA 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 Ataluren zieht der G-BA die Zulassung und Unterlagen zur Zulassung (insbesondere den EPAR [5]), die Studie, auf deren Grundlage die Zulassung des Arzneimittels beruht, sowie das Dossier des pharmazeutischen Unternehmers heran und bewertet die darin enthaltenen Daten nach Maßgabe der im 5. Kapitel § 5 Absatz 7 Nr. 1 bis 4 VerfO festgelegten Kriterien im Hinblick auf ihre therapeutische Relevanz.

Der Unterausschuss Arzneimittel hat die Nutzenbewertung zum Wirkstoff Ataluren in seiner Sitzung am 23. August 2016 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 1. Juni 2016 das IQWiG beauftragt, ergänzend eine Bewertung der Angaben des pharmazeutischen Unternehmers in Modul 3 zu folgenden Gesichtspunkten durchzuführen:

- Kosten der Therapie für die gesetzliche Krankenversicherung
- Anzahl der GKV-Patienten in der Zielpopulation

Die Nutzenbewertung wird am 01. 09 2016 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 Einführung

Duchenne-/Becker-Muskeldystrophie (ICD 10 G71.0)

Es handelt sich um X-chromosomal-rezessiv vererbte Erkrankungen, die durch zunehmende proximale Muskelschwäche, bedingt durch Degeneration von Muskelfasern, charakterisiert sind (Sarnat et al., 2007 [15]). Die Becker-Muskeldystrophie (BMD) manifestiert sich später und verläuft milder als die Duchenne-Muskeldystrophie (DMD). Eine weitere Erkrankung, die durch Mutationen im DMD-Gen ausgelöst wird, ist die dilatative Kardiomyopathie. Die Prävalenz der DMD wird mit 1 pro 3.600 bis 1 pro 6.000 lebendgeborenen Jungen angegeben (Bushby et al., 2010 [2]).

Die Erkrankung wird durch eine Mutation im Dystrophin-Gen (DMD-Gen) verursacht. Dystrophin ist ein Strukturprotein in der Muskelfasermembran und wird auf dem X-Chromosom kodiert. Es ist Teil eines Proteinkomplexes, der das Zellskelett einer Muskelfaser mittels Actin mit dem umgebenden Gewebe durch die Zellmembran verbindet. Dystrophin wird auch in anderen Organen exprimiert (Gehirn, Niere, Hoden, Lunge, teilweise auch in Herz und Leber). Es sind bisher ca. 1.000 Mutationen des Dystrophin-Gens bekannt. 60–70 % der Mutationen betreffen Deletionen von einem oder mehreren Exonen. Bei 5–10 % liegen Duplikationen vor, die zu Frameshift-Transkriptionen führen. Häufig sind auch Punktmutationen einzelner Nukleotide (25–35 %). Zu dieser Mutationsart gehören auch sogenannte Nonsense-Mutationen (nm-Mutationen), bei denen statt einer Aminosäure ein Nonsense-Codon entsteht, so dass es zum Abbruch der Proteinsynthese an dieser Stelle kommt.¹

Das typische Manifestationsalter der DMD liegt im 4. Lebensjahr (Bushby et al. 2010 [2]). Initial kommt es zu Muskelschwäche der proximalen Muskulatur im Beckengürtelbereich. In diesem Alter ist das Gower-Phänomen zu beobachten, d. h. beim Aufrichten aus der Bauchlage schieben sich die Kinder entlang der Oberschenkel in die aufrechte Position. Die Kinder laufen auf den Zehen, entwickeln einen watschelnden Gang (Trendelenburg-Zeichen) und eine Lordose. Es kommt häufig zu Stürzen mit Frakturen an Armen und Beinen. Rennen, hüpfen, Treppen gehen und vom Boden aufstehen sind zunehmend erschwert. Die Erkrankung verläuft progressiv, die Kinder entwickeln nach Verlust der Gehfähigkeit Skoliosen, die häufig chirurgisch stabilisiert werden müssen, sowie Kontrakturen (Sprunggelenke, Kniegelenke, Hüfte), an den Waden ist eine Pseudohypertrophie durch Ersatz von Muskel- durch Fett- und Bindegewebe zu beobachten. 50–80 % der Patienten entwickeln eine Kardiomyopathie. Bei vielen Kindern entwickeln sich zudem Lernstörungen und weitere intellektuelle Beeinträchtigungen. Mit zwölf Jahren sind die meisten Kinder auf einen Rollstuhl angewiesen und sterben in der Regel in der dritten Lebensdekade an kardiorespiratorischen Komplikationen.

Die Diagnose ergibt sich aus der typischen Symptomatik, dem Erkrankungsalter, ggf. der Anamnese sowie aus diagnostischen Tests (Elektromyographie, Muskelbiopsie, Erhöhung der Kreatinkinase). Durch die Bestimmung der Dystrophin-Konzentration (nicht nachweisbar bei DMD, reduzierte Konzentration bei BMD) und/oder eine DNA-Mutationsanalyse kann die Diagnose bestätigt werden.

Eine spezifische Therapieoption zur Behandlung der DMD existiert bisher nicht. Die Behandlung ist symptomatisch und richtet sich auf den längstmöglichen Funktionserhalt, die

¹ vgl. <http://www.ncbi.nlm.nih.gov/books/NBK1119>, Zugriff am 18.12.2014

Gewichtskontrolle und die Vermeidung von Komplikationen. Zudem sind psychosoziale Maßnahmen (inklusive Psychotherapie, soziale und pflegerische Betreuung, Psychopharmakologie) zur Begleitung der Patienten und ihrer Familien erforderlich. Atemmasken und elektive Tracheotomien werden zur Unterstützung der Atemfunktion eingesetzt. Medikamentöse Therapien werden zur Verbesserung der Herzfunktion bei Kardiomyopathie (ACE-Hemmer, β -Blocker) eingesetzt. Kortikosteroide werden bei Kindern ab einem Alter von 5 Jahren gegeben, bei denen sich die Muskelfunktion verschlechtert. Die Gehfähigkeit soll durch die frühzeitig begonnene Langzeitgabe von Kortikosteroiden, die der Myopathie entgegenwirken, bis zu drei bis vier Jahre länger erhalten werden können, zudem die Lungen- und Herzfunktion stabilisieren und orthopädische Komplikationen reduzieren. Allerdings treten auch gravierende Nebenwirkungen auf (u. a. „Vollmondgesicht“, Gewichtszunahme, Frakturanfälligkeit aufgrund von Osteoporose, Hypertonie).

2 Nutzenbewertung

2.1 Fragestellung

Ataluren (Translarna®) ist ein Arzneimittel mit dem Status „Orphan Drug“. Gemäß § 35a Abs. 1 Satz 10 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 50 Millionen Euro in den letzten zwölf Kalendermonaten nicht übersteigen, das Ausmaß des Zusatznutzens auf der Grundlage der Zulassung und der die Zulassung begründenden Studien.

Ataluren ermöglicht ein ribosomales Durchlesen eines überflüssigen Stop-Codons infolge einer Nonsense-Mutation, so dass funktionsfähiges Dystrophin transkribiert wird. In der Fachinformation wird eine Gesamtdosis von 40 mg pro kg Körpergewicht pro Tag, aufgeteilt auf drei Dosen (10/10/20 mg/kg KG) im Abstand von jeweils 6 Stunden, 6 Stunden und 12 Stunden empfohlen.

Ataluren wurde 2014 aufgrund der placebokontrollierten, doppelblinden Phase-2b-Studie PTC124-GD-007-DMD mit dem Ziel der Verbesserung der Gehfähigkeit im Vergleich zu Placebo anhand der 6-Minuten-Gehstrecke (6MWD) von der EMA zugelassen und mit Markteintritt in Deutschland vom G-BA bewertet. Mit Beschluss vom 21. Mai 2015 wurde das Ausmaß des Zusatznutzens für diese Indikation als „gering“ festgelegt und der Beschluss gleichzeitig bis zum 1. Juni 2016 befristet. Die Befristung ergab sich aus der Zulassung von Ataluren durch die EMA „unter besonderen Bedingungen“. Bezüglich der zu erbringenden Nachweise forderte die EMA, die Ergebnisse einer noch laufenden Phase-3-Studie im Anwendungsgebiet vorzulegen (Studie PTC124-GD-020-DMD). Der Abschluss dieser Studie war für das Ende des Jahres 2015 geplant. Die Ergebnisse der Studie PTC124-GD-020-DMD liegen inzwischen vor und sind Grundlage der vorliegenden Nutzenbewertung. Um eine zeitnahe Einbeziehung der aufgrund der bedingten Zulassung gegenüber der Zulassungsbehörde zu erbringenden Nachweise hinsichtlich Unbedenklichkeit und Wirksamkeit in die Nutzenbewertung des Arzneimittels nach § 35a SGB V zu ermöglichen, wurde eine Frist von 6 Monaten nach dem Abschluss der von der EMA geforderten klinischen Untersuchung als angemessen erachtet.

2.2 Liste der verwendeten Quellen

Zur Bestimmung des Ausmaßes des Zusatznutzens für Ataluren wurden folgende Unterlagen herangezogen:

- Studienbericht (CSR) inkl. Studienprotokoll, Common Technical Documents (CTD) und statistischem Analyseplan (SAP) der Studie PTC124-GD-020-DMD [12,13,14]
- Herstellerdossier zu Ataluren vom 30. Mai 2016 [11]

Zum Zeitpunkt der Dossierbewertung lag ein aktualisierter EPAR seitens der EMA nicht vor.

2.3 Studiencharakteristika

Die Unterlagen zum Nachweis des Ausmaßes des Zusatznutzens für Ataluren basieren auf der Studie PTC124-GD-020-DMD (Studie 020). Studiendesign, Baseline-Charakteristika und Interventionen werden in den Tabellen 1 bis 4 dargestellt.

Tabelle 1: Charakterisierung der Studie 020

Charakteristikum	Beschreibung
Design	<p>Es handelt sich um eine multizentrische, randomisierte, placebokontrollierte, doppelt verblindete Studie der Phase 3 im Parallelgruppendesign (1:1) mit zwei Behandlungsarmen. Bestandteil der Studie war eine 2-wöchige Screening-Phase, eine 48-wöchige verblindete Behandlungsphase und eine 6-wöchige Nachbeobachtungsphase. Patienten wurden stratifiziert für die Faktoren Alter, Dauer der Einnahme von Kortikosteroiden und 6MWT-Wert zu Baseline.</p> <p>Im Interventionsarm wurde eine tägliche Gesamtdosis von 40 mg/kg Körpergewicht (KG), aufgeteilt auf drei Dosen, eingesetzt. Das primäre Ziel war der Nachweis einer Verbesserung der Gehfähigkeit im Vergleich zu Placebo anhand der 6-Minuten-Gehstrecke (6MWT). Sekundäre Ziele waren die Untersuchung der Wirkung von Ataluren auf die Gehfähigkeit gemessen als Zeit bis zu einer 10%-Verschlechterung im 6MWT; proximale Muskelfunktion gemessen mittels ‚Timed Function Tests‘; körperliche Funktion gemessen mit dem North Star Ambulatory Assessment (NSAA); Elternberichtete gesundheitsbezogene Lebensqualität mittels Pediatric Outcomes Data Collection Instrument (PODCI); Aktivitäten des täglichen Lebens sowie Sicherheit; Compliance mit der Studienmedikation sowie Plasmakonzentration von Ataluren.</p> <p>Es waren Interimsanalysen nach jeweils 6 Monaten durch das Data Monitoring Committee für die Auswertung von Wirksamkeits- bzw. Sicherheitsaspekten vorgesehen.</p>
Population	<p>Einschlusskriterien</p> <p>Nur Patienten, die alle nachgenannten Kriterien erfüllen, waren für die Studie geeignet:</p> <ul style="list-style-type: none"> • unterschriebene Einverständniserklärung • männlich, ≥ 7 Jahre bis ≤ 16 Jahre • für DMD typische klinische Symptome, erhöhte CK, Probleme beim Gehen, die sich mit 6 Jahren manifestiert haben • Nachweis einer Nonsense-Mutation durch Sequenzierung des Dystrophin-Gens durch ein zertifiziertes Labor <ul style="list-style-type: none"> ○ Zertifizierung durch College of American Pathologists, Clinical Laboratory Improvement Act/Amendment oder äquivalente Organisation • dokumentierte Blutentnahme zur Bestätigung einer Nonsense-Mutation im Dystrophin-Gen <ul style="list-style-type: none"> ○ Fähigkeit, ≥ 150 m ohne Hilfe zu gehen im Screening-6MWT und ≤ 80 % des nach Alter und Größe vorausgesagten Wertes <ul style="list-style-type: none"> ○ Ergebnisse von zwei 6MWT-Werten zu Baseline müssen vorliegen und höchstens 20 % voneinander bzw. vom Screening-6MWT abweichen • Einnahme systemischer Kortikosteroide (Prednison, Prednisolon, Deflazacort) unmittelbar vor Studienbeginn für mindestens sechs Monate und ohne signifikante Dosisänderung (abgesehen von Anpassung an das Körpergewicht) für mindestens drei Monate unmittelbar vor Studienbeginn sowie begründete Erwartung, dass die Dosis während der Studienlaufzeit stabil bleibt • keine auffälligen Laborwerte: <ul style="list-style-type: none"> ○ bezogen auf Leber, Nebenniere, Niere, Serumelektrolyte • Bereitschaft zur Konzeptionsverhütung während der Behandlungs- und Nachbeobachtungsphase • Bereitschaft zur Einhaltung der geplanten Studienvisiten, Einnahmepläne für die Prüfmedikation, Labortests, Restriktionen und Prozeduren während der Studie <p>Ausschlusskriterien</p> <p>Patienten, die eines der nachgenannten Kriterien erfüllen, waren für die Studie ungeeignet:</p>

Charakteristikum	Beschreibung					
	<ul style="list-style-type: none"> • Behandlung mit Aminoglykosiden < 3 Monate vor Studienbeginn • Beginn der Behandlung mit systemischen Glukokortikoiden < 6 Monate vor Studienbeginn oder nicht-adaptive Dosiserhöhung (z. B. Dosisanpassung aufgrund von Wachstum und Gewichtszunahme) < 3 Monate vor Studienbeginn • Änderung bei Prophylaxe/Therapie aufgrund einer Herzinsuffizienz < 3 Monate vor Studienbeginn • laufende Behandlung mit Cumarin-Derivaten, Phenytoin, Tolbutamid oder Paclitaxel • vorherige Therapie mit Ataluren • bekannte Hypersensitivität für einen der Inhalts- oder Hilfsstoffe der Studienmedikation • Einnahme einer anderen Prüfmedikation < 3 Monate vor Studienbeginn • größerer chirurgischer Eingriff < 6 Wochen vor Studienbeginn • laufende immunsuppressive Therapie (außer Glukokortikoide) • gleichzeitige Teilnahme an einer anderen klinischen Therapiestudie (außer für PTC Therapeutics) • zu erwartender größerer chirurgischer Eingriff während der 12 Monate Behandlungsphase (z. B. aufgrund einer Skoliose) • tagsüber erforderliche Atemhilfe • nicht kontrollierte klinische Symptome von Herzinsuffizienz (ACC-Stadium C oder D) • frühere oder aktuelle sonstige Erkrankung (z. B. psychiatrische Erkrankung, Alkoholabhängigkeit, Drogenabhängigkeit), Anamnese-, Untersuchungsbefunde, EKG-Befunde oder Laborbefunde, die in der Einschätzung des Prüfarztes die Sicherheit des Patienten gefährden, die komplette Studienteilnahme gefährden oder die Endpunkterhebung beeinträchtigen 					
Interventionen und Zahl der Patienten	<table border="1"> <tr> <td data-bbox="406 1205 900 1429">Ataluren 10/10/20 mg/kg KG Ataluren 40 mg/kg KG (aufgeteilt auf drei Dosen mit 10, 10 und 20 mg/kg KG) täglich: morgens, mittags, abends im Abstand von 6 h, 6 h und 12 h; Einnahme innerhalb von 30 Minuten nach einer Mahlzeit</td> <td data-bbox="911 1205 1378 1429">Placebo Einnahmeschema wie Prüfintervention</td> </tr> <tr> <td data-bbox="406 1435 900 1464">N=115</td> <td data-bbox="911 1435 1378 1464">N=115</td> </tr> </table>	Ataluren 10/10/20 mg/kg KG Ataluren 40 mg/kg KG (aufgeteilt auf drei Dosen mit 10, 10 und 20 mg/kg KG) täglich: morgens, mittags, abends im Abstand von 6 h, 6 h und 12 h; Einnahme innerhalb von 30 Minuten nach einer Mahlzeit	Placebo Einnahmeschema wie Prüfintervention	N=115	N=115	<p>Insgesamt wurden 291 Patienten für den Studieneinschluss evaluiert. Davon wurden 230 Patienten randomisiert (115 je Studienarm). Je ein Patient in jedem Studienarm wurde nicht in der ITT-Analyse ausgewertet, weil keine Nonsense-Mutation im Dystrophin-Gen nachgewiesen werden konnte. Drei weitere Patienten im Placebo- und vier weitere Patienten im Atalurenarm brachen die Studie vorzeitig ab, so dass insgesamt 221 Patienten ausgewertet werden konnten.</p> <p>Anzahl einzuschließender Patienten je Gruppe: Ausgehend von der Hypothese, dass der Unterschied in der Gehstrecke gemessen mittels 6MWT zwischen der Ataluren- und der Placebogruppe nach 48 Wochen 30 Meter beträgt, sollten je Gruppe 105 Patienten eingeschlossen werden. Dabei wurde von einer Standardabweichung von 72 Metern ausgegangen (Wert aus der Studie 007). Die Power würde damit 85 % auf dem Signifikanzniveau $p=0,05$ betragen.</p> <p>Um einen vorzeitigen Studienabbruch von ca. 5 % der Teilnehmer zu kompensieren, sollten insgesamt 220 Patienten (110 je Studienarm) rekrutiert werden.</p>
Ataluren 10/10/20 mg/kg KG Ataluren 40 mg/kg KG (aufgeteilt auf drei Dosen mit 10, 10 und 20 mg/kg KG) täglich: morgens, mittags, abends im Abstand von 6 h, 6 h und 12 h; Einnahme innerhalb von 30 Minuten nach einer Mahlzeit	Placebo Einnahmeschema wie Prüfintervention					
N=115	N=115					
Ort und Zeitraum der	Studiendauer: 26. März 2013 (erster Patient zur ersten Visite) bis 20. August 2015 (letzter Patient zur letzten Visite)					

