

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

Datum der Veröffentlichung: 2. März 2015

Inhalt

Tabellenverzeichnis.....	3
Abkürzungsverzeichnis.....	4
Hintergrund	6
1 Einführung	8
2 Nutzenbewertung.....	10
2.1 Fragestellung	10
2.2 Zulassungsbegründende Studie	10
2.3 Liste der verwendeten Quellen	10
2.4 Studiencharakteristika	11
2.5 Endpunkte und Erhebungsinstrumente.....	18
2.6 Ergebnisse zum Zusatznutzen	30
2.7 Einschätzung zum Ausmaß des Zusatznutzens durch den pU	47
3 Methodische Anmerkungen und Bewertung der Unterlagen	48
3.1 Design und Methodik.....	49
3.2 Endpunkte zur Wirksamkeit.....	51
3.3 Lebensqualität.....	56
3.4 Sicherheit.....	57
4 Anforderungen an eine qualitätsgesicherte Anwendung	59
5 Zusammenfassung der Bewertung	60
Referenzen.....	63

Tabellenverzeichnis

Tabelle 1: Charakterisierung der Studie 007

Tabelle 2: Charakteristika der Studienpopulation in der Studie 007

Tabelle 3: Charakteristika der Studienpopulation nach Stratifizierungsfaktoren
in der Studie 007

Tabelle 4: Charakterisierung der Intervention in der Studie 007

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

Tabelle 6: Verzerrungspotential der Endpunkte in der Studie 007

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 007 (ITT-Population)

Tabelle 11: Ergebnisse zu sekundären Endpunkten in der Studie 007 (ITT-Population)

Tabelle 12: Ergebnisse zur Lebensqualität in der Studie 007 (ITT-Population)

Tabelle 13: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE in Studie 007

Tabelle 14: UE in der Studie 007 unabhängig von der Kausalität, die bei > 5% der Patienten
auftraten (as-treated-Population)

Tabelle 15: UE Grad 3 in der Studie 007 (as-treated-Population)

Tabelle 16: Mit der Behandlung in möglich oder wahrscheinlich in Zusammenhang stehende
UE in der Studie 007, die bei > 1 Patienten auftraten (as-treated-Population)

Tabelle 17: Schwerwiegende UE in der Studie 007 (as-treated-Population)

Tabelle 18: Zusammenfassende Darstellung der Studienergebnisse
(ITT-Population, präspezifizierte Auswertung)

Abkürzungsverzeichnis

6MWD	6-Minuten-Gehstrecke (6-minutes walk distance)
6MWT	6-Minuten-Gehtest
AM-NutzenV	Arzneimittelnutzenbewertungsverordnung
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
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
kITT	korrigierte Intention-to-Treat-Analyse
LOCF	last observation carried forward
LS	Least Square
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
PedsQL	Pediatric Quality of Life Inventory
PT	Preferred Term

pU	Pharmazeutischer Unternehmer
SAM	Step Activity Monitoring
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), 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 24. Februar 2015 zur Kenntnis genommen. Darüber hinaus wurde mit Schreiben vom 27. November 2014 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 2. März 2015 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 (Kliegman et al. 2007). 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 wird mit 1 pro 3.600 bis 1 pro 6.000 lebendgeborenen Jungen angegeben (Bushby et al. 2010).

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. 2011). 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, Treppengehen und vom Boden aufstehen ist 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 DNA-Mutationsanalyse kann die Diagnose bestätigt werden.

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

Eine spezifische Therapieoption zur Behandlung der DMD existiert bisher nicht. Die Behandlung ist symptomatisch und richtet sich auf den längstmöglichen Funktionserhalt, 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. In die Bewertung werden gemäß Zulassungsstatus folgende Patienten in der folgenden Indikation eingeschlossen:

- Muskeldystrophie vom Typ Duchenne infolge einer Nonsense-Mutation im Dystrophin-Gen bei gehfähigen Patienten im Alter ab 5 Jahren

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.

2.2 Zulassungsbegründende Studie

Pivotal Studie

- PTC124-GD-007-DMD (im Folgenden Studie 007): *“Placebo-controlled, double-blind, Phase 2b, evaluation of efficacy and safety of ataluren administered for 48 weeks in males with nmDBMD ≥5 years”*. Es handelt sich um eine placebokontrollierte, doppelblinde kontrollierte Studie der Phase 2b mit dem Ziel, die Verbesserung der Gehfähigkeit im Vergleich zu Placebo anhand der 6-Minuten-Gehstrecke (6MWD) nachzuweisen.

2.3 Liste der verwendeten Quellen

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

- Studienbericht (CSR) inkl. Studienprotokoll und statistischem Analyseplan (SAP) der Studie PTC124-GD-007-DMD.
- Unterlagen der Zulassungsbehörde, insbesondere Committee for Medicinal Products for Human Use Assessment Report (CHMP 2014).
- Herstellerdossier zu Ataluren.

2.4 Studiencharakteristika

Die Unterlagen zum Nachweis des Ausmaßes des Zusatznutzens für Ataluren basieren auf der Zulassungsstudie PTC124-GD-007-DMD (Studie 007). Studiendesign, Baseline-Charakteristika und Interventionen werden in den Tabellen 1 bis 4 dargestellt.

Tabelle 1: Charakterisierung der Studie 007

Charakteristikum	Beschreibung
Design	<p>Es handelt sich um eine, multizentrische, randomisierte placebokontrollierte, verblindete Dosisfindungsstudie der Phase 2b im Parallelgruppendesign (1:1:1) mit drei Behandlungsarmen. Bestandteil der Studie war eine 6-wöchige Screening-Phase, eine 48-wöchige verblindete Behandlungsphase und eine 6-wöchige Nachbeobachtungsphase. Patienten wurden stratifiziert für die Faktoren Alter, Einnahme von Kortikosteroiden und dem 6MWT-Wert zu Baseline.</p> <p>In den beiden Interventionsarmen wurde eine tägliche Gesamtdosis von 40 mg/kg Körpergewicht (KG) bzw. 80 mg/kg KG, jeweils aufgeteilt auf drei Dosen, eingesetzt. Das primäre Ziel war der Nachweis einer Verbesserung der Gehfähigkeit im Vergleich zu Placebo anhand der 6-Minuten-Gehstrecke (6MWD). Weitere Ziele waren die Untersuchung des Effekts von Ataluren auf die klinischen Manifestationen und die Pathophysiologie der DMD.</p> <p>Es waren zwei Interimsanalysen 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, ≥ 5 Jahre • für DMD typische Symptome, erhöhte CK, Probleme beim Gehen <ul style="list-style-type: none"> ○ zum Zeitpunkt der Randomisierung war keine Muskelbiopsie, keine Differenzierung zwischen DMD und BMD (Becker-MD, mildere Form) erforderlich • 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 • Fähigkeit, ≥ 75 m ohne Hilfe zu gehen im Screening-6MWT • 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, Restriktion und Prozeduren während der Studie <p>Ausschlusskriterien</p> <p>Patienten, die eines der nachgenannten Kriterien erfüllen, waren für die Studie ungeeignet:</p> <ul style="list-style-type: none"> • Behandlung mit Aminoglykosiden < 3 Monate vor Studienbeginn • Beginn der Behandlung mit systemischen Glukokortikoiden < 6 Monate vor