Charakteristikum	Beschreibung
Durchführung	<p>Studiendauer: 48 Wochen</p> <p>Die Studie wurde in 54 Zentren in 18 Ländern durchgeführt: USA (N=68), Kanada (N=7), Südamerika (Brasilien N=5, Chile N=14), Europa (Belgien N=6, Deutschland N=13, England N=13, Frankreich N=17, Italien N=11, Polen N=8, Schweden N=7, Schweiz N=3, Spanien N=22, Tschechien N=3, Türkei N=16), Israel (N=3), Australien (N=10), Südkorea (N=4).¹</p>
Randomisierung/ Verblindung	<p>Nach Unterzeichnung der Einverständniserklärung und dem Vorliegen eines validen Screening-6MWT erfolgte die Randomisierung durch ein Interactive-Voice-Response/Interactive-Web-Response-(IVR/IWR)-System 1:1 in die folgenden Behandlungsarme:</p> <ul style="list-style-type: none"> • Placebo • Ataluren 10, 10, 20 mg/kg KG Dosis <p>Entsprechend der vorgesehenen Stratifizierung nach Alter (< 9 / ≥ 9 J.), Dauer der Einnahme von Kortikosteroiden (≥ 12 Monate / 6–12 Monate) und 6MWT-Wert zu Baseline (< 350 / ≥ 350 m) wurde eine adaptive Randomisierung vorgenommen, um balancierte Gruppen zu gewährleisten.</p> <p>Patienten, Eltern, Studienpersonal und Angestellte von PTC Therapeutics waren verblindet bis zum Zeitpunkt des Datenbankschlusses. Die Verblindung wurde nur aufgehoben bei Notfällen, auf Verlangen des Data Monitoring Committees zum Zeitpunkt der Interimsanalyse und nach Abschluss der Studie. Lediglich ein unabhängiger Statistiker sowie das Data Monitoring Committee hatten Zugang zu den unverblindeten Daten.</p>
Primärer Endpunkt, sekundäre Endpunkte und explorative Endpunkte gemäß Studienprotokoll	<p>Primärer Endpunkt</p> <ul style="list-style-type: none"> • 6-Minuten-Gehtest (6MWT) im Vergleich zu Baseline nach 48 Wochen <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Zeit bis zu einer mindestens 10%igen Verschlechterung im 6MWT nach 48 Wochen (Progressoren) • Änderung in der proximalen Muskelfunktion in ‚Timed Function Tests‘, d. h. so schnell wie möglich <ul style="list-style-type: none"> ○ 10 m laufen/gehen ○ 4 Treppenstufen hinauf-/hinuntersteigen <p>Explorative Endpunkte</p> <ul style="list-style-type: none"> • North Star Ambulatory Assessment (NSAA) <ul style="list-style-type: none"> ○ Messung der körperlichen Funktion (insgesamt 17 Fähigkeiten) • Änderung der Alltagsaktivitäten und Survey der Krankheitssymptome gemessen mit einem von PTC Therapeutics entwickelten Erhebungsinstrument <ul style="list-style-type: none"> ○ erfasst u. a. Gehen, Gleichgewicht, persönliche Hygiene/Körperpflege, An- und Ausziehen, selbständiges Essen, Badezimmernutzung, mit der Hand schreiben, Leistungen in der Schule, Energieniveau sowie Krankheitssymptome • Änderung in der Lebensqualität gemessen mit dem Pediatric Outcomes Data Collection Instrument (PODCI) in den Domänen <ul style="list-style-type: none"> ○ Transfer/Basismobilität ○ Sport / körperliche Funktion <p>Sonstige</p> <ul style="list-style-type: none"> • Ausmaß der Exposition <ul style="list-style-type: none"> ○ wird bestimmt anhand der Anzahl der Patienten und der jeweils eingenommenen Dosen sowie der Dauer der Einnahme der Studienmedikation • Sicherheit <ul style="list-style-type: none"> ○ Erfassung von unerwünschten Ereignissen (schwerwiegend, nicht

Charakteristikum	Beschreibung
	schwerwiegend) gemäß CTCAE Version 3.0 <ul style="list-style-type: none"> ○ Entnahme von Blutproben für hämatologische und biochemische Analysen ○ Vitalzeichen ○ Ultraschalluntersuchung der Nieren und ableitenden Harnwege: während Screening und nach 48 Wochen ○ körperliche Untersuchung
Subgruppenanalysen	Subgruppenanalysen des primären Endpunkts wurden a priori für die folgenden Variablenausprägungen zu Baseline geplant (vgl. SAP, S. 17): <ul style="list-style-type: none"> • 6MWT (≥ 350 m / < 350 m) • 6MWT-Gruppe (< 300 m, ≥ 300 bis < 400 m, ≥ 400 m) • Alter (< 9 / ≥ 9 Jahre) • Dauer Kortikosteroidtherapie zu Baseline (≥ 6 Monate bis < 12 Monate / ≥ 12 Monate)

¹⁾ Quelle: CSR Tab. 8, S. 55

Entsprechend dem X-chromosomalen Erbgang und dem natürlichen Verlauf der Erkrankung wurden ausschließlich männliche Patienten im Alter von ≥ 7 bis ≤ 16 Jahren in die Studie eingeschlossen. Darunter befanden sich sechs Geschwisterpaare. Zwei dieser Paare wurden in die Placebo-Gruppe, vier Paare in den Ataluren-Studienarm randomisiert. Das Vorhandensein einer Nonsense-Mutation wurde bei allen randomisierten Patienten mittels Sequenzierung in einem zertifizierten Labor bestätigt. Je ein Patient wurde nach der Randomisierung aus der ITT-Population ausgeschlossen, weil keine Nonsense-Mutation nachgewiesen werden konnte.

Tabelle 2: Charakteristika der Studienpopulation in der Studie 020

Studie 020	Placebo (N=115)	Ataluren 10/10/20 mg/kg KG (N=115)
<i>Alter (Jahre)</i>		
Mittelwert (SD)	9,0 (1,65)	8,9 (1,79)
Median	9,0	9,0
Spanne	7–14	7–14
<i>Abstammung, N (%)</i>		
kaukasisch	86 (74,8)	89 (77,4)
schwarz	1 (0,9)	1 (0,9)
asiatisch	6 (5,2)	7 (6,1)
hispanisch	8 (7,0)	4 (3,5)
andere	4 (3,5)	7 (6,1)
nicht berichtet	10 (8,7)	7 (6,1)
<i>Körpergröße (cm)</i>		
Mittelwert (SD)	126,3 (10,4)	126,1 (10,7)
Median	126	125,6
Spanne	101,8–151	93,9–163,5
<i>Körpergewicht (kg)</i>		
Mittelwert (SD)	30,6 (10,4)	31,4 (10,8)
Median	27	29,3

Dossierbewertung für Orphan Drugs

Ataluren (Neubewertung nach Fristablauf)

Studie 020	Placebo (N=115)	Ataluren 10/10/20 mg/kg KG (N=115)
Spanne	18,1–68	15,8–63
<i>Body-Mass-Index (kg/m²)</i>		
Mittelwert (SD)	18,8 (4,1)	19,3 (4,4)
Median	17,9	18,4
Spanne	13–36	11,3–36,2
<i>Alter bei Diagnose (Jahre)</i>		
Mittelwert (SD)	4,4 (2,5)	4,4 (2,1)
Median	4,0	4,0
Spanne	0–11	0–9
<i>Zeit von Diagnosestellung bis Randomisierung (Jahre)</i>		
Mittelwert (SD)	4,6 (2,5)	4,6 (2,4)
Median	4,7	4,8
Spanne	0,4–12,2	0,2–11,7
<i>Phänotypische Diagnose, N (%)</i>		
Watschelgang	76 (66,1)	83 (72,2)
Gower-Zeichen	91 (79,1)	83 (72,2)
Wadenhypertrophie	92 (80)	91 (79,1)
<i>6MWT zu Baseline, m</i>		
Mittelwert (SD)	362,7 (81,4)	364 (73,3)
Median	370,5	375,2
Spanne	142,5–526	166,8–511

Quelle: CSR, S. 58f

Tabelle 3: Charakteristika der Studienpopulation nach Stratifizierungsfaktoren in der Studie 020

Studie 020	Placebo (N=115)	Ataluren 10/10/20 mg/kg KG (N=115)
<i>Alter (Jahre), N (%)</i>		
< 9	53 (46,1)	57 (49,6)
≥ 9	62 (53,9)	58 (50,4)
<i>Dauer der Einnahme von Kortikosteroiden vor Studienbeginn, N (%)</i>		
≥ 6 – < 12 Monate	19 (16,5)	19 (16,5)
≥ 12 Monate	96 (83,5)	96 (83,5)
<i>6MWT (Meter), N (%)</i>		
≥ 350	73 (63,5)	73 (63,5)
< 350	42 (36,5)	42 (36,5)

Quelle: CSR, S. 56

Die Charakteristika der Patienten in den beiden Studienarmen waren insgesamt gut balanciert. Lediglich das Gewicht zu Baseline war in der Ataluren-Gruppe geringfügig höher (Tab. 2). Auch hinsichtlich der Begleittherapien mit Kortikosteroiden und Herzmedikamenten zeigten sich keine relevanten Unterschiede zwischen den Gruppen (CSR, S. 61).

Tabelle 4: Charakterisierung der Intervention in der Studie 020

Intervention	Kontrolle
<p>Behandlungsarm 1: Ataluren 40 mg/kg KG (aufgeteilt auf drei Dosen mit 10, 10 und 20 mg/kg KG) täglich: morgens, mittags, abends im Abstand von (\pm 1 Std.) 6, 6 und 12 Std.; Einnahme innerhalb von 30 Minuten nach einer Mahlzeit. Ataluren wurde als weißes Pulver mit Vanillegeschmack zum Auflösen in Flüssigkeit (Wasser oder Milch) zur Verfügung gestellt. Packungsgrößen waren 125, 250 oder 1.000 mg.</p> <p>Die konkrete Dosierung wurde zentral vom IVR/IWR auf Basis der Angaben zu den Patienten kalkuliert und vorgegeben und von klinischen Pharmakologen dispensiert. Nach 24 Wochen erfolgte eine Nachberechnung, um eventuelle Änderungen im Körpergewicht zu berücksichtigen. Die Eltern/Angehörigen wurden in die Präparation der Suspension eingewiesen. Wenn eine Dosis verpasst wurde, dann sollte sie dennoch eingenommen werden, solange nicht mehr als 4 Stunden vergangen waren. Bei mehr als 4 Stunden Verzögerung wurde die Einnahmezeit der nachfolgenden Dosis angepasst.</p>	<p>Einnahmeschema sowie Aussehen/Verpackung, Geruch und Geschmack waren identisch mit der Prüfintervention und somit nicht von ihr zu unterscheiden.</p>
<p>Vorbehandlung, erlaubte und nicht erlaubte (Begleit-)Medikationen</p> <p>Kortikosteroide sollten möglichst standardisiert eingenommen werden, um Confounding zu vermeiden; eine Neuverordnung oder auch Beendigung sollte möglichst nicht während der Studie vorgenommen werden, es sei denn, es war medizinisch unvermeidlich. Gewichtsbedingte Dosisanpassungen waren erlaubt. Medikamentenwechsel sollte möglichst vermieden werden. Die Gesamtdosis sollte möglichst 0,75 mg/kg KG für Prednison bzw. 0,9 mg/kg KG für Deflazacort nicht überschreiten (Appendix A zum Studienprotokoll).</p> <p>Auch Herzmedikamente (z. B. ACE-Hemmer und β-Blocker) sollten möglichst standardisiert weiterverordnet bzw. möglichst nicht neu begonnen oder abgesetzt werden. Bei Prophylaxe oder Therapie einer Herzinsuffizienz sollte die Dosis in den 3 Monaten vor Studienbeginn möglichst nicht verändert worden sein.</p> <p>Aminoglykoside und andere potentiell nephrotoxische Antibiotika sollten aufgrund des Risikos von renalen Funktionsstörungen vermieden werden. Bei schwerwiegenden Infektionen sollten möglichst andere Antibiotika verwendet werden.</p>	

2.4 Endpunkte und Erhebungsinstrumente

2.4.1 In der Nutzenbewertung berücksichtigte Endpunkte

In der Nutzenbewertung wurden folgende auch vom pharmazeutischen Unternehmer (pU) als patientenrelevant eingestuften Endpunkte berücksichtigt (Tabelle 5).

Tabelle 5: Vom pU im Dossier als patientenrelevant bezeichnete und ergänzend in der Nutzenbewertung berücksichtigte Endpunkte

Endpunkt
Morbidität
Primärer Endpunkt Veränderung im 6MWT im Vergleich zu Baseline nach 48 Wochen
Sekundäre Endpunkte <ul style="list-style-type: none"> • Zeit bis zu einer mindestens 10%igen Verschlechterung im 6MWT nach 48 Wochen (Progressoren) • Änderung in der proximalen Muskelfunktion in ‚Timed Function Tests‘, d. h. so schnell wie möglich <ul style="list-style-type: none"> ○ 10 m laufen/gehen ○ 4 Treppenstufen hinauf-/hinuntersteigen
Explorative Endpunkte <ul style="list-style-type: none"> • North Star Ambulatory Assessment (NSAA)
Lebensqualität
Änderung in der Lebensqualität gemessen mit dem Pediatric Outcomes Data Collection Instrument (PODCI) in den Domänen <ul style="list-style-type: none"> • Transfer/Basismobilität • Sport / körperliche Funktion
Sicherheit
Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse

Im Herstellerdossier wird die Zeit zum Aufstehen aus Rückenlage in Verbindung mit dem NSAA als patientenrelevanter Endpunkt aufgeführt (vgl. Modul 4, S. 162). Dieser Endpunkt ist in der Studie 020 Teil des NSAA (Item 11), in Studie 007 wurde er separat erhoben. Daher wird er in dieser Nutzenbewertung als Komponente des NSAA-Scores berücksichtigt.

2.4.2 Verzerrungspotential auf Studien- und Endpunktebene

Bei der Studie PTC124-GD-020-DMD handelt es sich um eine randomisierte kontrollierte Studie der Phase 3. Die Zuteilung auf die Behandlungsgruppen erfolgte mithilfe von Interactive Voice Response bzw. Interactive Web Response (IVR/IWR). Patienten, Eltern, Studienpersonal und Angestellte von PTC Therapeutics waren während der gesamten Zeit der Studiendurchführung verblindet. Ataluren und Placebo waren nicht unterscheidbar. Alle Studienmedikationen wurden in ihrer konkreten Dosierung zentral vom IVR/IWR auf der Basis der Angaben zu den Patienten kalkuliert und vorgegeben.

Das Verzerrungspotential auf Studienebene wird für die Studie 020 als niedrig eingeschätzt. Das Verzerrungspotential der relevanten Endpunkte wird in Tabelle 6 dargestellt. Die Einschätzung des Verzerrungspotentials bezieht sich auf die a priori geplante Auswertung der ITT-Population.

Tabelle 6: Verzerrungspotential der Endpunkte in der Studie 020

Endpunkt	Verblindung adäquat	ITT adäquat	Keine ergebnis- gesteuerte Berichterstattung	Keine sonstigen Aspekte	Verzerrungs- potential Endpunkt
Morbidität					
Primärer Endpunkt					
Veränderung im 6MWT	ja	ja	ja	ja	gering
Sekundäre Endpunkte					
Zeit bis zu einer mindestens 10%igen Verschlechterung im 6MWT	ja	ja	ja	ja	gering
proximale Muskelfunktionstests (TFT)					
10 m laufen/gehen	ja	ja	ja	ja	gering
4 Treppenstufen hinaufsteigen	ja	ja	ja	ja	gering
4 Treppenstufen hinuntersteigen	ja	ja	ja	ja	gering
körperliche Funktion					
North Star Ambulatory Assessment (NSAA)	ja	ja	ja	ja	gering
Lebensqualität					
Pediatric Outcomes Data Collection Instrument (PODCI)	ja	ja	ja	ja	gering
Sicherheit					
UE	ja	ja	ja	ja	gering
SUE	ja	ja	ja	ja	gering

2.4.3 Charakterisierung der berücksichtigten Endpunkte und Einschätzung der Patientenrelevanz

Laut VerfO des G-BA gilt der medizinische Zusatznutzen durch die Zulassung als belegt. Nachweise gemäß § 5 Absatz 1 bis 6 sind nicht verpflichtend vorzulegen. Zur Beurteilung der Patientenrelevanz der vorgelegten Endpunkte wurden das Dossier, die zugehörigen Studienberichte und die in diesen Quellen zitierte Literatur herangezogen. Ferner wurden eigene Recherchen zur Validität und Patientenrelevanz der Endpunkte durchgeführt. Das Ergebnis der Betrachtung der Endpunkte findet sich aufgeteilt nach Endpunktkategorien in den Tabellen 7 bis 9.