Charakteristikum	Beschreibung						
	<p>Studienbeginn oder nicht-adaptive Dosiserhöhung (z. B. Dosisanpassung aufgrund von Wachstum und Gewichtszunahme)</p> <ul style="list-style-type: none"> • Änderung bei Prophylaxe/Therapie aufgrund einer Herzinsuffizienz < 3 Monate vor Studienbeginn • Behandlung mit Warfarin < 1 Monat vor Studienbeginn • vorherige Therapie mit Ataluren • bekannte Hypersensitivität für einen der Inhalts- oder Hilfsstoffe der Studienmedikation • Einnahme einer anderen Prüfmedikation < 2 Monate vor Studienbeginn • größerer chirurgischer Eingriff < 30 Tage vor Studienbeginn • laufende immunsuppressive Therapie (außer Glukokortikoide) • gleichzeitige Teilnahme an einer anderen klinischen Therapiestudie • zu erwartender größerer chirurgischer Eingriff während der 12 Monate Behandlungsphase (z. B. aufgrund einer Skoliose) • tagsüber erforderliche Atemhilfe • 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" data-bbox="454 1115 1364 1205"> <thead> <tr> <th data-bbox="454 1115 758 1160">Ataluren 10/10/20 mg</th> <th data-bbox="758 1115 1061 1160">Ataluren 20/20/40 mg</th> <th data-bbox="1061 1115 1364 1160">Placebo</th> </tr> </thead> <tbody> <tr> <td data-bbox="454 1160 758 1205">N=57</td> <td data-bbox="758 1160 1061 1205">N=60</td> <td data-bbox="1061 1160 1364 1205">N=57</td> </tr> </tbody> </table> <p>Insgesamt wurden 185 Patienten für den Studieneinschluss evaluiert. Davon wurden 174 Patienten randomisiert. Ein Patient in der Ataluren-20/20/40-Gruppe schied nach 6 Wochen aufgrund von Non-Compliance aus.</p> <p><i>Anzahl einzuschließender Patienten je Gruppe:</i></p> <p>Auf Basis einer Pilotstudie wurde von einer durchschnittlichen Gehstrecke von 268 ± 92 m im 6MWT ausgegangen. Eine Verbesserung von 10-15 % wurde als klinisch relevant angesehen. Hieraus leitete sich eine angestrebte Verbesserung im 6MWT von 30 m ab (ca. 10 %). Die SD wurde mit 50 m angenommen. Daraus ergab sich ein Bedarf für 150 Patienten (50 je Studienarm), um eine Differenz von 30 m Gehstrecke im 6MWT mit einer Power von > 85 % im zweiseitigen Dunnett-t-Test auf einem Signifikanzniveau von $p=0,042$ nachzuweisen. Um einen vorzeitigen Studienabbruch von 10 % der Teilnehmer zu kompensieren, sollten bis zu 165 Patienten (55 je Studienarm) rekrutiert werden.</p>	Ataluren 10/10/20 mg	Ataluren 20/20/40 mg	Placebo	N=57	N=60	N=57
Ataluren 10/10/20 mg	Ataluren 20/20/40 mg	Placebo					
N=57	N=60	N=57					
Ort und Zeitraum der Durchführung	<p>Studiendauer: 28. Februar 2008 (erster Patient zur ersten Visite) bis 17. Dezember 2009 (letzter Patient zur letzten Visite).</p> <p>Studiendauer: 48 Wochen.</p> <p>Die Studie wurde in 37 Zentren in 11 Ländern durchgeführt: USA (N=78), Kanada (N=10), Europa (Belgien 5, Deutschland 11, England 21, Frankreich 8, Italien 13, Schweden 8, Spanien 6), Israel (N=3), Australien (N=11).</p>						
Primärer Endpunkt, sekundäre Endpunkte und	<p>Primärer Endpunkt</p> <ul style="list-style-type: none"> • 6-Minuten-Gehtest (6MWT) im Vergleich zu Baseline nach 48 Wochen. <ul style="list-style-type: none"> ○ unterstützend wurde der Anteil der Patienten mit einer mindestens 10%igen Verbesserung nach 48 Wochen (Responder) bzw. einer mind. 10%igen 						

Charakteristikum	Beschreibung
explorative Endpunkte gemäß Studienprotokoll	<p>Verschlechterung nach 48 Wochen (Progressoren) sowie Zeit bis zur 10%igen Verbesserung/Verschlechterung kalkuliert</p> <p>Sekundäre Endpunkte</p> <ul style="list-style-type: none"> • Änderung in der proximalen Muskelfunktion in ‚Timed Function Tests‘, d. h. so schnell wie möglich <ul style="list-style-type: none"> ○ Aufstehen aus Rückenlage ○ 10 m laufen/gehen ○ Treppenstufen hinauf-/hinuntersteigen • Änderung der Muskelkraft gemessen mittels Myometrie während <ul style="list-style-type: none"> ○ Knieflexion bzw. -extension ○ Ellenbogenflexion bzw. -extension ○ Schulterabduktion • Änderung der Alltagsaktivitäten gemessen mittels ‚Step Activity Monitoring‘ (Pedometer am Knöchel) an ≥ 9 Tagen nach jeder Studienvisite <ul style="list-style-type: none"> ○ mittlere Aktivitätsperiode/Tag/Visite ○ mittlere Anzahl Schritte je Tag/Visite während der aktiven Perioden ○ mittlere Anzahl Schritte pro Stunde während der aktiven Perioden an den Tagen der Studienvisite ○ maximale Anzahl an Schritten während kontinuierlicher Aktivitätsphasen (je 10, 20, 30, 60 Minuten) ○ Zeitanteil während der aktiven Perioden, der mit geringer Aktivität (≤ 15 Schritte pro Minute) verbracht wurde ○ Zeitanteil während der aktiven Perioden, der mit mittlerer Aktivität (16-30 Schritte pro Minute) verbracht wurde ○ Zeitanteil während der aktiven Perioden, der mit hoher Aktivität (> 30 Schritte pro Minute) verbracht wurde • patientenberichtete Endpunkte <ul style="list-style-type: none"> ○ Änderung in der Lebensqualität gemessen mit dem PedsQL (physikalische, emotionale, soziale, schulische Subskalen) sowie PedsQL Fatigue Scale ○ Änderung in der Zufriedenheit der Eltern/Pflegepersonen gemessen mit dem TSQM ○ Änderung in der Häufigkeit von Stürzen pro Tag (Patiententagebuch) • kognitive Funktion <ul style="list-style-type: none"> ○ verbales Gedächtnis, Aufmerksamkeit gemessen mittels ‚Digit Span Task‘ • kardiale Funktion <ul style="list-style-type: none"> ○ Änderung der Herzfrequenz vor, während und nach dem 6MWT mittels Herzfrequenzmonitoring • Pharmakodynamik <ul style="list-style-type: none"> ○ Änderung der Serum-CK-Konzentration ○ Änderung der Dystrophin-Expression im Muskel (M. biceps) mittels Immunfluoreszenz • 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 schwerwiegend) gemäß CTCAE Version 3.0 ○ Entnahme von Blutproben für hämatologische und biochemische Analysen

Charakteristikum	Beschreibung
	<ul style="list-style-type: none"> ○ Vitalzeichen ○ Ultraschalluntersuchung der Nieren und ableitenden Harnwege: während Screening und nach 48 Wochen ○ körperliche Untersuchung
Subgruppenanalysen	Subgruppenanalysen des primären Endpunkts wurden für die folgenden Variablen geplant (vgl. CSR, S. 121): <ul style="list-style-type: none"> • Alter (< 9 / ≥ 9 Jahre) • 6MWT (≥ 350 / < 350 m) • Kortikosteroide zu Baseline (ja / nein) • Altersverteilung (5-6, 7-8, 9-11, ≥ 12 Jahre) • Lokalisation der Nonsense-Mutation (1-39, 40-79) • Art des Stopp-Codons (UGA, UAG, UAA)

Entsprechend dem X-chromosomalen Erbgang und dem natürlichen Verlauf der Erkrankung wurden ausschließlich männliche Patienten im Alter von 5-20 Jahren in die Studie eingeschlossen. Darunter befanden sich sechs Geschwisterpaare. Vier dieser Paare wurden in die Placebo-Gruppe, jeweils ein Paar in die beiden Ataluren-Studienarme randomisiert. Das Vorhandensein einer Nonsense-Mutation wurde bei allen randomisierten Patienten mittels Sequenzierung in einem zertifizierten Labor bestätigt.

Tabelle 2: Charakteristika der Studienpopulation in der Studie 007

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
<i>Alter (Jahren)</i>			
Mittelwert (SD)	8,3 (2,33)	8,8 (2,91)	8,4 (2,53)
Median	8,0	8,0	8,0
Spanne	5-15	5-20	5-16
<i>Abstammung, N (%)</i>			
kaukasisch	54 (94,7)	53 (93)	50 (83,3)
schwarz	0	1 (1,8)	1 (1,7)
asiatisch	1 (1,8)	1 (1,8)	4 (6,7)
hispanisch	1 (1,8)	1 (1,8)	2 (3,3)
andere	1 (1,8)	1 (1,8)	3 (5,0)
<i>Körpergröße (cm)</i>			
Mittelwert (SD)	123,4 (11,8)	124,5 (15,3)	126,2 (13,8)
Median	122,1	121,1	125,9
Spanne	104-163	99-173	99-173
<i>Körpergewicht (kg)</i>			
Mittelwert (SD)	28,6 (9,1)	31,2 (12,1)	31,9 (12,8)
Median	25,6	27,0	27,6
Spanne	16-55	16-76	17-84

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
<i>Body-Mass-Index (kg/m²)</i>			
Mittelwert (SD)	18 (3,7)	19 (3,5)	19 (4,8)
Median	17,4	18,8	18,2
Spanne	13-29	14-31	14-41
<i>Alter bei Diagnose (Jahre)</i>			
Mittelwert (SD)	3,9 (2,3)	3,3 (1,8)	3,8 (2,0)
Median	4,0	3,0	3,5
Spanne	0-10	0-9	0-8
<i>Zeit von Diagnosestellung bis Randomisierung (Jahre)</i>			
Mittelwert (SD)	4,4 (2,5)	5,4 (3,4)	4,6 (3,1)
Median	5,0	5,0	4,0
Spanne	0-11	0-17	0-14
<i>Phänotypische Diagnose, N (%)</i>			
proximale Muskelschwäche	50 (88)	49 (86)	52 (87)
Watschelgang	44 (77)	43 (75)	49 (82)
Gower-Zeichen	47 (83)	50 (88)	51 (85)
Wadenhypertrophie	56 (98)	48 (84)	55 (92)
Andere ¹⁾	12 (21)	19 (33)	13 (22)
<i>6MWT²⁾ zu Baseline, m</i>			
Mittelwert (SD)	359,6 (87,7)	350,0 (97,6)	358,2 (104,0)
Median	354,0	362,1	368,0
Spanne	159-533	75-525	90-554

Quelle: CSR, S. 81

¹⁾ Inklusive Zehen(spitzen)gang, Muskelkrämpfe, Lordose und Entwicklungsverzögerung.

²⁾ 2 Patienten hatten sich vor der Baseline-Erhebung am Bein verletzt (je 1 Patient in der Placebo- und der Ataluren-20/20/40mg-Gruppe), für die post-hoc-Analyse der korrigierten ITT-Population wurden Werte aus der Screening-Erhebung eingesetzt, die sich wie folgt auf die Mittelwerte auswirkten: Placebo [Mittelwert (SD)]: 361,1 (87,5); Ataluren 10/10/20 mg/kg: 350,0 (97,6) [= unverändert]; Ataluren 20/20/40 mg/kg: 361,2 (99,7).