Tabelle 7: Beschreibung der Endpunkte der Kategorie Morbidität

Morbidität	
Endpunkt	Primärer Endpunkt: Veränderung im 6MWT im Vergleich zu Baseline nach 48 Wochen
Operationalisierung und Validierung: Begründung seitens des pU	<p>Operationalisierung: Standardisierte Erhebung nach der Methode der Universität von Kalifornien, wobei es sich um eine gegenüber der Version der American Thoracic Society modifizierte bzw. adaptierte Version handelt (McDonald et al. 2010 [10]): Es wird ein einführender Videofilm vor dem Test gezeigt, während des Tests werden die Patienten ermuntert, außerdem geht u. a. aus Sicherheitsgründen eine Person hinter den Patienten her („Safety Chaser“).</p> <p><u>Durchführung:</u> Die Messung erfolgt in Metern. Hilfsmittel sind nicht erlaubt. Die Testpersonen tragen angemessene Schuhe (festes Schuhwerk) und komfortable Kleidung. Der Test soll in einem Gebäude mit einem Flur von wenigstens 6 Fuß Breite und mit festem, glatten Untergrund durchgeführt werden. Die Gehstrecke wird mit Klebeband markiert, es handelt sich um jeweils 25 m, an deren Ende gewendet wird. Die markierte Linie soll nicht überschritten werden. Laufen/Rennen ist nicht erlaubt. Nach jeder Minute wird eine Markierung gesetzt, ebenso an Stellen, an denen die Kinder stürzen. Die Zeit wird mittels Stoppuhr erfasst. Der Gehstest wird mittels Videokamera aufgezeichnet. Die Anzahl von Runden wird erfasst, auch pro Minute. Außerdem wird die Herzfrequenz mit Hilfe eines tragbaren Monitors erfasst. Wenn sich die Testperson ausruht oder stürzt, wird die Uhr nicht angehalten. Bei einem Sturz, wenn keine Verletzung vorliegt, kann der ‚Safety Chaser‘ der Testperson aufhelfen. Kann die Testperson nicht weiterlaufen (wegen Erschöpfung oder Sturz), dann ist der Test zu Ende. Der ‚Safety Chaser‘ ermuntert die Testperson zudem unter Nutzung standardisierter Phrasen (<i>“Fantastic job (subject name)! Keep Going!”</i>) und setzt die Minutenmarkierungen sowie Markierungen bei Stürzen.</p> <p>Ein ‚Clinical Evaluator‘ nimmt jeweils eine Einschätzung vor, ob ein Testergebnis valide ist.</p> <p>Wenn die Testperson den Instruktionen nicht folgen kann, soll der Test am Folgetag erneut versucht werden. Liegen mehrere valide Testergebnisse zu einem Erhebungszeitpunkt vor, dann zählt das beste erzielte Resultat.</p> <p>Ein Testergebnis wird als <u>nicht valide</u> eingestuft, wenn</p> <ul style="list-style-type: none"> • den Instruktionen nicht Folge geleistet wurde, • die Testperson nicht auf der markierten Strecke bleibt, • die Testperson die Gehrichtung ändert, • die Testperson den Test vor Ablauf von 6 Minuten abbricht, aufgrund von Noncompliance oder aus anderen Gründen als Erschöpfung (Verletzung nach Sturz, Hinsetzen). <p><u>Erhebungszeitpunkte:</u> Der 6MWT wurde beim Screening, zu Baseline (vor Beginn der Studienmedikation) und nach jeweils 8 Wochen Behandlung erhoben. Der Wert sollte jeweils als letzter Test am ersten Tag während des zweitägigen Klinikaufenthalts erhoben werden.</p> <p>Unterstützend wurde der Anteil der Patienten mit einer mindestens 10%igen Verschlechterung nach 48 Wochen (Progressoren) kalkuliert.</p> <p><u>Begründung des pU gemäß der im Studienprotokoll und im Dossier vorhandenen Angaben:</u> Der Verlust der Gehfähigkeit wird als ernsthafteste Komplikation im Krankheitsverlauf von DMD angesehen. Daher ist der Erhalt der Gehfähigkeit ein zentrales Ziel der Therapie. Eine Änderung im 6MWT wird als sensitives Maß für die</p>

Morbidität	
	<p>Krankheitsprogression von DMD-Patienten angesehen, der über den vorgesehenen Beobachtungszeitraum von 48 Wochen prognostische Aussagen zur Entwicklung der Gehfähigkeit erlaubt. In mehreren Studien wurde gezeigt, dass DMD-Patienten eine durchschnittlich geringere Gehstrecke im 6MWT aufweisen als gesunde Gleichaltrige (bspw. in einer Studie mit 55 Jungen mit einem Altersmedian von 8–9 Jahren 363 ± 83 m (Jungen mit DMD) vs. 621 ± 68 m (Gesunde); zitiert nach Studienprotokoll Version 2.0, S. 65). Die pathophysiologische Begründung findet sich darin, dass, basierend auf dem Wirkmechanismus von Ataluren (Dystrophinerhöhung), der Effekt besser durch repetitive exzentrische Muskelkontraktionen (i. e. gehen) zu messen ist als durch isometrische (i. e. Myometrie).</p> <p>Der 6MWT wurde bereits als Endpunkt in Zulassungsstudien für Arzneimittel anderer neuromuskulärer und pulmonaler Erkrankungen eingesetzt, bspw. Mukopolysaccharidose I, II und IV, Morbus Pompe.</p> <p>Die Test-Retest-Reliabilität beträgt $r=0,91$ für Messungen im Wochenabstand mit einem 95%-KI von -13 bis +21 Meter (Mittelwert).</p> <p>Erfahrungen bei Patienten mit pulmonalen und kardialen Erkrankungen legen nahe, dass eine Verbesserung um 4,7–11,6 % (entsprechend 23 bis 45 m) als klinisch relevant angesehen wird. Vorerfahrungen bzgl. einer MCID für den 6MWT bei DMD-Patienten konnten in der Studie 007 gesammelt werden. Der pU wendet verteilungsbasierte Methoden an und berechnet ein Drittel der Standardabweichung des 6MWT zu Baseline für die Ermittlung der MCID. Daraus ergeben sich Werte von 28,5 und 31 m für die MCID. Dieser Wert liegt auch außerhalb der gemessenen Test-Retest-Fluktuation (Studienprotokoll Version 2.0, S. 66f).</p> <p>Es wird darauf hingewiesen, dass der 6MWT ein etablierter Endpunkt ist, der den Status der für das Gehen relevanten Körperfunktionen (neuromuskulär, kardiovaskulär, pulmonal) und damit die Krankheitsprogression adäquat messen kann (S. 68).</p>
Einschätzung zur Qualität u. Patientenrelevanz	<p>Definition und Erhebung des Endpunkts:</p> <p>Die Operationalisierung des Endpunkts ist nachvollziehbar, inklusive der Stratifizierung der Gehstrecke. Es ist anzumerken, dass der 6MWT im Kontext der Arzneimittelentwicklung für die Spezifika der DMD-Patientenpopulation adaptiert wurde.</p> <p>Validität und Patientenrelevanz:</p> <p>Die EMA-Guideline 2015 [4] empfiehlt die Erhebung des funktionellen Status ("functional improvement") als wichtigsten klinischen Endpunkt sowie ergänzend Aktivitäten des täglichen Lebens. Zwei Endpunkte sollten für die Domänen Muskelkraft und motorische Funktion erhoben werden, davon je einer als primärer und sekundärer Endpunkt. Der 6MWT wird als Erhebungsinstrument erwähnt, allerdings mit Vorbehalten bezogen auf einen möglichen Lerneffekt, inter- und intrapersonelle Variabilität und die Definition eines klinisch relevanten Effekts. Der 6MWT, wenn er als primärer Endpunkt ausgewählt wird, sollte in seiner Aussagekraft mit komplementären Tests, wie bspw. 'Timed Function Tests' unterstützt werden.</p> <p>Ein Guideline-Entwurf der FDA hält den 6MWT für Kinder mit DMD ab ca. 5 Jahren ebenfalls für geeignet zur Erfassung von Muskelkraft und Ausdauer, verweist aber auf mögliche Probleme bei der Auswertung, wenn Patienten während der Studie ihre Gehfähigkeit verlieren und der 6MWT dann Nullwerte ausweist (FDA. Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment Guidance for Industry. 2015 [16]).</p> <p>Bzgl. der Validität verweist der pU im Modul 4 auf die Ergebnisse einer Konferenz aus dem Jahr 2014 (Lynn et al., 2015 [7]). Demnach wurden Reliabilität, Normwerte, Änderungssensitivität und Korrelation mit dem PODCI in verschiedenen Studien für den 6MWT untersucht. Die Korrelation mit dem PODCI-Score wurde in einer Studie von Henricson et al., 2013 [6] an 24 Kindern mit DMD und an 36 gesunden Kontrollen</p>

Morbidität	
	<p>untersucht. Der 6MWT-Baseline-Wert der Kinder mit DMD war darin deutlich korreliert mit den PODCI-Baseline-Scores (Gesamtscore $r=0,68$; Mobilitätsskala $r=0,79$; Sport und körperliche Funktionsskala $r=0,68$). 1-Jahres-Follow-up-Daten konnten von 13 Kindern mit DMD (54 %) und 18 gesunden Kindern (50 %) erhoben werden. Für Kinder mit DMD zeigte sich eine signifikante Verschlechterung um 54 m im 6MWT ($p=0,027$) und um 5 Punkte im PODCI-Gesamtscore ($p=0,027$) und 10 Punkte auf der PODCI-Mobilitätsskala ($p=0,013$). Die Korrelationsanalyse zwischen der Veränderung im 6MWT nach einem Jahr und der Lebensqualität in der Gruppe mit DMD zeigte, dass eine Veränderung des 6MWT am stärksten mit einer Veränderung im PODCI-Gesamtscore ($r=0,76$) und auf der PODCI-Mobilitätsskala ($r=0,93$) assoziiert war. Mit der PODCI-Sport-und-körperliche-Funktionsskala zeigte sich eine mittlere Korrelation ($r=0,45$).</p> <p>Die Patientenrelevanz des 6MWT als Surrogat für „Meilensteine“ des Funktionsverlusts ist entsprechend der Expertenempfehlungen nachvollziehbar (Bushby & Connor, 2011 [1], Lynn et al., 2015 [7]). Die gewählte individuelle Relevanzschwelle von 30 m ergibt sich aus verschiedenen Berechnungsansätzen und wird vom pU deutlich oberhalb der zu erwartenden Variabilität angesetzt. In einer Publikation von Yost & Eton, 2005 [17] stimmte die MCID auf der Basis von 0,5 SD besser mit einem ankerbasierten Ansatz überein. Die Patientenrelevanz der vom pU gewählten MCID auf der Basis von 0,3 SD ist daher nicht abschließend einzuschätzen, entspricht aber eher einer moderaten Effektstärke (entsprechend dem Vorschlag Cohens). Bei der Anwendung von 0,5 SD wäre die MCID deutlich höher anzusetzen. Es ist zu beachten, dass es sich um eine individuelle Responseschwelle handelt und nicht um eine MCID für Gruppenunterschiede, d. h. die vom pU gewählte Schwelle für die MCID ist als konservativ einzuschätzen.</p>
Endpunkt	Änderung in der proximalen Muskelfunktion (,Timed Function Tests‘) zu Woche 48
Operationalisierung und Validierung: Begründung seitens des pU	<p>Operationalisierung</p> <p>Die Änderung in der proximalen Muskelfunktion in ,Timed Function Tests‘ (TFT) wurde in drei Testsituationen^a aufgeteilt, d. h. die Patienten sollten so schnell wie möglich</p> <ul style="list-style-type: none"> • 10 m laufen/gehen, • 4 Treppenstufen hinauf-/hinuntersteigen. <p>Das Zeitlimit betrug jeweils 30 Sekunden. Wenn die Aufgabe nicht innerhalb dieser Zeit bewältigt wurde, dann wurden 30 Sekunden notiert. Im Rahmen von Sensitivitätsanalysen wurde ein oberes Limit von 45 Sekunden bzw. der jeweils höchste Wert ausgewertet.</p> <p>Für die Tests wird die Zeit bis zur Komplettierung erfasst. Außerdem werden für jeden Test die folgenden Charakteristika erfasst (CSR, S. 35f.):</p> <ul style="list-style-type: none"> • 10 m laufen/gehen: Einteilung in die folgenden 6 Kategorien: 1 = unfähig, ohne Hilfe zu gehen; 2 = unfähig, ohne Hilfe zu gehen, aber mit Unterstützung durch Orthese oder einer Begleitperson möglich; 3 = ausgeprägt adaptierter Gang mit ausgeprägter Lordose, Geschwindigkeit kann nicht erhöht werden; 4 = mäßig adaptierter Gang, Geschwindigkeit kann erhöht werden, laufen nicht möglich; 5 = Geschwindigkeit kann erhöht werden, laufen nur mit Standphase möglich (d. h. ein Fuß immer auf dem Boden); 6 = laufen möglich, beide Füße vom Boden abgehoben. Höhere Werte repräsentieren ein besseres Ergebnis. • 4 Treppenstufen hinaufsteigen: Einteilung in die folgenden 6 Kategorien: 1 = unfähig, 4 Treppenstufen (Standardhöhe) zu steigen; 2 = steigt 4 Treppenstufen (Standardhöhe) im Stillstand (= mit jeweils einem Fuß und beiden Füßen auf einer Stufe vor dem Ersteigen der nächsten Stufe), beide Arme an einer oder beiden Handläufen; 3 = steigt 4 Treppenstufen (Standardhöhe) im Stillstand (= mit jeweils einem Fuß und beiden Füßen auf einer Stufe vor dem Ersteigen der nächsten Stufe), ein

Morbidität	
	<p>Arm an einem Handlauf; 4 = steigt 4 Treppenstufen (Standardhöhe) im Stillstand (= mit jeweils einem Fuß und beiden Füßen auf einer Stufe vor dem Ersteigen der nächsten Stufe), ohne den Handlauf zu Hilfe zu nehmen; 5 = steigt 4 Treppenstufen (Standardhöhe) alternierend mit beiden Füßen, Handlauf wird unterstützend benötigt; 6 = steigt 4 Treppenstufen (Standardhöhe) alternierend mit beiden Füßen, ohne den Handlauf zu Hilfe zu nehmen. Höhere Werte repräsentieren ein besseres Ergebnis.</p> <ul style="list-style-type: none"> • 4 Treppenstufen hinuntersteigen: Charakterisierung analog zum Treppenhochsteigen: 1 = unfähig (...); 6 = ohne Hilfe möglich. Höhere Werte repräsentieren ein besseres Ergebnis. <p><u>Erhebungszeitpunkte:</u> Screening, zu Baseline an Tag 2 (vor Beginn der Studienmedikation), nach 8, 16, 24, 32, 40 und 48 Wochen.</p> <p>Begründung des pU gemäß der im Dossier vorhandenen Angaben: (siehe Modul 4, S. 83 und Studienprotokoll Version 2.0, S. 68)</p> <p>Die TFTs messen die Funktionalität der proximalen Muskelfunktion, die substantiell schlechter ist im Vergleich zu gesunden Kindern. Verschlechterungen in den TFT-Werten können mit dem Risiko der Rollstuhlpflichtigkeit korreliert werden und sind änderungssensitiv. TFTs sind gut reproduzierbar. Es bestehen langjährige Erfahrungen und eine hohe Akzeptanz in der Fachöffentlichkeit.</p> <p>Die MCID wird im Studienprotokoll mit ca. 1,5 Sekunden angegeben. Zur Begründung wird auf eine Publikation von Escolar et al., 2011 [3] verwiesen. Darin wurden 64 steroidnaive Kinder mit bestätigter DMD im Alter von 4–10 J. in je eine Gruppe mit täglicher vs. wöchentlicher Kortikosteroidmedikation über 12 Monate randomisiert. TFTs wurden als sekundäre Endpunkte erhoben. Es wurde eine Äquivalenzgrenze von $\pm 0,4 \log \text{Sek.}$ festgelegt (die Daten wurden aufgrund einer schiefen Verteilung log-transformiert; der Wert von ca. 1,5 ergibt sich aus der Rücktransformation $\ln(e^{0,4})$; vgl. auch CSR, S. 66, Fußnote 2), ohne jedoch zwischen verschiedenen Komponenten der TFTs zu differenzieren. Eine Herleitung dieses Wertes findet sich in der Publikation nicht. Im Modul 4 werden die folgenden MCIDs für die verschiedenen TFT-Komponenten angegeben: 4 Stufen hinauf- oder hinuntersteigen (2 verschiedene verteilungsbasierte Berechnungsmethoden): 2,1 bzw. 2,2 Sekunden; 10 Meter gehen/laufen: 2,3 bzw. 1,4 Sekunden. Für diese Werte wird auf eine Publikation von McDonald et al., 2013 [9] mit Ergebnissen der Studie 007 verwiesen (Modul 4, S. 165).</p>
Einschätzung zur Qualität u. Patientenrelevanz	<p>Definition und Erhebung des Endpunkts: Die Operationalisierung des Endpunkts bzw. der Einzeltests ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz: Die Funktionstests erfassen Fähigkeiten der Patienten, die für sie relevant sind und deren Verlust gemäß einem internationalen Expertenkonsens als „Meilensteine“ des Krankheitsverlaufs angesehen wird (Bushby & Connor, 2011 [1], Lynn et al., 2015 [7]). Dieser Expertenkonsens beruht auf mehreren Treffen europäischer und US-amerikanischer Experten in den Jahren 2009, 2010, 2013 und 2014 die u. a. das Anliegen der EMA nach Standards für Endpunkte bei DMD angesichts laufender Studien mit neuen Medikamenten adressierten. Die Zusammenkünfte wurden überwiegend aus öffentlichen Mitteln finanziert und beinhalteten Reviews von Studien verschiedener Arbeitsgruppen (u. a. Dänemark, Deutschland, England, Italien, USA), die zahlreiche Patientenverläufe repräsentieren, zum Verlauf der DMD sowie Auswertungen zur Reliabilität und Validität von Instrumenten zur Erhebung von Endpunkten. Aus den Daten zum natürlichen Verlauf der DMD ergeben sich Hinweise auf das durchschnittliche Alter, in dem Ereignisse, wie etwa Verlust der Gehfähigkeit, eintreten.</p>

Morbidität	
	<p>Lynn et al., 2015 [7] weisen darauf hin, dass TFTs weniger änderungssensitiv sind als der 6MWT.</p> <p>Der Begründung des pU für die Patientenrelevanz der TFTs kann insgesamt gefolgt werden. Die Validität der MCID für die TFT-Komponenten kann nicht abschließend eingeschätzt werden.</p>
Endpunkt	North Star Ambulatory Assessment (NSAA)
Operationalisierung und Validierung: Begründung seitens des pU	<p>Operationalisierung:</p> <p>Das Instrument besteht aus 17 Komponenten, das ein- und zweibeiniges Stehen (für jeweils 3 Sekunden), Gehen, Aufstehen aus verschiedenen Positionen, Klettern auf eine Kiste, Kopfanheben in Rückenlage, Hüpfen und 10 Meter laufen umfassen.^b Jede Komponente wird mit 0 (kann nicht ausgeführt werden), 1 (kann auf Umwegen bzw. modifiziert ausgeführt werden) und 2 (normal) bewertet. Daraus ergibt sich ein maximaler Score von 34. Je höher der Score, desto geringer sind die funktionellen Einschränkungen ausgeprägt. Dieser Score wird vom pU in einen linearen Score mit den Ausprägungen 0 (schlechtester Zustand) bis 100 (bester Zustand) transformiert (CSR S. 76, SAP S. 18). Der NSAA bewertet Aufgaben, die kurzzeitige Aktivitäten darstellen und ist daher komplementär zum 6MWT (Ausdauererprobung).</p> <p>Die Bewertung wird durch den Untersucher vorgenommen. Die Bearbeitungszeit beträgt etwa 10 Minuten.</p> <p>Begründung des pU gemäß der im Dossier vorhandenen Angaben:</p> <p>Beim NSAA handelt es sich um ein Instrument zur Ermittlung körperlicher Funktionen. Es wurde vom <i>North Star Clinical Network for Paediatric Neuromuscular Disease Management</i> in England entwickelt und hinsichtlich Reliabilität validiert. Das Instrument wurde bisher noch nicht in einer klinischen Studie eingesetzt (Studienprotokoll Version 2.0, S. 69).</p>
Einschätzung zur Qualität u. Patientenrelevanz	<p>Definition und Erhebung des Endpunkts:</p> <p>Es handelt sich um ein gut beschriebenes Erhebungsinstrument, die Operationalisierung ist nachvollziehbar. Das Instrument wird in der EMA-Guideline als geeignetes Instrument eingestuft (EMA 2015 [4]).</p> <p>Validität und Patientenrelevanz:</p> <p>Eckdaten zu psychometrischen Eigenschaften wurden in verschiedenen Studien ermittelt. Die Test-Retest-Reliabilität wird mit 100 % angegeben, mit Ausnahme der Items Hüpfen (80 %) und 10 Meter laufen (60 %). Der Intraclass-Korrelations-Koeffizient für den Gesamtscore beträgt 0,95. Zur Validität liegen keine Daten vor. (Quelle: Proquolid-Datenbank).</p> <p>In einer Studie mit 198 DMD-Patienten aus England mit stabiler täglicher oder intermittierender Glukokortikoidmedikation wurden die MCID und die Responsivität des Instruments untersucht. Die ordinalskalierten Scores (0–34) wurden mittels Rasch-Analyse auf eine Skala von 0–100 übertragen, um die zu den Scoreunterschieden korrespondierenden klinischen Zustände in metrischen Punktabständen abzubilden (d. h. jeder Scorewert auf der Originalskala wird einem Wert auf einer Skala von 1–100 zugewiesen). Die Auswertung beruht auf insgesamt 805 Untersuchungen. Die Studie zeigt, dass der NSAA bzgl. der abnehmenden Mobilität mit zunehmendem Alter und unterschiedlicher Kortikosteroidregime änderungssensitiv ist. Die MCID wurde als 0,5 der mittleren SD kalkuliert und liegt auf der transformierten Skala bei 7–9 Punkten. Es ist zu beachten, dass das Item „Kopfanheben in Rückenlage“ wegen fehlender Konstruktvalidität aus dem Score ausgeschlossen wurde. Klinisch korrespondiert eine Änderung im transformierten Score um 10 Punkte im Verlust einer Funktion, wie bspw. Hüpfen (Mayhew et al., 2013 [8]).</p> <p>Entsprechend § 2 Satz 3 der Arzneimittelnutzenbewertungsverordnung (AM-NutzenV)</p>

Morbidity	
	ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität.“

^a Die Komponente „Aufstehen aus Rückenlage“ wird zwar im Studienprotokoll Version 2.0 (S. 44, 56) erwähnt (im Amendment 1.2 ergänzt), aber nicht mehr im CSR oder im SAP. Diese Testkomponente ist allerdings Teil der NSAA-Testbatterie, so dass er dort berücksichtigt wurde. In der Studie 007 war diese Komponente Teil der ‚Timed Function Tests‘.

^b Eine detaillierte Beschreibung der einzelnen Komponenten findet sich hier:
http://www.musculardystrophyuk.org/assets/0002/5040/North_Star_Ambulatory_assessment.pdf

Tabelle 8: Beschreibung der Endpunkte der Kategorie Lebensqualität

Lebensqualität	
Endpunkt	Pediatric Outcomes Data Collection Instrument (PODCI)
Operationalisierung und Validierung: Begründung seitens des pU	<p>Operationalisierung: Der Fragebogen wurde ursprünglich von der <i>American Academy of Orthopaedic Surgeons</i> als <i>Pediatric Musculoskeletal Function Instrument</i> entwickelt (McDonald et al., 2010 [10]). Für die Messung der Lebensqualität wurde das <i>Pediatric Outcomes Data Collection Instrument (PODCI)</i> verwendet, allerdings nur die Domänen <i>Transfer/Basismobilität</i> und <i>Sport / körperliche Funktion</i>. Jede dieser Domänen wird mit 0 bis 100 Punkten bewertet. Höhere Scores bedeuten bessere Funktion.</p> <p>Der Fragebogen steht in allen benötigten Sprachen zur Verfügung. Er wird von den Eltern/Betreuungspersonen ausgefüllt, möglichst bei jedem Erhebungszeitpunkt von derselben Person. Erhebungszeitpunkte sind beim Screening, in den Wochen 1, 8, 16, 24, 32, 40 und 48.</p> <p>Begründung des pU gemäß der im Dossier vorhandenen Angaben: Die beiden Domänen wurden ausgewählt, weil sie mit der Krankheitsprogression und klinischen Endpunkten, die oft in klinischen Studien zu DMD eingesetzt werden, gut korrelieren (vgl. Ausführungen in Tab. 7 zum 6MWT). In der Studie von McDonald et al., 2010 [10] korrelieren die Domänen <i>Transfer/Basismobilität</i> und <i>Sport / körperliche Funktion</i> mäßig gut ($r=0,43$ bis $-0,53$ bzw. $0,46$ bis $-0,53$) mit TFTs.</p>
Einschätzung zur Qualität u. Patientenrelevanz	<p>Definition und Erhebung des Endpunkts: Eine genauere Beschreibung des Instruments findet sich nicht im Dossier.</p> <p>Validität und Patientenrelevanz: Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer Verbesserung der Lebensqualität.“ Angaben zur Testgüte fehlen ebenfalls. Auch eine MCID wird nicht angegeben. Die Qualität und Patientenrelevanz sind daher nicht abschließend beurteilbar.</p>

Tabelle 9: Beschreibung der Endpunkte der Kategorie Unerwünschte Ereignisse

Unerwünschte Ereignisse	
Endpunkt	Unerwünschte Ereignisse
Operationalisierung und Validierung:	<p>Operationalisierung: Die Auswertung erfolgte mittels der As-treated-Population. Die Erhebung unerwünschter Ereignisse (UE) und schwerwiegender unerwünschter Ereignisse (SUE)</p>

Unerwünschte Ereignisse	
Endpunkt	Unerwünschte Ereignisse
Begründung seitens des pU	<p>erfolgte gemäß des <i>Medical Dictionary for Regulatory Activities</i> (MedDRA) unter Verwendung der Systemorganklasse (SOC) und dem <i>Preferred Term</i> (PT) und wurde gemäß CTCAE Version 3.0 klassifiziert. <i>Treatment emergent adverse events</i> (TEAE) wurden definiert als Ereignisse, die nach der ersten Einnahme der Studienmedikation bis zu 6 Wochen nach der letzten Dosis auftraten oder sich verschlechterten. Wenn ein UE bei einem Patienten mehrmals auftrat, wurde es nur einmal mit dem höchsten CTCAE-Schweregrad dokumentiert. Begleitmedikamente wurden mittels WHO DRUG Dictionary kodiert. Für jedes Ereignis wurden die folgenden Merkmale erfasst: schwerwiegend oder nicht schwerwiegend, durch Studienmedikation verursacht, Schweregrad, Datum Beginn und Ende, ergriffene Maßnahmen, Verlauf.</p> <p>Da in der Studie 007 leichte Anstiege von Serum-Kreatinin, Harnstoff und Cystatin C beobachtet wurden sowie in einer anderen Studie veränderte Nierenwerte (bei kombinierter Gabe von Ataluren mit Aminoglykosiden), wurde ein zusätzliches Nierenmonitoring durchgeführt, um das nephrotoxische Potential von Ataluren (in Kombination mit anderen Arzneimitteln) zu überprüfen. Aufgrund der ebenfalls erhöhten Lipidwerte in der Studie 007 wurde auch ein zusätzliches Lipidprofil erhoben.</p> <p>Im Einzelnen wurden folgende Maßnahmen durchgeführt:</p> <ul style="list-style-type: none"> • Nierenmonitoring: <ul style="list-style-type: none"> ○ Serum-Kreatinin, Serum-Cystatin C, Serum-Harnstoff-Stickstoff zur Überwachung der Nierenfunktion (glomeruläre Filtrationsrate) ○ Urin-Protein:Kreatinin- und Urin-Protein:Osmolalität-Ratio (Proteinurie) ○ Blut im Urin ○ Serum-Elektrolyte ○ renaler Ultraschall • Monitoring der Lebertoxizität: <ul style="list-style-type: none"> ○ Serumwerte für ALT, AST, GGT, Bilirubin, CK • Monitoring der Lipide: <ul style="list-style-type: none"> ○ Serumwerte für Gesamtcholesterin, LDL, HDL, Triglyceride • Vitalzeichen, Größe, Gewicht beim Screening und zu jeder Studienvisite, körperliche Untersuchung zum Screening und nach Woche 1, 24, 48 und 6 Wochen nach Studienende • Hämatologie, Biochemie: zum Screening und bei jeder Studienvisite • Serum-ACTH, Kortisol, Renin, Aldosteron: zum Screening und bei gesonderter Indikationsstellung • 12-Kanal-EKG zum Screening, nach 48 Wochen und 6 Wochen nach Studienende • Urinanalyse: bei jeder Studienvisite
Einschätzung zur Qualität u. Patientenrelevanz	<p>Definition und Erhebung des Endpunkts: Die Operationalisierung des Endpunkts ist nachvollziehbar.</p> <p>Validität und Patientenrelevanz: Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, <u>der Verringerung von Nebenwirkungen</u> oder einer Verbesserung der Lebensqualität.“</p>

2.4.4 Statistische Methoden

Methoden zur Analyse des primären Endpunkts (Studienprotokoll Version 2.0, S. 55f, SAP Version 1.0, S. 16f)

Für die primäre Analyse (Änderung im 6MWT nach 48 Wochen) wird die mITT-Population herangezogen (siehe Auswertungspopulationen).

Die statistische Auswertung des primären Endpunkts testet die Änderung im 6MWT nach 48 Wochen im Gruppenvergleich mittels Kovarianzanalyse (ANCOVA) nach multipler Imputation fehlender Daten. Die Stratifizierungsfaktoren sowie der 6MWT zu Baseline werden in das Modell einbezogen und auf Interaktion geprüft (siehe Subgruppenanalyse). Aus dem Modell werden Mittelwerte und Standardfehler mittels der MIANALYZE-Prozedur errechnet und mittels t-Test-Statistik auf Signifikanz getestet. Mit diesem Modell werden auch By-visit-Analysen durchgeführt, d. h. zu jedem Erhebungszeitpunkt wird, zurückgehend von einem statistisch signifikanten Effektunterschied zu Woche 48, jeweils getestet, ab wann der Unterschied signifikant war, um den frühestmöglichen Zeitpunkt eines signifikanten Effekts zu ermitteln. Um den Effekt der Allokation von Zwillingen in denselben Studienarm zu untersuchen, wird die Analyse des primären Endpunkts mit den Daten des ersten Zwillings wiederholt. In weiteren Analysen wurden MMRM-Modelle mit einer unstrukturierten Kovarianzmatrix verwendet.

Präspezifizierte supportive Analysen des primären Endpunkts

Supportiv erfolgen die folgenden Analysen (Studienprotokoll, S. 56): Kaplan-Meier-Analyse der Zeit bis zu einer konsistenten Verschlechterung des 6MWT um 10 %, Vergleich der Responder (d. h. < 10 % Verschlechterung im 6MWT) mit den Progressoren (≥ 10 % Verschlechterung im 6MWT nach 48 Wochen).

In allen Analysen wird dem 6MWT ein Ergebnis von 0 zugewiesen, sobald die Patienten aufgrund fehlender Gehfähigkeit den Test nicht mehr absolvieren können.

Sensitivitätsanalyse

Eine Sensitivitätsanalyse wurde durchgeführt, um den Effekt der Allokation von Geschwisterpaaren in denselben Behandlungsarm zu untersuchen.

Subgruppenanalysen

Subgruppenanalysen des primären Endpunkts (Interaktionstests) werden für die folgenden Faktoren geplant (SAP, S. 17): Stratifizierungsfaktor für die Randomisierung für 6MWT (≥ 350 m / < 350 m), Gehstreckengruppe zu Baseline (< 300 m, ≥ 300 bis < 400 m, ≥ 400 m), Dauer der Kortikosteroidbehandlung zu Baseline (≥ 6 Monate bis < 12 Monate / ≥ 12 Monate), Altersgruppe (< 9 / ≥ 9 Jahre). Die Analysen wurden mittels eines ANCOVA-Modells durchgeführt (Änderung im 6MWT zu Woche 48 als abhängige Variable, Stratifizierungsfaktoren als unabhängige Variablen, Baseline-6MWT als Kovariate).

Auswertung der sekundären und explorativen Endpunkte

Sekundäre Endpunkte wurden anhand der ITT-Population im Wesentlichen mit dem Auswertungsmodell für den primären Endpunkt ausgewertet. Die Auswertung der Progressoren sowie des NSAA wurde abweichend wie folgt vorgenommen:

- Für den Endpunkt konsistente Verschlechterung des 6MWT um 10 % (Progressoren) wurde eine Kaplan-Meier-Analyse durchgeführt und der Effektschätzer mittels Cox-Proportional-Hazards-Modell errechnet. Patienten mit weniger als 10%-Verschlechterung nach 48 Wochen bzw. zum Zeitpunkt des letzten 6MWT wurden zensiert. Patienten mit Verlust der Gehfähigkeit wurden als Progressoren klassifiziert.
- Der NSAA-Gesamtscore ergibt sich aus der Summe von 17 Einzelaktivitäten, deren Durchführung mit 0 (kann nicht durchgeführt werden), 1 oder 2 (normal) bewertet wird (Maximalscore 34). Der Score des NSAA wurde in einer alternativen Analyse auf einen Wert von 0–100 linear transformiert, ansonsten aber mit dem ANCOVA-Modell für den primären Endpunkt ausgewertet. Wenn weniger als 13 der 17 vorgesehenen Aktivitäten ausgeführt werden konnten, wurde der Gesamtscore als Missing gezählt. Wenn zwischen 13 und 16 Aktivitäten (=x) ausgeführt wurden, wurde der Gesamtscore als Summe der durchgeführten Aktivitäten multipliziert mit $17/x$ kalkuliert.