Tabelle 3: Charakteristika der Studienpopulation nach Stratifizierungsfaktoren in der Studie 007

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
Alter (Jahre), N (%)			
< 9	32 (56)	32 (56)	34 (57)
≥ 9	25 (44)	25 (44)	26 (43)
Einnahme von Kortikosteroiden, N (%)			
ja	40 (70)	41 (72)	43 (72)
nein	17 (30)	16 (28)	17 (28)
6MWT (Meter), N (%)			
≥ 350	34 (60)	32 (56)	33 (55)
< 350 ¹⁾	23 ²⁾ (40)	25 (44)	27 (45)

Quelle: CSR, S. 78

¹⁾ Die Gehstreckenstratifizierung wurde im 3. Protokollamendment von 270 m auf 350 m geändert, nachdem bereits 42 Patienten randomisiert waren. Von diesen 42 Patienten wurden 13 von dem ursprünglichen Stratum ≥ 270 m in das Stratum < 350 m neu zugeteilt: Placebo-Gruppe N=4, Ataluren 10/10/20 N=4, Ataluren 20/20/40 N=5. Alle Auswertungen erfolgten nach Restratifizierung.

²⁾ 1 Patient, der in die Placebo-Gruppe randomisiert war, verletzte sich 1 Tag vor der Baseline-Visite, das Ergebnis des 6MWT (309 m) wurde fälschlicherweise als valide eingetragen, post hoc aber durch den Screening-Wert (395 m) ersetzt. In der post-hoc-Analyse wurde der Patient in die ≥ 350 m-Gruppe restratifiziert.

Die Studienarme waren insgesamt vergleichbar. Das Alter bei Diagnosestellung war in der hier relevanten Ataluren-10/10/20mg-Gruppe geringfügig niedriger, entsprechend war das Zeitintervall von Diagnosestellung bis Randomisierung in dieser Gruppe geringfügig größer im Vergleich zur Placebo-Gruppe. Dementsprechend waren auch Durchschnittsalter, Gewicht und Größe in dieser Gruppe etwas höher als im Placebo-Arm.

Post hoc wurde bei zwei Patienten (je ein Patient in der Placebo- bzw. in der Ataluren-20/20/40mg-Gruppe) der Baseline-Wert für den 6MWT durch den Screening-Wert ersetzt. Grund hierfür waren akute Verletzungen (Knie bzw. Sprunggelenk) wenige Tage vor der Erhebung der Baseline-Werte der unteren Extremität, die zu einem falsch-niedrigen, d. h. nicht repräsentativen Ergebnis zu Baseline führten (309 statt 395 m bzw. 125 statt 303 m bei der Screening-Visite). Die Werte wurden zunächst als gültig eingestuft und nachträglich durch die Screening-Werte ersetzt, was zur korrigierten ITT-Population (kITT) führte. Ergänzend wurden die Werte im Rahmen einer Sensitivitätsanalyse aus der Analyse eliminiert, die Ergebnisse werden ebenfalls dargestellt (vgl. CSR, S. 97f).

Tabelle 4: Charakterisierung der Intervention in der Studie 007

Intervention	Kontrolle
<u>Behandlungsarm 1:</u> 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, 6 und 12 Std.; Einnahme innerhalb von 30 Minuten nach einer Mahlzeit <u>Behandlungsarm 2:</u> Ataluren 80 mg/kg KG (aufgeteilt auf drei Dosen mit 20, 20 und 40 mg/kg KG) täglich: morgens, mittags, abends im Abstand von 6, 6	Einnahmeschema sowie Aussehen/Verpackung, Geruch und Geschmack identisch zur Prüfintervention und somit nicht von Prüfmedikation zu unterscheiden.

Dossierbewertung für Orphan Drugs

Ataluren

Intervention	Kontrolle
<p>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 sind 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 Dosis ausgelassen.</p>	
<p>Vorbehandlung, erlaubte und nicht erlaubte (Begleit-)Medikationen</p> <p>Wurden bereits bei Studieneintritt Kortikoidsteroiden eingenommen, dann sollte diese möglichst standardisiert erfolgen, 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 ACE-Hemmer und β-Blocker sollten möglichst standardisiert weiterverordnet oder 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 sollten vermieden werden, um Confounding zu vermeiden, da diese Arzneimittel das „Durchlesen“ von Ribosomen bei Nonsense-Mutationen ebenfalls induzieren können. Bei schwerwiegenden Infektionen sollten möglichst andere Antibiotika verwendet werden.</p> <p>Warfarin war nicht erlaubt, da dieses ebenfalls über CYP2C9 metabolisiert wird.</p> <p>Phenytoin war nur unter Monitoring des Plasmaspiegels erlaubt.</p>	

2.5 Endpunkte und Erhebungsinstrumente

2.5.1 In der Nutzenbewertung berücksichtigte Endpunkte

In der Nutzenbewertung wurden folgende auch vom pharmazeutischen Unternehmer (pU) als patientenrelevant eingestuft 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 <i>supportiv:</i> Anteil der Patienten mit einer mindestens 10%igen Verbesserung nach 48 Wochen (Responder) bzw. einer mind. 10%igen Verschlechterung nach 48 Wochen (Progressoren), Zeit bis zur 10%igen Verbesserung/Verschlechterung
Sekundäre Endpunkte Änderung in der proximalen Muskelfunktion in ‚Timed Function Tests‘, d. h. so schnell wie möglich <ul style="list-style-type: none"> • Aufstehen aus Rückenlage • 10 m laufen/gehen • 4 Treppenstufen hinauf-/hinuntersteigen
Änderung der Alltagsaktivitäten gemessen mittels ‚Step Activity Monitoring‘ (Pedometer am Knöchel) an ≥ 9 Tagen nach jeder Studiervisite
Änderung in der Häufigkeit von Rollstuhlnutzung (SAM-Patiententagebuch)
Änderung in der Häufigkeit von Stürzen pro Tag (Patiententagebuch)
Lebensqualität
Änderung in der Lebensqualität gemessen mit dem PedsQL (physikalische, emotionale, soziale, schulische Subskalen) sowie PedsQL Fatigue Scale
Sicherheit
Art und Häufigkeit unerwünschter Ereignisse und schwerwiegender unerwünschter Ereignisse

Die Endpunkte Aufstehen aus Rückenlage, 10 m laufen/gehen, Änderung der Alltagsaktivitäten, Änderung in der Häufigkeit von Stürzen pro Tag, verbales Gedächtnis/Aufmerksamkeit sowie Lebensqualität gemäß PedsQL wurden vom pU in Modul 4 des Dossiers als nicht patientenrelevant betrachtet und die Ergebnisse wurden daher nicht als Grundlage für die Herleitung des Zusatznutzens herangezogen (vgl. Modul 4, Tab. 4-29, S. 76). Dennoch werden die Ergebnisse für Aufstehen aus Rückenlage, 10 m laufen/gehen, Sturzhäufigkeit, Lebensqualität und Anzahl der Tage, an denen ein Rollstuhl benutzt wurde, in der Nutzenbewertung ergänzend dargestellt (vgl. Modul 4, S. 112).