Weitere Auswertungsmethoden

Interimsanalyse

Abhängig von der Patientenrekrutierung war nach jeweils etwa 6 Monaten eine Interimsanalyse der Sicherheitsdaten unter Aufrechterhaltung der Verblindung durch das DMC vorgesehen.

Metaanalyse

Für den primären Endpunkt sowie bestimmte sekundäre Endpunkte werden Metaanalysen der Ergebnisse der Studien 007 und 020 durchgeführt. Hierzu werden die Subgruppen der „ambulatory decline phase“ herangezogen (SAP, S. 20).

Auswertungspopulationen

Für die Endpunktauswertungen wurden folgende Auswertungspopulationen verwendet:

- **Intention-to-Treat-(ITT)-Population:** Alle randomisierten Patienten, unabhängig davon, ob sie eine Studienmedikation erhalten hatten. Die Patienten mussten einen gültigen 6MWT-Wert zu Baseline und aus mindestens einer Post-Baseline-Visite haben. Zwillinge wurden als unabhängige Studienteilnehmer behandelt (als mITT-Population bezeichnet). Die ITT-Population wurde für alle in der Nutzenbewertung berücksichtigten Endpunkte herangezogen.
- **As-treated-Population:** Alle randomisierten Patienten, die die Studienmedikation erhalten haben, wobei die Zuordnung zur Behandlungsgruppe der aktuell erhaltenen Medikation entspricht. Diese Population wird für die Evaluation der Sicherheit verwendet. Als *evaluable population* werden alle Patienten bezeichnet, die für die Auswertung eines Endpunkts ausreichende Daten aufweisen.
- **Per-Protocol-Population:** Alle randomisierten Patienten, die die Studienmedikation erhalten haben, wobei die Zuordnung zur Behandlungsgruppe der aktuell erhaltenen Medikation entspricht. Die Patienten haben an der gesamten Studie teilgenommen, mindestens 75 % der Studienmedikation eingenommen und weisen einen gültigen Wert im 6MWT zu Baseline und zu Woche 48 auf. Außerdem durften sie keine unerlaubte Begleitmedikation eingenommen haben. Die Population wird für die supportiven Analysen des primären und von wichtigen sekundären Endpunkten herangezogen. Eine Liste mit

Patienten in der Per-Protocol-Population wurde vor der Entblindung finalisiert und im Trial Master File dokumentiert.

2.5 Ergebnisse zum Zusatznutzen

Die Einnahme von Kortikosteroiden während der Laufzeit der Studie war in den beiden Behandlungsarmen balanciert und stabil. Kardiale Medikamente wurden von 32 Patienten (27,8 %) in der Placebo- und von 22 Patienten (19,1 %) in der Atalurengruppe eingenommen; die Einnahme von ACE-Hemmern war häufiger in der Placebogruppe (24 vs. 16 Patienten). Die Inanspruchnahme nicht-medikamentöser Begleittherapien (z. B. Physiotherapie, Nutzung von Rollstühlen, Beschäftigungstherapie) war üblich und zwischen den Gruppen vergleichbar (CSR, S. 61).

Die Compliance mit der Studienmedikation war insgesamt hoch und in beiden Gruppen vergleichbar. Insgesamt wurden 95,1 % aller geplanten Dosierungen in der Placebo-Gruppe und 95,7 % in der Ataluren-Gruppe eingenommen (patientenberichtete Daten). Nach 24 Wochen wurde bei 68 Patienten (59,6 %) in der Placebo- und bei 39 Patienten (34,2 %) in der Atalurengruppe eine Dosisanpassung aufgrund von Gewichtsveränderungen vorgenommen (über das IVR-/IWR-System).

2.5.1 Ergebnisse für patientenrelevante Endpunkte

Mortalität

Es traten keine Todesfälle im Studienzeitraum auf. Mortalität war nicht als Endpunkt vorgesehen.

Morbidität

Im Studienverlauf verloren 23 Patienten (davon 9 im Ataluren- und 14 im Placeboarm) ihre Gehfähigkeit (d.h. der 6MWT wurde mit 0 bewertet). 19 dieser Patienten hatten zu Baseline eine Gehstrecke im 6MWT von <300 Meter; die übrigen 4 Patienten waren im Placeboarm in der Gruppe mit ≥300 bis <400 Meter Gehstrecke.

Tabelle 10: Ergebnisse zum primären Endpunkt in der Studie 020 (ITT-Population)

Studie 020	Placebo (N=114)	Ataluren (N=114)	Effektschätzer (95%-KI), LS Means p-Wert
Absolute Veränderung im 6MWT zu Woche 48^{1,2)}			
Baseline (m), Mittelwert (SD)	363,5 (81,3)	365,2 (72,6)	-
Woche 48 (m), Mittelwert (SD)	306,5 (157,1)	326,3 (134,1)	
Veränderung Woche 48 gegenüber Baseline (m), Mittelwert (SD)	-57,6 (98,8)	-42,2 (85,0)	
Veränderung im 6MWT im Gruppenvergleich zu Woche 48³⁾			
Gehstreckendifferenz (m)			12,98 (-7,44;33,39) p=0,213

Studie 020	Placebo (N=114)	Ataluren (N=114)	Effektschätzer (95%-KI), LS Means p-Wert
Subgruppenanalyse⁴⁾			
< 300 Meter			
N	21	24	
Baseline (m), Mittelwert (SD)	233,5 (51,2)	256,9 (39,8)	
Woche 48 (m), Mittelwert (SD)	96,8 (115,2)	129,2 (112,3)	
Veränderung Woche 48 gegenüber Baseline (m), Mittelwert (SD)	-134,3 (96,9)	-135,3 (111,8)	-7,71 (-54,93;39,51) p=0,749
≥ 300 bis < 400 Meter			
N	52	47	
Baseline (m), Mittelwert (SD)	354,5 (29,7)	356,7 (26,7)	
Woche 48 (m), Mittelwert (SD)	278,0 (120,7)	328,9 (70,3)	
Veränderung Woche 48 gegenüber Baseline (m), Mittelwert (SD)	-75,5 (106,3)	-28,3 (56,0)	42,89 (11,75;74,03) p=0,007
≥ 400 Meter			
N	41	43	
Baseline (m), Mittelwert (SD)	441,6 (31,6)	435,0 (27,7)	
Woche 48 (m), Mittelwert (SD)	441,8 (51,5)	426,9 (66,0)	
Veränderung Woche 48 gegenüber Baseline (m), Mittelwert (SD)	1,2 (40,8)	-8,4 (56,8)	-9,51 (-43,19;24,18) p=0,58
< 350 Meter			
N	41	41	
Baseline (m), Mittelwert (SD)	276,7 (58,4)	285,6 (46,6)	
Woche 48 (m), Mittelwert (SD)	153,0 (120,3)	189,3 (115,7)	
Veränderung Woche 48 gegenüber Baseline (m), Mittelwert (SD)	-124,8 (98,2)	-101,1 (99,1)	21,89 (-12,32;56,11) p=0,21
≥ 350 Meter			
N	73	73	
Baseline (m), Mittelwert (SD)	412,3 (42,0)	409,9 (38,0)	
Woche 48 (m), Mittelwert (SD)	392,0 (99,5)	399,7 (70,9)	
Veränderung Woche 48 gegenüber Baseline (m), Mittelwert (SD)	-20,2 (77,4)	-10,7 (55,3)	7,98 (-1,53;33,48) p=0,54

¹⁾ Quelle: CSR Tab. 14.2.1.2.1.1, S. 244

²⁾ Insgesamt 23 Patienten verloren im Studienverlauf ihre Gehfähigkeit (Placebo N=14, Ataluren N=9 (p=0,404)), die 6MWT-Werte dieser Patienten wurden ab dem Zeitpunkt des Verlusts der Gehfähigkeit mit 0 registriert. Quelle: CSR, S. 66

³⁾ Quelle: CSR Tab. 14.2.1.3.1, S. 409. ANCOVA-Modell mit Änderung von Baseline als abhängige Variable, unabhängige Variablen sind Stratifizierungsfaktoren (Alter < 9 / ≥ 9 J.; Baseline 6MWT < 350 / ≥ 350 M.; Dauer Kortikosteroidtherapie (≥ 6 bis < 12 / ≥ 12 Monate), Therapie, Baseline 6MWT als Kovariate

⁴⁾ Quelle: CSR Tab. 14.2.1.2.2.2, S. 283; Tab. 14.2.1.3.3.2, S. 414, Tab. 14.2.1.2.2.1, S. 269, Tab. 14.2.1.3.3.1, S. 411. Ein Interaktionstest für die präspezifizierten Subgruppen zeigte ein signifikantes Ergebnis von p=0,05 (CSR Tab. 14.2.1.3.3.2, S. 419).

Das primäre Ziel der Studie 020, der Nachweis der Überlegenheit von Ataluren vs. Placebo in der Behandlung von Patienten mit DMD für den primären Endpunkt 6MWT, wurde nur für die präspezifizierte Subgruppe mit einer Gehstrecke zu Baseline von ≥ 300 bis < 400 Meter erreicht, wobei die untere Grenze des 95%-Konfidenzintervalls mit 11,75 m allerdings unter der Relevanzschwelle lag. Der Unterschied im 6MWT für die Gesamtgruppe zwischen dem Ataluren-10/10/20mg/kg-KG- und dem Placebo-Arm von 12,98 m zugunsten von Ataluren war in der primären Analyse statistisch nicht signifikant und blieb unter der vom pU ermittelten klinischen Relevanzschwelle.

Ergebnisse für den im SAP vorgesehenen Vergleich der Responder (d. h. < 10 % Verschlechterung im 6MWT) mit den Progressoren (≥ 10 % Verschlechterung im 6MWT nach 48 Wochen) werden nicht berichtet.

Ergebnisse der präspezifizierten Subgruppenanalysen für den primären Endpunkt

Subgruppenanalysen des primären Endpunkts (Interaktionstests) wurden für die folgenden Faktoren durchgeführt: Stratifizierungsfaktor 6MWT (≥ 350 / < 350 m), Gehstreckengruppe zu Baseline (< 300 m, ≥ 300 bis < 400 m, ≥ 400 m), Dauer der Kortikosteroidbehandlung zu Baseline (≥ 6 Monate bis < 12 Monate / ≥ 12 Monate) und Altersgruppe (< 9 / ≥ 9 Jahre). Ein statistisch signifikanter p-Wert zeigte sich für die Behandlung-Subgruppe-Interaktion ($p=0,05$). Für die präspezifizierte Subgruppe mit einer Gehstrecke zu Baseline ≥ 300 bis < 400 Meter lag der p-Wert bei 0,007 (CSR Tab. 14.2.1.3.3.2, S. 419). Für die anderen Subgruppen zeigte sich keine signifikante Interaktion mit Ausnahme der Altersgruppen (< 9 / ≥ 9 Jahre), hier lag der p-Wert für die Behandlung-Subgruppe-Interaktion bei $p=0,076$; bei den < 9 -Jährigen lag die Differenz in der Gehstrecke zu Woche 48 bei 16,26 (95%-KI -13,2;45,71) Metern, bei den ≥ 9 -Jährigen bei 9,92 (95%-KI -18,48;38,31) Metern (CSR Tab. 14.2.1.3.5, S. 441, Modul 4, S. 125).

Sensitivitätsanalyse für den primären Endpunkt

Die präspezifizierte Sensitivitätsanalyse ergab ein vergleichbares Ergebnis wie die Primäranalyse des primären Endpunkts.

Tabelle 11: Ergebnisse zu sekundären und explorativen Endpunkten in der Studie 020 (ITT-Population)

Studie 020	Placebo (N=114)	Ataluren (N=114)	Effektschätzer HR/LSM (95%-KI) p-Wert
Zeit bis zu einer persistierenden Verschlechterung um mindestens 10 %¹⁾			
Anzahl Ereignisse (%)	52 (45,6)	49 (43,0)	HR 0,75 (0,51;1,12) p=0,1603
Subgruppenanalyse²⁾			
< 300 Meter			
N	21	24	
Median (Tage), 95%-KI	56 (1,0;111,0)	164 (1,0;225,0)	HR 0,48 (0,24;0,93) p=0,031

Studie 020	Placebo (N=114)	Ataluren (N=114)	Effektschätzer HR/LSM (95%-KI) p-Wert
≥ 300 bis < 400 Meter			
N	52	47	
Median (Tage), 95%-KI	280 (169,9;-)	(280;-)	HR 0,79 (0,44;1,41) p=0,42
≥ 400 Meter			
N	41	43	
Median (Tage), 95%-KI	-	-	HR 1,52 (0,59;3,91) p=0,39
Änderung in der proximalen Muskelfunktion (,Timed Function Tests') zu Woche 48			
10 m laufen/gehen ³⁾			
Baseline (Sekunden), Mittelwert	6,8	6,6	Differenz (95%-KI): -1,07 (-2,4;0,27) p=0,12
48 Wochen	10,3	8,8	
benötigte Zeit, um 4 Treppenstufen hinaufzusteigen ⁴⁾			
Baseline (Sekunden), Mittelwert	6,4	5,9	Differenz (95%-KI): -1,4 (-2,9;0,05) p=0,06
48 Wochen	10,6	8,8	
benötigte Zeit, um 4 Treppenstufen hinunterzusteigen ⁵⁾			
Baseline (Sekunden), Mittelwert	4,8	5,0	Differenz (95%-KI): -1,97 (-3,52;-0,43) p=0,01
48 Wochen	9,3	7,4	
North Star Ambulatory Assessment (NSAA)⁶⁾			
Baseline, Gesamtscore (SD)	21,9 (8,0)	22,2 (7,8)	Differenz (95%-KI): 0,80 (-0,23;1,82) p=0,13
48 Wochen	18,4 (10,0)	20,1 (8,8)	
Δ nach 48 Wochen (SD)	-3,7 (4,3)	-2,7 (4,0)	
Baseline, linearisierter Score (SD)	60,2 (18,4)	60,9 (18,0)	Differenz (95%-KI): 1,51 (-1,16;4,17) p=0,27
48 Wochen	52,2 (23,8)	55,7 (20,6)	
Δ nach 48 Wochen (SD)	-8,4 (10,7)	-6,3 (10,6)	

¹⁾ Quelle: CSR Tab. 14.2.1.5.1, S. 445. Cox-proportional-hazards-Modell mit Behandlung, Stratifizierungsfaktoren und Baseline-6MWT als Kovariate.

²⁾ Quelle: CSR Tab. 14.2.1.5.3, S. 447.

³⁾ Quelle: CSR Tab. 14.2.2.2.1, S. 457, Tab. 14.2.2.2.2, S. 461

⁴⁾ Quelle: CSR Tab. 14.2.2.3.1, S. 514, Tab. 14.2.2.3.2, S. 518

⁵⁾ Quelle: CSR Tab. 14.2.2.4.1, S. 571, Tab. 14.2.2.4.2, S. 575

⁶⁾ Quelle: CSR Tab. 14.2.3.1, S. 640, Tab. 14.2.3.2, S. 644, Tab. 14.2.3.3, S. 645, Tab. 14.2.3.4, S. 652

Die Ergebnisse der sekundären und explorativen Endpunkte zeigen, mit Ausnahme des Endpunkts „Zeit bis zu einer persistierenden Verschlechterung um mindestens 10 %“ in der Subgruppe < 300 Meter Gehstrecke im 6MWT zu Baseline sowie der TFFT-Komponente „benötigte Zeit, um 4 Treppenstufen hinunterzusteigen“, keine statistisch signifikanten Unterschiede. Die Verschlechterung der Gehstrecke um 10 % in der < 300-Meter-Subgruppe trat in der Atalurengruppe nach im Median 164 (95%-KI 1,0;225,0) Tagen ein, in der Placebogruppe nach 56 Tagen (95%-KI 1,0;111,0), HR 0,48 (95%-KI 0,24;0,93), $p=0,0306$. Der Unterschied in der Zeit zum Heruntersteigen von 4 Treppenstufen betrug -1,97 (95%-KI -3,52;-0,43) Sekunden, $p=0,012$, zugunsten der Atalurengruppe.

Lebensqualität

Tabelle 12: Ergebnisse zur Lebensqualität in der Studie 020 (ITT-Population)

Studie 020	Placebo (N=114)	Ataluren (N=114)	LS-MW-Differenz (95%-KI) p-Wert
Änderung in der Lebensqualität gemessen mit dem <i>Pediatric Outcomes Data Collection Instrument (PODCI)</i> (Domänen Transfer/Basismobilität, Sport / körperliche Funktion)			
<i>Subskala Transfer/Basismobilität¹⁾</i>			
Baseline, Mittelwert (SD)	81,4 (15,8)	83,9 (13,1)	1,64 (-2,11;5,39) p=0,39
Woche 48	72,4 (24,8)	77,2 (22,2)	
Δ nach 48 Wochen	-8,8 (15,8)	-6,6 (14,8)	
<i>Subskala Sport / körperliche Funktion²⁾</i>			
Baseline, Mittelwert (SD)	56,0 (20,9)	56,2 (18,9)	2,15 (-1,75;6,05) p=0,28
Woche 48	48,3 (26,0)	51,1 (22,8)	
Δ nach 48 Wochen	-7,3 (15,9)	-5,6 (15,9)	

¹⁾ Quelle: CSR, Tab. 14.2.4.1.1, S. 737, Tab. 14.2.4.1.2, S. 741

²⁾ Quelle: CSR, Tab. 14.2.4.2.1, S. 784, Tab. 14.2.4.2.2, S. 788

Die Ergebnisse zur Lebensqualität, gemessen in den Domänen Transfer/Basismobilität, Sport / körperliche Funktion des PODCI zeigen keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen.

Unerwünschte Ereignisse

Die Sicherheitspopulation (“as-treated”) umfasste alle randomisierten Patienten, die ein Prüfpräparat erhalten hatten (115 je Gruppe).

Ausmaß der Exposition:

Die mediane Dauer der Behandlung war in den Gruppen vergleichbar: Placebo 338,0 (Spanne 43–365) Tage und Ataluren 338,0 (36–364) Tage. Je ein Patient pro Behandlungsarm beendete die Therapie vorzeitig aufgrund unerwünschter CTCAE-Grad-2-Ereignisse (Obstipation bei einem Patienten in der Ataluren- und Krankheitsprogression bei einem Patienten in der Placebogruppe).

Tabelle 13: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE in Studie 020

Patienten mit ...	Placebo (N=115)	Ataluren (N=115)
mindestens 1 UE, N (%)	101 (87,8)	103 (89,6)
UE nach Schweregrad, N (%)		
Grad 1 (mild)	54 (47,0)	61 (53,0)
Grad 2 (moderat)	37 (32,2)	35 (30,4)
Grad 3 (schwer)	9 (7,8)	7 (6,1)
Grad 4 (lebensbedrohlich)	0	0
Therapieabbruch aufgrund von UE	1 (0,9)	1 (0,9)
schwerwiegenden UE, N (%)	4 (3,5)	4 (3,5)
Todesfälle	0	0

Quelle: CSR, S. 90, Tab. 20, Tab. 14.3.1.1, S. 853; für einen Patienten kann mehr als ein Ereignis berichtet werden.

Unerwünschte Ereignisse

Das Nebenwirkungsprofil war zwischen den Behandlungsarmen vergleichbar. Therapieabbrüche aufgrund von UE traten bei je einem Patienten pro Gruppe auf. Es gab keine Todesfälle.

UE, die bei mehr als 5 % der Patienten in einem der Behandlungsarme auftraten, sind in Tabelle 14 nach Systemorganklasse (SOC) nach MedDRA dargestellt. Die häufigsten UE wurden in den SOC Infektionen (Nasopharyngitis), Gastrointestinaltrakt (Erbrechen, Durchfall), Verletzungen (Stürze), Muskuloskelettal (Extremitäten- und Rückenschmerzen) und Respirationstrakt (Husten), Allgemeinerkrankungen (Fieber) und Nervensystem (Kopfschmerzen) beobachtet.

Insgesamt wurden 23 Ereignisse (15 in der Placebo-, 8 in der Atalurengruppe) berichtet, aufgrund deren die Studienmedikation unterbrochen wurde. Gründe waren u. a. Urinauffälligkeiten, Leberwertauffälligkeiten, Bauchschmerzen, Husten, Infektionen, Obstipation.

Tabelle 14: UE in der Studie 020 unabhängig von der Kausalität, die bei > 5 % der Patienten auftraten (As-treated-Population)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=115)	Ataluren (N=115)
Gastrointestinaltrakt	48 (41,7)	52 (45,2)
Erbrechen	21 (18,3)	26 (22,6)
Diarrhoe	10 (8,7)	20 (17,4)
Oberbauchschmerzen	13 (11,3)	9 (7,8)
Übelkeit	7 (6,1)	7 (6,1)
Obstipation	10 (8,7)	3 (2,6)
Bauchschmerzen	5 (4,3)	7 (6,1)
Allgemeine Erkrankungen	32 (27,8)	29 (25,2)
Fieber	12 (10,4)	16 (13,9)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=115)	Ataluren (N=115)
Krankheitsprogression	14 (12,2)	9 (7,8)
Infektionen	50 (43,5)	63 (54,8)
Nasopharyngitis	22 (19,1)	24 (20,9)
Infektion des oberen Respirationstraktes	6 (5,2)	11 (9,6)
Rhinitis	4 (3,5)	8 (7,0)
Verletzung, Vergiftung, prozedurale Komplikation	34 (29,6)	35 (30,4)
Sturz	20 (17,4)	21 (18,3)
Muskuloskelettal und Bindegewebe	32 (27,8)	32 (27,8)
Extremitätenschmerzen	14 (12,2)	10 (8,7)
Rückenschmerzen	8 (7,0)	11 (9,6)
Nervensystem	23 (20,0)	28 (24,3)
Kopfschmerzen	21 (18,3)	21 (18,3)
Respirationstrakt, Brustkorb, Mediastinum	30 (26,1)	34 (29,6)
Husten	13 (11,3)	19 (16,5)
oropharyngeale Schmerzen	6 (5,2)	7 (6,1)

Quelle: CSR, S. 93, Tab. 22 und Tab. 14.3.1.2.1, S. 854

Der Schweregrad der meisten UE war mild oder moderat. Insgesamt wurden 22 Grad-3- oder höher Ereignisse berichtet (je 11 in der Placebo- und der Atalurengruppe). Das häufigste als Grad 3 eingestufte UE war Krankheitsprogression (6 in der Placebo-, 2 in der Atalurengruppe). Lebensbedrohliche UE oder Todesfälle traten nicht auf (Tabelle 15).

Tabelle 15: UE Grad 3 in der Studie 020 (As-treated-Population)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=115)	Ataluren (N=115)
Blut und lymphatisches System	0	1 (0,9)
Neutropenie	0	1 (0,9)
Gastrointestinaltrakt	1 (0,9)	2 (1,7)
Verstopfung	0	1 (0,9)
Diarrhoe	0	1 (0,9)
Erbrechen	1 (0,9)	0
Allgemeine Erkrankungen	7 (6,1)	4 (3,5)
Krankheitsprogression	6 (5,2)	2 (1,7)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=115)	Ataluren (N=115)
Gehunfähigkeit (Abasie)	0	1 (0,9)
peripheres Ödem	0	1 (0,9)
Fieber	0	1 (0,9)
Gangstörungen	1 (0,9)	0
Infektionen	0	1 (0,9)
virale Gastroenteritis	0	1 (0,9)
Verletzung, Vergiftung, prozedurale Komplikation	0	2 (1,7)
Sturz	0	1 (0,9)
posttraumatische Schmerzen	0	1 (0,9)
Muskuloskelettal und Bindegewebe	3 (2,6)	1 (0,9)
Extremitätenschmerzen	0	1 (0,9)
Rückenschmerzen	1 (0,9)	0
Muskelschwäche	1 (0,9)	0
Skoliose	1 (0,9) ¹⁾	0

Quelle: CSR, S. 95, Tab. 23 und Tab. 14.3.1.8, S. 1148

¹⁾ Dieses Skolioseereignis wurde fälschlich als Grad-4-Ereignis kodiert; die korrekte Kodierung ist Grad 1. Die Reklassifizierung wurde allerdings erst nach Datenbankschluss vorgenommen.

Renale oder hepatische behandlungsassoziierte Ereignisse (unabhängig von der Kausalität) traten häufiger in der Ataluren- als in der Placebogruppe auf (Tabelle 16), insbesondere Hämaturie (Placebo 1 vs. Ataluren 7 Ereignisse) und periphere Ödeme (Placebo 0 vs. Ataluren 3 Ereignisse, davon je ein moderates bzw. schwerwiegendes).

Tabelle 16: Renale und hepatische behandlungsassoziierte UE in der Studie 020 (As-treated-Population)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=115)	Ataluren (N=115)
Patienten mit ≥ 1 UE	9 (7,8)	15 (13,0)
Allgemeine Erkrankungen	0	3 (2,6)
peripheres Ödem	0	3 (2,6)
hepatobiliäre Erkrankungen	1 (0,9)	0
hepatische Steatose	1 (0,9)	0
Renale und Harnwegserkrankungen	8 (7,0)	12 (10,4)
Hämaturie	1 (0,9)	7 (6,1)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=115)	Ataluren (N=115)
Urinauffälligkeit	2 (1,7)	1 (0,9)
Harninkontinenz	1 (0,9)	2 (1,7)
Nephrolithiasis	2 (1,7)	0
Myoglobulinurie	1 (0,9)	0
Pollakisurie	1 (0,9)	0
Polyurie	0	1 (0,9)
Nierenzyste	0	1 (0,9)

Quelle: CSR, S. 96, Tab. 24 und Tab. 14.3.1.6, S. 1142

Bei der Analyse der UEs in 8-Wochen-Intervallen zeigte sich kein Hinweis auf zunehmende Toxizität im Zeitverlauf (CSR, S. 97).

Schwerwiegende UE (SUE)

Schwerwiegende UE traten bei insgesamt 8 Patienten auf, jeweils bei 4 Patienten in der Placebo-Gruppe und bei 4 Patienten in der Atalurengruppe (Tabelle 17).

Tabelle 17: Schwerwiegende UE in der Studie 020 (As-treated-Population)

Preferred Term MedDRA, N (%)	Placebo (N=115)	Ataluren (N=115)
Patienten insgesamt	4 (3,5)	4 (3,5)
Ereignisse insgesamt	5 (4,3)	7 (6,1)
Myokarditis	1 (0,9)	0
Leberfunktionsstörung	1 (0,9)	0
Pneumonie	0	2 (1,7)
Bronchiolitis	0	1 (0,9)
Gastroenteritis	1 (0,9)	0
posttraumatische Schmerzen	0	1 (0,9)
Femurfraktur	1 (0,9)	0
Unterschenkelfraktur	1 (0,9)	0
Sehnenerkrankung	0	1 (0,9)
adenoide Hypertrophie	0	1 (0,9)
Nasenmuschelhypertrophie	0	1 (0,9)

Quelle: CSR, S. 101, Tab. 27 und Tab. 14.3.2.1, S. 1150. Patienten können mehrfach gezählt werden.

Weitere sicherheitsrelevante Aspekte

Aufgrund der Beobachtung in der Studie 007 und einer weiteren Studie von leichten Anstiegen von Serum-Kreatinin, Harnstoff und Cystatin C sowie veränderten Nierenwerten (bei kombinierter Gabe von Ataluren mit Aminoglykosiden) wurde ein zusätzliches Nierenmonitoring durchgeführt, um das nephrotoxische Potential von Ataluren (in Kombination mit anderen Arzneimitteln) zu überprüfen. Aufgrund der ebenfalls erhöhten Lipidwerte in der Studie 007 wurde auch ein zusätzliches Lipidprofil erhoben.

In der Studie 020 fanden sich keine Hinweise auf hämatologische Toxizität. Während der Studienlaufzeit wurden durchweg erhöhte Werte für CK, AST, ALT und LDH sowie niedrige Baselinewerte für Kreatinin, die im Verlauf im Atalurenarm leicht anstiegen, beobachtet, die mit dem natürlichen Verlauf der DMD (Muskelyse bzw. Kortikosteroideinnahme) konsistent sind. Auch Cystatin C und Harnstoff zeigten im Atalurenarm im Studienverlauf leichte Anstiege. Die Werte für HDL, LDL, Gesamtcholesterin und Triglyceride sowie für GGT, AP und Bilirubin änderten sich kaum im Studienverlauf. Bei 2 Patienten im Placeboarm fand sich eine geringgradige Hämaturie, im Atalurenarm bei 2 Patienten eine mäßige und bei 10 Patienten eine ausgeprägte Hämaturie. Bei einem Patienten im Atalurenarm bestand die Hämaturie bereits bei Baseline und hielt während der gesamten Studiendauer an. Insgesamt wurden keine klinisch relevanten, behandlungsbedürftigen Laborveränderungen beobachtet.

Bei einem Drittel der Studienteilnehmer fand sich bereits zu Baseline ein erhöhter Blutdruck, der im Studienverlauf weiter leicht anstieg. Der Anteil von Patienten, der die Definition einer Hypertonie erfüllte, betrug zu Baseline im Placebo- bzw. Atalurenarm 18,3 bzw. 24,3 %, nach 48 Wochen 20,9 bzw. 26,4 % (CSR, S. 109). Bei zwei Patienten im Atalurenarm zeigte sich im EKG zu Studienende eine Veränderung, die mit einer rechtsventrikulären Hypertrophie kompatibel war. Bei den Ultraschalluntersuchungen zeigte sich bei einem Patienten im Atalurenarm eine leicht vergrößerte Milz, die aber klinisch als nicht relevant eingestuft wurde.

Die EMA weist auf die Interaktion mit potentiell nephrotoxischen Substanzen hin, insbesondere auf intravenös applizierte Aminoglykoside, so dass diese Medikamente im Rahmen der Studie 007 nicht erlaubt waren (EPAR [5], S. 58). Außerdem bemerkt die EMA, dass keine Studien an Patienten mit renaler oder hepatischer Funktionseinschränkung durchgeführt wurden, so dass die Sicherheit bei diesen Patientenpopulationen nicht untersucht ist. Ataluren wird zu 50 % über die Niere ausgeschieden und kann daher bei renaler Funktionseinschränkung akkumulieren. Entsprechendes gilt für die hepatische Elimination (EPAR [5], S. 59).

2.5.2 Ergebnisse der Metaanalyse

Die Ergebnisse der im SAP erwähnten Metaanalyse finden sich nicht im CSR. Es wird daher auf die Darstellung im Modul 4 des Dossiers (Abschnitt 4.4.2.5) sowie auf die Common Technical Documents (CTD) 2.7.3 und ISE verwiesen.

Für den primären Endpunkt sowie bestimmte sekundäre Endpunkte sollten Metaanalysen der Ergebnisse der Studien 007 und 020 durchgeführt werden. Hierzu sollten die Subgruppen der „*ambulatory decline phase*“ herangezogen werden (SAP, S. 20), weil die Studienpopulationen hinsichtlich der Einschlusskriterien Alter (Studie 007 ≥ 5 J. vs. Studie 020 ≥ 7 bis ≤ 16 J.), Gehstrecke zu Baseline (Studie 007 ≥ 75 m vs. Studie 020 ≥ 150 m und ≤ 80 % des

altersbezogenen Erwartungswerts) und Kortikosteroideinnahme bei Studienbeginn (Studie 007: keine Bedingung vs. Studie 020: erforderliche Bedingung) nur bedingt vergleichbar sind.

Der pU zieht in seiner Metaanalyse allerdings nicht, wie im SAP vorgesehen, Patienten der post-hoc gebildeten Subgruppe der so genannten „ambulatory decline phase“ (ADP) aus der Studie 007 (Placebo N=31, Ataluren N=32) heran (sowie die komplette ITT-Population aus der Studie 020, die hinsichtlich der Einschlusskriterien der ADP-Population entspricht (vgl. Modul 4 der Studie 007, S. 105, EPAR [5] S. 42)), sondern die Patienten mit einer Gehstrecke zu Baseline von ≥ 300 bis < 400 m. Die post-hoc gebildete Subgruppe, auf die im EPAR verwiesen wird, bezog sich auf eine Gehstrecke von unter/über 350 m. Die Subgruppe aus der Studie 007 war insgesamt etwas älter, insbesondere weil der Anteil der über 9-Jährigen um rund 10 Prozentpunkte höher war (ca. 60 % vs. ca. 50 %). Dementsprechend waren die durchschnittlichen Werte für den 6MWT sowie der TFTs zu Baseline niedriger bzw. die Subgruppenverteilung war in der Population der Studie 007 prognostisch ungünstiger.

Entsprechend dem präspezifizierten Vorgehen werden die Daten der Studienpopulationen gepoolt, deren Gehstrecke im 6MWT zwischen 300 und 400 Metern liegt und damit der ADP im engeren Sinne entspricht. Daher wird hier nur die Metaanalyse dieser Subgruppe dargestellt. Diese Subpopulationen wurden im Sinne einer individuellen Patientendaten-Metaanalyse gepoolt, so dass 74 Patienten im Placebo- und 69 Patienten im Atalurenarm ausgewertet wurden (entsprechend 44 Patienten aus der Studie 007 [N=22 Placebo, N=22 Ataluren] und 99 Patienten aus der Studie 020 [N=52 Placebo, N=47 Ataluren]). Diese Subgruppen waren hinsichtlich demographischer Merkmale und weiterer Charakteristika gut vergleichbar (Modul 4, S. 197f.; CTD-ISE Tab. 2.1.1-2, S. 28; Tab. 2.1.2-2, S. 107).

Die Ergebnisse der Metaanalyse in den Subgruppen unter 300 bzw. über 400 Meter Gehstrecke im 6MWT zu Baseline waren nicht signifikant (vgl. Tab. 4-70 im Modul 4, S. 203).

In der Tabelle 18 finden sich die Ergebnisse der Metaanalyse für die Subgruppen der Studien 007 und 020 entsprechend der ADP.