2.5.2 Verzerrungspotential auf Studien- und Endpunktebene

Bei der Studie PTC124-GD-007-DMD handelt es sich um eine randomisiert-kontrollierte Studie der Phase IIb. Die Zuteilung auf die Behandlungsgruppen erfolgte mithilfe von Interactive Voice Response bzw. Interactive Web Response (IVR/IWR). Patienten, Untersucher und Forschungspersonal 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 007 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 (nicht korrigierten) ITT-Population.

Tabelle 6: Verzerrungspotential der Endpunkte in der Studie 007

Endpunkt	Verblindung adäquat	ITT adäquat	Keine ergebnis- gesteuerte Berichte Erstattung	Keine sonstigen Aspekte	Verzerrungs- potential Endpunkt
Morbidität					
Primärer Endpunkt: Veränderung im 6MWT	ja	ja	ja	ja	gering
<i>proximale Muskelfunktionstests (TFT)</i>					
Aufstehen aus Rückenlage	ja	ja	ja	ja	gering
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
<i>patientenberichtete Morbiditätsendpunkte</i>					
Alltagsaktivitäten	ja	ja	ja	ja	gering
Stürze	ja	ja	ja	ja	gering
Rollstuhlnutzung	ja	ja	ja	nein ¹⁾	unklar
Lebensqualität					
Veränderung der Lebens- qualität (Peds-QL)	ja	ja	ja	ja	gering
PesQL Fatigue Scale	ja	ja	ja	ja	gering
Sicherheit					
UE	ja	ja	ja	ja	gering
SUE	ja	ja	ja	ja	gering

¹⁾ Im CSR (S. 113) wird unter der Rubrik Alltagsaktivitäten auch der prozentuale Anteil der Tage mit Rollstuhlbenutzung als patientenberichteter Endpunkt dargestellt. Es ist davon auszugehen, dass die Angaben zur Rollstuhlnutzung an jeweils 7 Tagen nach einer Studiervisite erhoben wurden. Ob dieser vergleichsweise kurze Zeitraum für die Rollstuhlnutzung repräsentativ ist, ist unklar.

2.5.3 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><u>Operationalisierung:</u> 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): Es wird in 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, glattem 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 Gehrichtung ändert, - den Test vor Ablauf von 6 Minuten abbricht, aufgrund von Noncompliance oder

Morbidität	
	<p>aus anderen Gründen als Erschöpfung (Verletzung nach Sturz, Hinsetzen).</p> <p><i>Erhebungszeitpunkte:</i> Der 6MWT wurde beim Screening, zu Baseline (vor Beginn der Studienmedikation) und nach jeweils 6 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 Verbesserung nach 48 Wochen (Responder) bzw. einer mind. 10%igen Verschlechterung nach 48 Wochen (Progressoren) kalkuliert.</p> <p><u>Begründung des pU gemäß der im Dossier vorhandenen Angaben:</u></p> <p>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 direkte Einschätzung der Gehfunktion bzw. als sensitives Maß für die Krankheitsprogression von DMD-Patienten angesehen. In mehreren Studien wurde gezeigt, dass DMD-Patienten eine durchschnittlich geringere Gehstrecke im 6MWT aufweisen als gesunde Gleichaltrige (bspw. in einer Studie 268 ± 92 m vs. 468 ± 42 m; unpublizierte Daten von 52 Kindern, zitiert nach Studienprotokoll Version 3.0, S. 87).</p> <p>Der 6MWT wurde als Endpunkt in Zulassungsstudien für Arzneimittel anderer neuromuskulärer und pulmonaler Erkrankungen eingesetzt, bspw. Mukopolysaccharidose I, II und IV, pulmonale Hypertonie.</p> <p>Des Weiteren wird darauf verwiesen, dass die National Institutes of Health den 6MWT als „globales Instrument der Gehfähigkeit und Ausdauer“ einstufen (S. 78). Der Erhalt der Muskelfunktion wurde in einer Umfrage für DMD-Patienten als wichtiger eingeschätzt als die Verlängerung der Lebenserwartung (S. 78).</p> <p>Vorerfahrungen bzgl. einer MCID für den 6MWT bei DMD-Patienten existieren bisher nicht. Erfahrungen bei Patienten mit pulmonaler Hypertonie und Mukopolysaccharidose legen nahe, dass eine Verbesserung um 10-15 % als klinisch relevant angesehen wird. Bezogen auf eine durchschnittliche Gehstrecke von ca. 270 m ergibt sich damit eine anzustrebende absolute Verbesserung um 30 m im Vergleich zu Placebo.</p> <p>In Modul 4 des Dossiers verweist der pU zudem auf eine italienische multizentrische Kohortenstudie mit einer Laufzeit von 24 Monaten, mit dem Ziel, das Spektrum von krankheitsbedingten Veränderungen bei DMD-Patienten mit mind. 75 m Gehstrecke mit verschiedenen Erhebungsinstrumenten zu erfassen. Insgesamt wurden 113 Kinder (Durchschnittsalter 8,2 Jahre, Spanne 4,1-17) eingeschlossen. Der Gehstreckenverlust betrug im Durchschnitt 87,3 m über 2 Jahre, bei beträchtlicher Variabilität (SD 123,1 m), besonders ausgeprägt bei Kindern über 7 Jahren (vs. < 7). Demnach liegt das Risiko, innerhalb von 2 Jahren die Gehfähigkeit zu verlieren, bei Kindern mit einer Gehstrecke im 6MWT von < 330 m bei 30 % (Mazzone et al. 2013).</p> <p>Hinsichtlich der MCID wird auf Untersuchungen von McDonald et al. verwiesen, die zeigen, dass eine Gehstreckenverkürzung um 30 m das Risiko, die Gehfähigkeit innerhalb von 2 Jahren zu verlieren, erheblich erhöht.</p> <p>Außerdem verweist der pU auf die Studie von Henricson et al. (2013) (s. u.) und Pane et al. (2014), um die prognostische Relevanz der Gehstrecke über/unter 350 m im 6MWT aufzuzeigen. Auch in der vorliegenden Studie 007 zeigt sich die Relevanz dieser Unterteilung in der Placebo-Gruppe. Es kann unterschieden werden in Patienten, die sich in einer stabilen Phase und solche, die sich in einer „Decline“-Phase befinden (vgl. Abb. 7 und 8, S. 82). Daraus folgert der pU, dass der Effekt von Ataluren in der Studie 007 vermutlich unterschätzt wird und eine Stratifizierung gerechtfertigt erscheint.</p>