Tabelle 18: Ergebnisse der Metaanalyse der Patienten aus den Studien 007 und 020 in der Subgruppe ≥ 300 bis < 400 Meter (ADP) im 6MWT im Vergleich zu Placebo

Endpunkt	Studie 007 Δ LS Mean, SE	Studie 020 Δ LS Mean, SE	Metaanalyse Δ LS Mean, SE/HR 95%-KI p-Wert
Absolute Veränderung im 6MWT zu Woche 48 (m) ¹⁾	48,7 (21,635)	42,9 (15,886)	44,9 (12,805) (19,836;70,031) p=0,0004
Zeit bis zu persistierender Verschlechterung um mind. 10 % ²⁾	HR 0,27	HR 0,79	HR 0,66 (0,39;1,12) p=0,1243
<i>Änderung in der proximalen Muskelfunktion (,Timed Function Tests‘) zu Woche 48¹⁾</i>			
10 m laufen/gehen (s)	-3,2 (1,543)	-1,8 (0,999)	-2,2 (0,839) (-3,877;-0,590) p=0,0077

Endpunkt	Studie 007 Δ LS Mean, SE	Studie 020 Δ LS Mean, SE	Metaanalyse Δ LS Mean, SE/HR 95%-KI p-Wert
benötigte Zeit, um 4 Treppenstufen hinaufzusteigen (s)	-3,4 (1,742)	-3,5 (1,150)	-3,4 (0,960) (-5,316;-1,554) p=0,0003
benötigte Zeit, um 4 Treppenstufen hinunterzusteigen (s)	-4,1 (1,870)	-4,4 (1,165)	-4,3 (0,989) (-6,223;-2,347) p=<0,0001

¹⁾ Quelle: CTD 2.7.3, S. 70, 77, CTD-ISE Tab. 3.1.2, S. 140

²⁾ Quelle: CTD 2.7.3, S. 70, CTD-ISE Tab. 3.2.1, S. 143

3 Methodische Anmerkungen und Bewertung der Unterlagen

3.1 Zulassungsstatus und Zulassungspopulation von Ataluren

Ataluren wurde von der EMA unter „besonderen Bedingungen“ (conditional approval) auf Antrag des pU im Mai 2014 für die Indikation „...*treatment of Duchenne muscular dystrophy resulting from a nonsense mutation in the dystrophin gene, in ambulatory patients aged 5 years and older*“ zugelassen. Die empfohlene tägliche Dosis ist 40 mg/kg KG aufgeteilt auf drei Einzeldosen (10, 10, 20 mg).

Die Zulassung unter “besonderen Bedingungen” war mit der Erwartung weiterer Studienergebnisse verknüpft: *“To complete a multicentre, randomised, double-blind, placebo-controlled confirmatory study to examine efficacy and safety of ataluren 10, 10, 20 mg/kg in patients with non-sense mutation Duchenne muscular dystrophy (Study PTC124-GD-020-DMD)”*. Dementsprechend war der Beschluss des G-BA bis zum 30.6.2016 befristet. Die Studienergebnisse der Studie 020 wurden dem G-BA fristgerecht zur Verfügung gestellt.

Bei der Studie 020 handelt es sich um eine Phase-3-Studie mit 230 Patienten, in der Ataluren 10/10/20 mg/kg KG mit Placebo verglichen wird. Primärer Endpunkt ist der Unterschied im 6MWT nach 48 Wochen. Im Unterschied zu Studie 007 ist das Alter der einzuschließenden Patienten mit 7–16 Jahren etwas höher, außerdem wurde die Mindestgehstrecke bei Einschluss auf 150 m und $\leq 80\%$ des nach Alter und Größe vorausgesagten Wertes festgelegt. Die Studie sollte Patienten in der so genannten *decline phase* erfassen, bei denen eine bessere Wirksamkeit von Ataluren erwartet wird.

Die folgenden Ausführungen beziehen sich auf die Phase-3-Studie 020 sowie auf eine Metaanalyse der Studien 007 und 020, deren Ergebnisse der pU im Dossier als Grundlage für seine Ausführungen zum Ausmaß des Zusatznutzens heranzieht.

3.2 Design und Methodik

Bei der Studie 020 handelt es sich um eine multizentrische, randomisierte, placebokontrollierte, doppelt verblindete Studie der Phase 3 im Parallelgruppendesign (1:1) mit zwei Behandlungsarmen (Ataluren 10/10/20 mg/kg KG vs. Placebo). Bestandteil der Studie war eine 2-wöchige Screening-Phase, eine 48-wöchige verblindete Behandlungsphase und eine 6-wöchige Nachbeobachtungsphase. Patienten wurden stratifiziert für die Faktoren Alter, Dauer der Einnahme von Kortikosteroiden und 6MWT-Wert zu Baseline. Das primäre Ziel der Studie war der Nachweis der Verbesserung der Gehfähigkeit im Vergleich zu Placebo anhand der 6-Minuten-Gehstrecke (6MWT).

Eingeschlossen wurden gehfähige Patienten (≥ 150 m Gehstrecke ohne Hilfe im 6MWT) im Alter ab 7 Jahren bis einschließlich 16 Jahre mit bestätigter Diagnose einer DMD (Nachweis einer Nonsense-Mutation durch Sequenzierung des Dystrophin-Gens). Die Studie wurde an 54 Zentren in den USA, Kanada, Australien, Südamerika, Europa, Südostasien und Israel durchgeführt. Die Randomisierung erfolgte durch ein IVR-/IWR-System. Entsprechend der vorgesehenen Stratifizierung wurde eine adaptive Randomisierung vorgenommen, um balancierte Gruppen zu gewährleisten. Patienten, Eltern, Studienpersonal und Angestellte von PTC Therapeutics waren verblindet bis zum Zeitpunkt des Datenbankschlusses. Der erste Patient wurde am 26.03.2013 eingeschlossen, die letzte Studienvisite fand am 20.08.2015 statt. Aufgrund des Designs und der

Methodik der Studie 020 wird von einem niedrigen Verzerrungspotential auf Studienebene ausgegangen.

Die Patientenpopulationen der Studien 007 und 020 unterscheiden sich hinsichtlich Mindestalter, Gehstrecke zu Baseline und Kortikosteroideinnahme (Studie 007 ≥ 5 J. vs. Studie 020 ≥ 7 bis ≤ 16 J., Gehstrecke zu Baseline Studie 007 ≥ 75 m vs. Studie 020 ≥ 150 m und ≤ 80 % des altersbezogenen Erwartungswerts, Kortikosteroideinnahme bei Studienbeginn Studie 007: keine Bedingung vs. Studie 020: erforderliche Bedingung). Daher wurden aus den beiden Studien die Daten der Patienten gepoolt, deren Gehstrecke zu Baseline im 6MWT zwischen 300 und 400 Metern lag und damit der ADP im engeren Sinne entsprach. Diese Subgruppen waren hinsichtlich demographischer Merkmale und weiterer Charakteristika gut vergleichbar.

3.3 Wirksamkeit

Morbidität

Primärer Endpunkt – Veränderung im 6MWT im Vergleich zu Baseline nach 48 Wochen

Der 6MWT wurde hinsichtlich seiner Validität und Patientenrelevanz bereits im Nutzenbewertungsverfahren vom 2.3.2015 bewertet. Zusammengefasst kann die Messung der Gehstrecke als patientenrelevanter Endpunkt und die Messung der Gehstrecke mit dem 6MWT als valide eingeschätzt werden. Die EMA-Guideline 2015 [4] empfiehlt die Erhebung des funktionellen Status als wichtigsten klinischen Endpunkt sowie ergänzend Aktivitäten des täglichen Lebens. Zwei Endpunkte sollten für die Domänen Muskelkraft und motorische Funktion erhoben werden, davon je einer als primärer und sekundärer Endpunkt. Der 6MWT wird als Erhebungsinstrument erwähnt, allerdings mit Vorbehalten bezogen auf einen möglichen Lerneffekt, inter- und intrapersonelle Variabilität und die Definition eines klinisch relevanten Effekts. Der 6MWT, wenn er als primärer Endpunkt ausgewählt wird, sollte in seiner Aussagekraft mit komplementären Tests, wie bspw. ‚Timed Function Tests‘, unterstützt werden. Ein Guideline-Entwurf der FDA hält den 6MWT für Kinder mit DMD ab ca. 5 Jahren ebenfalls für geeignet zur Erfassung von Muskelkraft und Ausdauer, verweist aber auf mögliche Probleme bei der Auswertung, wenn Patienten während der Studie ihre Gehfähigkeit verlieren und der 6MWT dann Nullwerte ausweist [16]. Dementsprechend wurde der 6MWT bei Verlust der Gehfähigkeit mit 0 bewertet.

Die gewählte Relevanzschwelle von 30 m ergibt sich aus verschiedenen Berechnungsansätzen (1/3 SD bzw. abgeleitet aus dem Standardfehler) und wird vom pU deutlich oberhalb der zu erwartenden Variabilität angesetzt. In einer Publikation von Yost & Eton (2005) [17] stimmte die MCID auf der Basis von 0,5 SD besser mit einem ankerbasierten Ansatz überein. Die Patientenrelevanz der vom pU gewählten MCID auf der Basis von 0,3 SD ist daher nicht abschließend einzuschätzen, entspricht aber eher einer moderaten Effektstärke (entsprechend dem Vorschlag Cohens). Bei der Anwendung von 0,5 SD wäre die MCID deutlich höher anzusetzen.

Der Nachweis der Überlegenheit für den primären Endpunkt wurde nicht erreicht. Der geschätzte Behandlungsunterschied zwischen dem Ataluren- und dem Placebo-Arm von 12,98 Metern (95%-KI -7,44;33,39) zugunsten von Ataluren war in der primären (präspezifizierten Analyse) statistisch nicht signifikant. Lediglich in der Subgruppe der Patienten mit einer Gehstrecke zu Baseline von

≥ 300 bis < 400 Meter zeigte sich ein signifikanter und klinisch relevanter Effekt von 42,89 (95%-KI 11,75;74,03) Metern ($p=0,007$), wobei das untere Konfidenzintervall allerdings unter der Relevanzschwelle von 30 Metern lag.

Weitere Analysen für den primären Endpunkt bezogen sich auf präspezifizierte Subgruppen (6MWT zu Baseline, Dauer der Kortikosteroidbehandlung und Altersgruppe) sowie den Anteil der Progressoren (Zeit bis zu einer persistierenden Verschlechterung um mindestens 10 %). Zudem wurde eine Sensitivitätsanalyse des Effekts der Allokation von Geschwisterpaaren in denselben Behandlungsarm durchgeführt.

Die Subgruppenanalysen des primären Endpunkts (Interaktionstests) zeigten einen signifikanten Wert für die Behandlung-Subgruppe-Interaktion ($p=0,05$). Für die präspezifizierte Subgruppe mit einer Gehstrecke zu Baseline ≥ 300 bis < 400 Meter lag der p-Wert bei 0,007. Für die anderen Subgruppen zeigte sich keine signifikante Interaktion mit Ausnahme der Altersgruppen (< 9 / ≥ 9 Jahre), hier lag der p-Wert für die Behandlung-Subgruppe-Interaktion bei $p=0,076$ mit einer Differenz der Gehstrecke von 16,26 (95%-KI -13,2;45,71) Metern für < 9-Jährige und 9,92 (95%-KI -18,48;38,31) Metern bei den ≥ 9-Jährigen. Die präspezifizierte Sensitivitätsanalyse ergab ein vergleichbares Ergebnis wie die Primäranalyse des primären Endpunkts.

Im Hinblick auf die bereits in der Studie 007 gezeigten Effekte bei den Patienten in der so genannten *ambulatory decline phase* (ADP) sind die Ergebnisse für die Subgruppe der Patienten mit einer Gehstrecke von ≥ 300 bis < 400 Meter plausibel. Auch die Ergebnisse der ergänzenden Metaanalyse zeigen ein damit konsistentes Ergebnis (Gehstrecke 44,9 [95%-KI 19,8;70,0] Meter) in dieser Subgruppe (vgl. Tab. 19). Demnach ist die Wirksamkeit von Ataluren nicht nachweisbar, wenn bereits zu viel Muskelgewebe im Krankheitsverlauf abgebaut ist (korreliert mit dem Alter und mit einer Gehstrecke zu Baseline < 300 Meter). Besteht eine Gehstrecke von ≥ 400 Metern zu Baseline, dann bleibt diese Gehstrecke im kommenden Jahr auch ohne Therapie weitgehend stabil, so dass ebenfalls ein Effekt von Ataluren nicht im Rahmen des gewählten Studiendesigns nachweisbar ist.

Sekundäre Endpunkte

Zu den sekundären Endpunkten gehören die Zeit bis zu einer mindestens 10%igen Verschlechterung im 6MWT nach 48 Wochen (Progressoren) sowie Timed Function Tests (TFT), die die Messung der Muskelfunktion ermöglichen. Die Messung der Alltagsaktivitäten mittels des NSAA wurde als explorativer Endpunkt mitgeführt.

Zeit bis zu einer mindestens 10%igen Verschlechterung im 6MWT nach 48 Wochen (Progressoren)

- In Studien hat sich gezeigt, dass eine Verschlechterung um 10 % im 6MWT mit einem signifikant erhöhten Risiko des Gehverlusts in den folgenden 4 Jahren einhergeht, d. h. dass dieser Endpunkt ergänzend zur absoluten Veränderung im 6MWT herangezogen werden kann [9]. In der Progressorenanalyse zeigte sich im Gegensatz zu der Subgruppenanalyse des primären Endpunkts ein signifikanter Unterschied in der Subgruppe mit < 300 Meter im 6MWT zu Baseline zugunsten von Ataluren. Die Verschlechterung der Gehstrecke trat in der Atalurengruppe nach im Median 164 (95%-KI 1,0;225,0) Tagen ein, in der Placebogruppe bereits nach 56 Tagen (95%-KI 1,0;111,0), HR 0,48 (95%-KI 0,24;0,93), $p=0,0306$. Die Bedeutung dieses Ergebnisses im Kontext der Überlegungen zur Wirkungsweise von Ataluren ist unklar. Der pU führt in seinem Dossier

aus, dass „*der deutliche Behandlungseffekt von Ataluren bei den Patienten mit einer Gehstrecke von ≥ 300 bis < 400 m verglichen mit den Patienten am oberen bzw. unteren Spektrum der Erkrankung eine höhere Sensitivität gegenüber den Testsystemen widerspiegelt*“; aufgrund des Wirkmechanismus von Ataluren sei eine „*Wirksamkeit über das gesamte Krankheitsspektrum zu erwarten*“ (Modul 4, S. 225).

Änderung in der proximalen Muskelfunktion („Timed Function Tests“) zu Woche 48

- Die Messung der proximalen Muskelfunktion in TFTs wurde in drei Testsituationen aufgeteilt, d. h. die Patienten sollten so schnell wie möglich (innerhalb von max. 30 Sekunden) 10 m laufen/gehen und vier Treppenstufen hinauf- bzw. hinuntersteigen. Diese Funktionstests erfassen Fähigkeiten der Patienten, deren Verlust gemäß einem internationalen Expertenkonsens als „Meilensteine“ und ausreichend sensitive Prädiktoren des Krankheitsverlaufs angesehen werden (McDonald et al., 2013 [9], Lynn et al., 2015 [7]).
- Die Ergebnisse zeigten eine statistisch nicht signifikante Tendenz (Differenz -1,4 Sekunden [95%-KI -2,9;0,05], $p=0,058$) zugunsten von Ataluren für die benötigte Zeit, um 4 Treppenstufen hinaufzusteigen, und einen statistisch signifikanten Unterschied zugunsten von Ataluren für die benötigte Zeit, um 4 Treppenstufen hinunterzusteigen (Differenz -1,97 Sekunden [95%-KI -3,52;-0,43], $p=0,012$). Dieses Ergebnis liegt oberhalb der vom pU angegebenen klinischen Relevanzschwelle von 1,5 Sekunden. Auch in der Metaanalyse zeigten sich, in der ADP-Subgruppe, in den TFTs statistisch signifikante und klinisch relevante Effekte zugunsten der Atalurengruppe. Die Gültigkeit der vom pU angegebenen Relevanzschwelle kann allerdings nicht abschließend eingeschätzt werden.

Explorative Endpunkte

Als explorativer Endpunkt wurde das *North Star Ambulatory Assessment* (NSAA) eingesetzt. Beim NSAA handelt es sich um ein Instrument zur Ermittlung körperlicher Funktionen, das 17 Komponenten enthält, die verschiedene kurzzeitige Alltagsaktivitäten messen (bspw. Aufstehen, Gehen, Stehen) und vom pU als komplementär zum 6MWT angesehen wird, der eher die Ausdauer misst. Das Instrument wurde vom *North Star Clinical Network for Paediatric Neuromuscular Disease Management* in England entwickelt und hinsichtlich Reliabilität validiert. Der Score (max. 34 Punkte) wurde im Rahmen einer separaten Studie auf eine Skala von 0–100 linearisiert, um die zu den Scoreunterschieden korrespondierenden klinischen Zustände in metrische Punktabstände abzubilden. Die MCID wird mit 7–9 Punkte auf der linearisierten Skala angegeben. Die Differenz in der Auswertung betrug 1,51 Punkte (95%-KI -1,16;4,17, $p=0,268$). Da nach Angaben des pU ein Scoreunterschied von 10 Punkten (linearisiert) in etwa mit dem Verlust einer Funktion korrespondiert, ist der geringe Unterschied in der gewählten klinischen Population plausibel; während der Studie verloren relativ wenige Patienten ihre Gehfähigkeit (insgesamt 23 Patienten, davon 9 im Ataluren- und 14 im Placeboarm), d. h. die Änderungen im Krankheitsverlauf in der Studie 020 wurden durch den NSAA offenbar nicht erfasst.