Morbidität	
Einschätzung zur Qualität u. Patientenrelevanz	<p><u>Definition und Erhebung des Endpunkts:</u> 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><u>Validität und Patientenrelevanz:</u> Die EMA-Draft-Guideline zu DMD/BMD (Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy, EMA/CHMP/236981/2011) von 2013 empfiehlt die Erhebung des funktionellen Status ("functional mobility") sowie ergänzend Aktivitäten des täglichen Lebens. Ein ko-primärer Endpunkt sollte definiert werden, der sowohl muskuläre Funktion wie auch Muskelkraft erfasst. 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.</p> <p>Der pU verweist im Modul zudem auf eine Studie von Henricson et al. (2013), in der zwei Lebensqualitätsinstrumente – das Pediatric Quality of Life Inventory (PedsQL) und das Pediatric Outcomes Data Collection Instrument (PODCI) – genutzt werden, um eine individuelle MCID für den 6MWT für Patienten mit DMD herzuleiten, die sich auch in eine Verbesserung der Lebensqualität übersetzen lässt. Das PedsQL ist ein 23-Item umfassendes Instrument, das von betroffenen Kindern oder deren Eltern ausgefüllt werden kann (s. u. Lebensqualität). Das PODCI-Instrument kann ebenfalls von betroffenen Kindern oder deren Eltern ausgefüllt werden. Es misst die Funktionsfähigkeit bei Kindern mit orthopädischen Einschränkungen. Skalen des Instruments umfassen Funktion der oberen Extremitäten, Mobilität, Sport und körperliche Funktion, Komfort und Schmerz, Glück und Zufriedenheit, und die behandlungsbezogenen Erwartungen. In die Studie wurden 24 Jungen mit DMD und 36 gesunde Jungen zwischen 4 und 12 Jahren eingeschlossen. Alle Patienten mussten zu Studienbeginn eine Gehstrecke von mindestens 10 m zurücklegen können. Die Fragebögen PedsQL und PODCI wurden ausgegeben und von den Eltern ausgefüllt. Die Werte der 6MWD sowie der Lebensqualitätsbögen wurden zum Baseline und nach einem Jahr erhoben. Aus den Baseline-Werten wurden für die DMD-Gruppe verteilungsbasierte MCID für den 6MWT sowie für die PedsQL- und PODCI-Skalen berechnet. Dafür wurde 1/3 der Standardabweichung zugrunde gelegt. Der 6MWT-Baseline-Wert der Kinder mit DMD war deutlich korreliert mit den PODCI-Baseline-Scores (Gesamtscore $r=0,68$; Mobilitätsskala $r=0,79$; Sport und körperliche Funktionsskala $r=0,68$). Für den PedsQL konnte für Patienten mit DMD keine Korrelation mit dem 6MWT nachgewiesen werden (Gesamtscore $r=0,21$; körperliche Funktionsskala $r=0,18$). 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 Veränderung war jeweils größer als die aus den Baseline-Werten berechnete MCID. Für die gesunden Kinder zeigten sich keine statistisch signifikanten Veränderungen. 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 körperlichen Funktionsskala zeigte sich eine mittlere Korrelation ($r=0,45$). Mit dem PedsQL-Gesamtscore zeigte sich keine Korrelation. Es wurde eine leicht negative Korrelation zwischen der PedsQL körperlichen Funktionsskala und der Veränderung des 6MWT beobachtet ($r=-0,36$). Es zeigt sich, dass bei niedrigen Ausgangswerten auf der PODCI-Mobilitätsskala geringere Veränderungen des 6MWT notwendig sind, um eine klinisch relevante Verbesserung auf der PODCI-Mobilitätsskala von 4,5 Punkten zu</p>

Morbidität	<p>erreichen, wohingegen bei höheren Ausgangswerten auf der PODCI-Mobilitätsskala größere Veränderungen des 6MWT notwendig sind. Bei niedrigen PODCI-Werten ergibt sich demnach eine notwendige Verbesserung des 6MWT von 3,7 %. Die prozentual notwendige Veränderung des 6MWT steigt bis 10 % an, wenn initial hohe Werte auf der PODCI-Mobilitätsskala erreicht wurden. Die in der Studie 007 vorliegenden Mittelwerte für den 6MWT zu Baseline (350-360 m) würden daher eine Änderung von ca. 36 m erfordern, um eine klinisch relevante Änderung im PODCI-Mobilitätsscore von 4,5 Punkten zu erreichen. Allerdings genügt für jeden der untersuchten PODCI-Scorebereiche eine Änderung um max. 10 %, um eine klinisch relevante Änderung zu erreichen. Es wird nicht dargelegt, warum bei der verteilungsbasierten Berechnung der MCID nicht der gebräuchliche Wert von 1/2 der Standardabweichung angewendet, sondern eine MCID von 1/3 der Standardabweichung gewählt wurde. Zudem liegen nur von 54 % der DMD-Patienten (13 von 24) und von 50 % der Kontrollen (18 von 36) in der Studie Daten für das 1-Jahres-Follow-up vor. Eine Imputation fehlender Daten wurde nicht vorgenommen. Die Aussagekraft der Studie zu Veränderungen der 6MWT und der Lebensqualitätsmaße ist daher eingeschränkt.</p> <p>Die Patientenrelevanz des 6MWT als Surrogat für „Meilensteine“ des Funktionsverlusts ist entsprechend der Expertenempfehlungen nachvollziehbar (Bushby 2011), allerdings ist die Begründung für die gewählte Relevanzschwelle von 30 m aufgrund der methodischen Einschränkungen der Assoziationsstudie mit dem PODCI-Instrument derzeit nur eingeschränkt als valide zu betrachten, insbesondere weil der Zusammenhang bisher lediglich indirekt untersucht wurde. Zudem wurde das PODCI-Instrument in der Studie 007 nicht eingesetzt.</p>
Endpunkt	Änderung in der proximalen Muskelfunktion (,Timed Function Tests‘) zu Woche 48
Operationalisierung und Validierung: Begründung seitens des pU	<p><u>Operationalisierung</u></p> <p>Die Änderung in der proximalen Muskelfunktion in ,Timed Function Tests‘ (TFT), wurde in vier Testsituationen aufgeteilt, d. h. die Patienten sollten so schnell wie möglich</p> <ul style="list-style-type: none"> - Aufstehen aus Rückenlage, - 10 m laufen/gehen, - Treppenstufen hinauf-/hinuntersteigen. <p>Das Zeitlimit beträgt jeweils 30 Sekunden. Wenn die Aufgabe nicht innerhalb dieser Zeit bewältigt wurde, dann werden 30 Sekunden notiert.</p> <p>Für die Tests wird zum einen die Zeit bis zur Komplettierung erfasst. Außerdem wird für jeden Test ein Score gebildet (Studienprotokoll Version 3.0, S. 59f):</p> <p>Aufstehen aus Rückenlage: Einteilung in die folgenden 6 Kategorien: 1 = unfähig, aus Rückenlage aufzustehen; 2 = Aufstehen mit abstützen an Möbeln (assisted Gowers); 3 = auf Bauch drehen, Aufrichten mittels abstützen an den Beinen (full Gowers); 4 = auf Bauch drehen, Aufrichten mittels Abstützen mit einer Hand an den Beinen (half Gowers); 5 = auf die Seite drehen und/oder aufstehen nach abstützen auf dem Boden, aber nicht an den Beinen; 6 = Aufstehen ohne drehen oder abstützen. Höhere Werte repräsentieren ein besseres Ergebnis.</p> <p>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</p>

Morbidität	
	<p>Werte repräsentieren ein besseres Ergebnis.</p> <p>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 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> <p>4 Treppenstufen hinuntersteigen: Der Score wurde analog zum Treppenhochsteigen gebildet: 1 = unfähig (...) 6 = ohne Hilfe möglich. Höhere Werte repräsentieren ein besseres Ergebnis.</p> <p>Erhebungszeitpunkte: zum Screening, an Tag 1 (vor Beginn der Studienmedikation), nach 6, 12, 18, 24, 30, 36, 42 und 48 Wochen.</p> <p><u>Begründung des pU gemäß der im Dossier vorhandenen Angaben:</u> (siehe Modul 4, S. 83 und Studienprotokoll Version 3.0, S. 89)</p> <p>Die TFTs messen die proximale Muskelfunktion, die substantiell schlechter ist im Vergleich zu gesunden Kindern. Diese Funktionstests werden als zusätzliche Endpunkte zum 6MWT eingesetzt.</p> <p>Der pU beruft sich auch auf Empfehlungen einer DMD-Expertengruppe. Demnach kann der Erhalt von Fähigkeiten wie Treppensteigen, selbständiges Gehen, Essen usw. als relevant für die Krankheitsprogression angesehen werden. Die in der Studie eingesetzten Tests werden in Studien bei DMD-Patienten üblicherweise verwendet. Sie sind reproduzierbar und können den Effekt von Steroiden auf die DMD erfassen.</p> <p>Der 10-Meter-Test ist breit akzeptiert und wird häufig eingesetzt, um die physische Spitzenaktivität zu messen. Bei einer Zeit von mehr als 12 Sekunden (in einer anderen Studie wurden 7 Sekunden ermittelt) ist das Risiko deutlich erhöht, innerhalb von 12 Monaten die Gehfähigkeit zu verlieren. Patienten, die weniger als 6 Sekunden benötigen, haben ein signifikant niedrigeres Risiko, die Gehfähigkeit zu verlieren.</p> <p>Weiterhin wird eine Beobachtungsstudie zitiert, die Assoziationen zwischen den Ergebnissen der TFTs und der physischen Funktion in zwei Instrumenten zur gesundheitsbezogenen Lebensqualität (PODCI, PedsQL) untersucht (zitiert nach McDonald 2013).</p> <p>Die Fähigkeit Treppen zu gehen hat sich ebenfalls als prädiktiv für den Verlust der Gehfähigkeit erwiesen; hier wurde eine Grenze von 8 Sekunden ermittelt, 40 % der Patienten, die länger als 8 Sekunden benötigten bzw. nicht in der Lage waren, Treppen zu steigen, verloren in den nächsten 12 Monaten ihre Gehfähigkeit (zitiert nach McDonald 2013).</p> <p>In der schon erwähnten Studie von Henricson et al. (2013) wurde eine MCID von 0,19 Meter/Sekunde für den 10-Meter-Gehtest ermittelt.</p>