3.3.1 Zusammenfassende Einschätzung zur Wirksamkeit

Das Verzerrungspotential für die Erhebung der Endpunkte ist (bezogen auf die präspezifizierten Auswertungen) insgesamt als niedrig einzuschätzen.

Der MCID für den 6MWT von 30 m kann grundsätzlich gefolgt werden. Für den primären Endpunkt ergibt sich kein statistisch signifikanter Effekt. Lediglich in der Subgruppe der Patienten mit einer Gehstrecke zu Baseline von ≥ 300 bis < 400 Meter zeigte sich sowohl in der Studie 020 wie auch in der Metaanalyse der Studien 007 und 020 ein statistisch signifikanter und klinisch relevanter Effekt von etwa 44 Metern, wobei das untere Konfidenzintervall jeweils unter der Relevanzschwelle lag. Für den Anteil der Progressoren (= Zeit bis zu einer mindestens 10%igen Verschlechterung nach 48 Wochen) zeigte sich ein statistisch signifikanter Vorteil in der Subgruppe mit einer Gehstrecke von < 300 Meter zu Baseline. Zu diesem Befund findet sich kein Erklärungsansatz in den Unterlagen des pU.

Die Ergebnisse der TFTs zeigten tendenziell einen Vorteil für die Atalurengruppe. Im NSAA zeigte sich kein Gruppenunterschied.

3.4 Lebensqualität

Für die Erfassung der Lebensqualität wurde das *Pediatric Outcomes Data Collection Instrument* (PODCI) verwendet mit den Domänen Transfer/Basismobilität und Sport / körperliche Funktion. Diese beiden Domänen zeigten in einer Studie eine ausreichende Korrelation mit dem 6MWT (McDonald et al., 2010 [10]). Aufgrund fehlender Angaben können die Qualität und Patientenrelevanz des Instruments für die vorliegende Patientenpopulation nicht abschließend eingeschätzt werden.

Für keine der beiden ausgewerteten Domänen fand sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

3.5 Sicherheit

Die Erhebung unerwünschter Ereignisse (UE) und schwerwiegender unerwünschter Ereignisse (SUE) erfolgte gemäß des Medical Dictionary for Regulatory Activities (MedDRA) unter Verwendung der Systemorganklasse (SOC) und dem Preferred Term (PT) und wurde gemäß CTCAE Version 3.0 klassifiziert. Begleitmedikamente wurden mittels WHO DRUG Dictionary kodiert. Besonderes Augenmerk wurde auf die Überprüfung der Serumwerte für Kreatinin, Harnstoff und Cystatin C sowie auf das nephrotoxische Potential (in Kombination mit anderen Arzneimitteln) von Ataluren gelegt. Daher wurde ein zusätzliches Nierenmonitoring durchgeführt. Außerdem wurde ein zusätzliches Lipidprofil erhoben, da sich in der Studie 007 erhöhte Lipidwerte gezeigt hatten. Für die Einschätzung der Sicherheit wurden Daten der As-treated-Population ausgewertet. Das Verzerrungspotential des patientenrelevanten Endpunkts UE wird als niedrig eingeschätzt.

Das Nebenwirkungsprofil war zwischen den Behandlungsarmen vergleichbar. Mindestens ein UE trat bei 89,6 % der Patienten in der Ataluren- und bei 87,8 % der Patienten in der Placebogruppe auf. Etwa die Hälfte der UE in jeder Gruppe hatten einen Schweregrad 1, jeweils etwa ein Drittel Schweregrad 2 und 6,1 % (Atalurengruppe) bzw. 7,8 % (Placebogruppe) ein Grad-3-UE. Je ein Patient pro Behandlungsarm brach die Therapie vorzeitig aufgrund unerwünschter CTCAE-Grad-2-Ereignisse ab (Obstipation bei einem Patienten in der Ataluren- und Krankheitsprogression bei einem Patienten in der Placebogruppe). Es gab keine lebensbedrohlichen (Grad 4) UE bzw. Todesfälle.

Die häufigsten UE waren Nasopharyngitis (20,9 im Ataluren- bzw. 19,1 % im Placeboarm), Erbrechen (22,6 bzw. 18,3 %) und Durchfall (17,4 bzw. 8,7 %), Stürze (18,3 bzw. 17,4 %, Extremitäten- und Rückenschmerzen (8,7/9,6 bzw. 12,2/7,0 %), Husten (16,5 bzw. 11,3 %), Fieber (13,9 bzw. 10,4 %) und Kopfschmerzen (jeweils 18,3 %), jeweils ohne signifikante Unterschiede zwischen den Gruppen. Hinsichtlich Dosisunterbrechungen oder Dosisreduzierungen zeigten sich keine signifikanten Gruppenunterschiede. Renale oder hepatische behandlungsassoziierte Ereignisse traten häufiger in der Ataluren- als in der Placebogruppe auf, insbesondere Hämaturie (Placebo 1 vs. Ataluren 7 Ereignisse) und periphere Ödeme (Placebo 0 vs. Ataluren 3 Ereignisse, davon je ein moderates bzw. schwerwiegendes). Allerdings zeigte sich kein Hinweis auf zunehmende Toxizität im Behandlungsverlauf (Erhebung in 8-Wochenintervallen).

Schwerwiegende UE traten bei insgesamt 8 Patienten auf, jeweils bei 4 Patienten in der Ataluren- und der Placebogruppe.

UE von besonderem Interesse waren Leberenzyme, Serum-Lipide und Nierenwerte. Die gemessenen Profile entsprachen den im Krankheitsverlauf zu erwartenden Werten. Insgesamt wurden keine klinisch relevanten, behandlungsbedürftigen Laborveränderungen beobachtet.

3.5.1 Zusammenfassende Einschätzung zur Sicherheit

Insgesamt ist Ataluren als gut verträglich einzuschätzen. Nebenwirkungen treten v. a. im Gastrointestinaltrakt auf, sind aber nach Einschätzung der EMA in der Regel beherrschbar. In Kombination mit Kortikosteroiden sind höhere Blutfettwerte und Blutdruckanstiege zu erwarten. Bei Patienten mit eingeschränkter Leber- und Nierenfunktion liegen nur begrenzte Erfahrungen vor, so dass diese Patienten sorgfältig beobachtet werden sollten. Interaktionen mit potentiell nephrotoxischen Substanzen sind zu beachten.

4 Anforderungen an eine qualitätsgesicherte Anwendung

Die Vorgaben der Fachinformation sind zu berücksichtigen.

Die europäische Zulassungsbehörde EMA stellt die Inhalte der Fachinformation zu Translarna® (Wirkstoff: Ataluren) unter folgendem Link frei zugänglich zur Verfügung (letzter Zugriff: 15. August 2016):

- http://www.ema.europa.eu/docs/de_DE/document_library/EPAR_-_Product_Information/human/002720/WC500171813.pdf

Die Einleitung und Überwachung der Behandlung mit Ataluren (Translarna®) soll durch in der Therapie von Patienten mit Duchenne-/Becker-Muskeldystrophie erfahrene Fachärzte erfolgen.

5 Zusammenfassung der Bewertung

Ataluren ermöglicht ein ribosomales Durchlesen eines überflüssigen Stop-Codons infolge einer Nonsense-Mutation, so dass funktionsfähiges Dystrophin transkribiert wird. Ataluren ist für das Anwendungsgebiet Muskeldystrophie vom Typ Duchenne infolge einer Nonsense-Mutation im Dystrophin-Gen bei gehfähigen Patienten im Alter ab 5 Jahren zugelassen. Die empfohlene Gesamtdosis ist 40 mg pro kg Körpergewicht pro Tag, aufgeteilt auf drei Dosen (10/10/20 mg/kg KG) im Abstand von 6, 6 und 12 Stunden.

Die Nutzenbewertung von Ataluren basiert auf der Studie PTC124-GD-020-DMD sowie einer ergänzenden Metaanalyse basierend auf den Ergebnissen der zulassungsbegründenden Studie 007 und der Studie 020. Bei der Studie 020 handelt es sich um eine multizentrische, randomisierte, placebokontrollierte, verblindete Studie der Phase 3 im Parallelgruppendesign mit zwei Behandlungsarmen. In die Studie wurden Patienten ab einem Alter von 7 Jahren mit DMD und einer Gehstrecke von ≥ 150 m ohne Hilfe und ≤ 80 % des nach Alter und Größe vorausgesagten Wertes im Screening-6MWT eingeschlossen. Aufgrund des Designs und der Methodik wird das Verzerrungspotential auf Studienebene für die präspezifizierte Auswertung insgesamt als niedrig eingeschätzt.

Die Ergebnisse der Studie 020 und der Metaanalyse werden in der folgenden Tabelle unter Berücksichtigung des Verzerrungspotentials zusammengefasst. Es wird dargestellt, ob unter Ataluren 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.

Tabelle 19: Zusammenfassende Darstellung der Ergebnisse der Studie 020 (ITT-Population, präspezifizierte Auswertung) sowie der Metaanalyse (Studien 007 und 020, ADP-Population)

Endpunkt	Ataluren 10/10/20 mg/kg KG vs. Placebo	Metaanalyse ¹⁾	Effekt
Mortalität			
Gesamtmortalität	nicht erhoben		
Morbidität			
Primärer Endpunkt			
Veränderung im 6MWT im Vergleich zu Baseline nach 48 Wochen (m) <i>alle Patienten</i>	12,98 (95%-KI -7,44;33,39) p=0,213		\leftrightarrow
<i>Subgruppenanalyse</i>			
< 300 Meter	-7,71 (95%-KI -54,93;39,51) p=0,749		\leftrightarrow
≥ 300 bis < 400 Meter	42,89 (95%-KI 11,75;74,03) p=0,007	44,9 (95%-KI 19,84;70,03) p=0,0004	$\uparrow\uparrow$

Endpunkt	Ataluren 10/10/20 mg/kg KG vs. Placebo	Metaanalyse ¹⁾	Effekt
≥ 400 Meter	-9,51 (95%-KI -43,19;24,18) p=0,58		↔
< 350 Meter	21,89 (95%-KI -12,32;56,11) p=0,21		↔
≥ 350 Meter	7,98 (95%-KI -1,53;33,48) p=0,54		↔
sekundäre Endpunkte			
Zeit bis zu einer persistierenden Verschlechterung um mindestens 10 % <i>alle Patienten</i>	HR 0,75 (95%-KI 0,51;1,12) p=0,1603		↔
<i>Subgruppenanalyse</i>			
< 300 Meter	HR 0,48 (95%-KI 0,24;0,93) p=0,0306		↑↑
≥ 300 bis < 400 Meter	HR 0,79 (95%-KI 0,44;1,41) p=0,4181	HR 0,66 (95%-KI 0,39;1,12) p=0,1243	↔
≥ 400 Meter	HR 1,52 (95%-KI 0,59;3,91) p=0,3872		
Änderung in der proximalen Muskelfunktion (,Timed Function Tests') zu Woche 48 ²⁾			
<i>10 m laufen/gehen (Sekunden, Mittelwert)</i>	Differenz -1,07 (95%-KI -2,4;0,27) p=0,117		↔
Patienten in der ADP (Gehstrecke im 6MWT zu Baseline ≥ 300 bis < 400 Meter)		Differenz -2,2 (95%-KI -3,877;-0,590) p=0,0077	↑↑
<i>benötigte Zeit, um 4 Treppenstufen hinaufzusteigen (Sekunden, Mittelwert)</i>	Differenz -1,4 (95%-KI -2,9;0,05) p=0,058		↔
Patienten in der ADP (Gehstrecke im 6MWT zu Baseline ≥ 300 bis < 400 Meter)		Differenz -3,4 (95%-KI -5,316;-1,554) p=0,0003	↑↑
<i>benötigte Zeit, um 4 Treppenstufen hinunterzusteigen (Sekunden, Mittelwert)</i>	Differenz -1,97 (95%-KI -3,52;-0,43) p=0,012		↑↑
Patienten in der ADP (Gehstrecke im 6MWT zu Baseline ≥ 300 bis		Differenz -4,3 (95%-KI -6,223;-2,347)	↑↑

Dossierbewertung für Orphan Drugs

Ataluren (Neubewertung nach Fristablauf)

Endpunkt	Ataluren 10/10/20 mg/kg KG vs. Placebo	Metaanalyse ¹⁾	Effekt
< 400 Meter)		p=<0,0001	
Explorative Endpunkte			
North Star Ambulatory Assessment (NSAA) Gesamtscore	Differenz 0,80 (95%-KI -0,23;1,82) p=0,128		↔
linearisierter Score	Differenz 1,51 (95%-KI -1,16;4,17) p=0,268		↔
Lebensqualität			
Änderung in der Lebensqualität gemessen mit dem <i>Pediatric Outcomes Data Collection Instrument</i> (PODCI)			
Domäne Transfer/Basismobilität (Baselinescore, Mittelwert)	1,64 (95%-KI -2,11;5,39) p=0,391		↔
Domäne Sport / körperliche Funktion (Baselinescore, Mittelwert)	2,15 (95%-KI -1,75;6,05) p=0,281		↔
Unerwünschte Ereignisse			
mindestens 1 UE, N (%)	103 (89,6) vs. 101 (87,8)		↔
mindestens 1 SUE, N (%)	4 (3,5) vs. 4 (3,5)		↔
Therapieabbruch aufgrund von UE, N (%)	1 (0,9) vs. 1 (0,9)		↔
UE Grad 3 oder höher, N (%)	9 (7,8) vs. 7 (6,1)		↔

¹⁾ Präspezifizierte Metaanalyse der Studie 007 und 020, nur in der Subgruppe der so genannten *ambulatory decline phase*, d. h. mit einer Gehstrecke zu Baseline im 6MWT von ≥ 300 bis < 400 Meter.

²⁾ Für diese Endpunkte wurden in der Studie 020 keine Subgruppenanalysen präspezifiziert.

Referenzen

1. **Bushby K, Connor E.** Clinical outcome measures for trials in Duchenne muscular dystrophy: report from International Working Group meetings. Clin Investig (Lond) 2011;1(9):1217-1235.
2. **Bushby K, Finkel R, Birnkrant DJ, Case LE, Clemens PR, Cripe L, et al.** Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. Lancet Neurol 2010;9(1):77-93.
3. **Escolar DM, Hache LP, Clemens PR, Cnaan A, McDonald CM, Viswanathan V, et al.** Randomized, blinded trial of weekend vs daily prednisone in Duchenne muscular dystrophy. Neurology 2011;77(5):444-452.
4. **European Medicines Agency (EMA).** Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy; EMA/CHMP/236981-2011, Corr. 1 [online]. 17.12.2015. London (GBR): EMA. [Zugriff: 14.07.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500199239.pdf.
5. **European Medicines Agency (EMA).** Translarna (ataluren): EPAR public assessment report EMEA/H/C/002720 [online]. 22.05.2014. London (GBR): EMA. [Zugriff: 23.06.2016]. URL: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002720/WC500171816.pdf.
6. **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.
7. **Lynn S, Aartsma-Rus A, Bushby K, Furlong P, Goemans N, De Luca A, et al.** Measuring clinical effectiveness of medicinal products for the treatment of Duchenne muscular dystrophy. Neuromuscul Disord 2015;25(1):96-105.
8. **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):1046-1052.
9. **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):357-368.
10. **McDonald CM, McDonald DA, Bagley A, Sienko Thomas S, Buckon CE, Henricson E, et al.** Relationship between clinical outcome measures and parent proxy reports of health-related quality of life in ambulatory children with Duchenne muscular dystrophy. J Child Neurol 2010;25(9):1130-1144.
11. **PTC Therapeutics.** Ataluren (Translarna): Dossier zur Nutzenbewertung gemäß § 35a SGB V; Modul 4A. PTC Therapeutics International Limited 30.05.2016.

12. **PTC Therapeutics.** Clinical Protocol. Protocol Number PTC124-GD-020-DMD. Version 2.0. PTC Therapeutics International Limited 2014.
13. **PTC Therapeutics.** Clinical Study Report. PTC124-GD-020-DMD. PTC Therapeutics International Limited 2015.
14. **PTC Therapeutics.** Statistical Analysis Plan. Protocol Number PTC124-GD-020-DMD. Version 1.0. PTC Therapeutics International Limited; 2015.
15. **Sarnat HB.** Muscular Dystrophies. In: Kliegman RM, Behrman RE, Jenson HB, Stanton BF (Eds). Nelson Textbook of pediatrics. Philadelphia: Saunders; 2007. S. 2540-2550.
16. **U.S. Food and Drug Administration (FDA).** Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment: Guidance for Industry [online]. 06.2015. Silver Spring (USA): FDA. [Zugriff: 08.08.2016]. URL: <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/UCM450229.pdf>.
17. **Yost KJ, Eton DT.** Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof 2005;28(2):172-191.