Morbidität	
Einschätzung zur Qualität u. Patientenrelevanz	<p><u>Definition und Erhebung des Endpunkts:</u> Die Operationalisierung des Endpunkts bzw. der Einzeltests ist nachvollziehbar.</p> <p><u>Validität und Patientenrelevanz:</u> 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 et al. 2011). Dieser Expertenkonsens beruht auf mehreren Treffen europäischer und US-amerikanischer Experten in den Jahren 2009 und 2010, 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 insgesamt 8 klinischen und Longitudinalstudien verschiedener Arbeitsgruppen (Dänemark, Deutschland, England, Italien, USA), die fast 2.000 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> <p>Bushby et al. weisen auch darauf hin, dass die Interpretation der Zeitangaben in heterogenen Populationen problematisch werden kann, wenn der Anteil von Nicht-Vollendern innerhalb des vorgegebenen Zeitfensters zu hoch wird, weshalb in diesen Fällen die Konvertierung der Zeitangaben in Geschwindigkeiten (Meter/Sekunde) vorgeschlagen wird. Die Tests werden am ehesten als sensitiv für Änderungen über 1 Jahr bei Kindern von 7-9 bzw. 10-12 Jahren.</p> <p>Der Begründung des pU für die Patientenrelevanz der TFTs kann insgesamt gefolgt werden. Die Validität der MCID kann nicht abschließend eingeschätzt werden.</p>
Endpunkt	Änderung der Alltagsaktivitäten gemessen mittels ‚Step Activity Monitoring‘ (Pedometer am Knöchel) an ≥ 9 Tagen nach jeder Studienvisite
Operationalisierung und Validierung: Begründung seitens des pU	<p><u>Operationalisierung:</u> Die Änderung der Alltagsaktivitäten wurde mittels ‚Step Activity Monitoring‘ anhand der folgenden Parameter gemessen (vgl. Studienprotokoll Version 3.0, S. 77, 88; SAP, S. 18):</p> <ul style="list-style-type: none"> - mittlere Aktivitätsperiode/Tag/Visite - mittlere Anzahl Schritte je Tag/Visite während der aktiven Perioden - mittlere Anzahl Schritte pro Stunde während der aktiven Perioden an den Tagen der Studienvisite - maximale Anzahl an Schritten während kontinuierlicher Aktivitätsphasen (je 10, 20, 30, 60 Minuten) - Zeitanteil während der aktiven Perioden, der mit geringer Aktivität (≤ 15 Schritte pro Minute) verbracht wurde - Zeitanteil während der aktiven Perioden, der mit mittlerer Aktivität (16-30 Schritte pro Minute) verbracht wurde - Zeitanteil während der aktiven Perioden, der mit hoher Aktivität (> 30 Schritte pro Minute) verbracht wurde <p>Der Pedometer wird bei Studienvisite 1 erstmals kalibriert (Eingabe von Größe und Gangcharakteristika) und bei jeder weiteren Visite wird die Kalibrierung überprüft. Es handelt sich um ein Accelerometer, das mit hoher Genauigkeit (in einer Studie wurde eine Genauigkeit von $> 99,9\%$ im Vergleich zu manueller Erfassung) Schritte erfassen kann. Das Gerät ist klein und leicht zu tragen und interferiert nicht mit Alltagsaktivitäten. Die Daten werden nicht direkt an den Träger zurückgemeldet, so dass eine Verhaltensänderung durch das Tragen des Geräts minimiert wird. Die</p>

Morbidität	<p>Daten werden zusammen mit dem SAM-Patiententagebuch (SAM Diary) interpretiert, um potentielle Einflussfaktoren (bspw. im Zusammenhang mit Wetter, Wochentag oder Compliance) berücksichtigen zu können.</p> <p>Der erste Tag nach der Studienvisite wurde jeweils nicht ausgewertet, um die Ergebnisse nicht durch die visite-bedingte Unterbrechung zu verfälschen. Daten für die Tage wurden gelöscht, an denen die Anzahl der Schritte < 10% bzw. die aktive Periode < 50% des Durchschnittswertes der übrigen Tage lag. Aktivität wurde zwischen 3.00 Uhr und 24 Uhr registriert, wenn mindestens 2 Schritte pro Minute registriert wurden. Es wurden jeweils mindestens 4, maximal 7 Tage ausgewertet. Daten zu weniger als 4 Tagen je Visite wurden als nicht evaluierbar gewertet.</p> <p>Erhebungszeitpunkte: zum Screening, an Tag 1 (vor Beginn der Studienmedikation), nach 6, 12, 18, 24, 30, 36, 42 und 48 Wochen.</p> <p><u>Begründung des pU gemäß der im Dossier vorhandenen Angaben:</u></p> <p>Der pU begründet die Messung der Alltagsaktivität mittels Step Activity Monitoring (SAM) als unterstützend für den primären Endpunkt. Die Quantifizierung mittels Pedometer wird als notwendig erachtet, da patientenberichtete Angaben alleine als ungenau angesehen werden.</p> <p>In Modul 4 wird der Endpunkt jedoch nicht als patientenrelevant eingestuft und daher nicht zum Nachweis des Zusatznutzens aufgeführt.</p>
Einschätzung zur Qualität u. Patientenrelevanz	<p><u>Definition und Erhebung des Endpunkts:</u></p> <p>Die Operationalisierung des Endpunkts ist nachvollziehbar. Die EMA empfiehlt die Erfassung der Aktivitäten des täglichen Lebens in ihrer Draft Guideline von 2013.</p> <p><u>Validität und Patientenrelevanz:</u></p> <p>Daten zur Validität (mit Ausnahme der Angabe der Genauigkeit der Übereinstimmung der SAM-Messung mit einer manuellen Schrittzählung) und Patientenrelevanz werden nicht vorgelegt, daher kann keine abschließende Einschätzung erfolgen.</p> <p>Entsprechend § 2 Satz 3 der Arzneimittelnutzenbewertungsverordnung (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.“</p>
Endpunkt	Patientenberichtete Änderung der Häufigkeit von Stürzen und der Rollstuhlnutzung (Patiententagebuch)
Operationalisierung und Validierung: Begründung seitens des pU	<p><u>Operationalisierung:</u></p> <p>Ein Patiententagebuch soll täglich von den Eltern/Pflegepersonen geführt werden. Darin werden Einnahme der Studienmedikation, Kortikosteroide sowie Stürze dokumentiert. Das Patiententagebuch wird bei jeder Studienvisite abgegeben.</p> <p>Die Sturzhäufigkeit wurde als Anzahl der Stürze pro Patient bezogen auf die Zeit unter Therapie kalkuliert (SAP, S. 32).</p> <p>Erhebungszeitpunkte: zum Screening, an Tag 1 (vor Beginn der Studienmedikation), nach 6, 12, 18, 24, 30, 36, 42 und 48 Wochen.</p> <p><u>Begründung des pU gemäß der im Dossier vorhandenen Angaben:</u></p> <p>Unbeabsichtigte Stürze stellen wichtige Ereignisse dar, da Verletzungsgefahr besteht. Frühere Studien haben gezeigt, dass Stürze mittels Patiententagebüchern zuverlässig erfasst werden können. Studien zum Effekt von Glukokortikoiden bei DMD-Patienten haben zudem gezeigt, dass die Sturzfrequenz mittels medikamentöser Therapie reduziert werden kann. Die Reduktion von Stürzen und damit asso-</p>

Morbidität	
	zierten Verletzungen (Frakturen) und Schmerzen wird daher als patientenrelevanter Nutzen angesehen, insbesondere bei Patienten mit glukokortikoidinduzierter Osteoporose (Studienprotokoll Version 3.0, S. 94).
Einschätzung zur Qualität u. Patientenrelevanz	<p><u>Definition und Erhebung des Endpunkts:</u> Zur Rollstuhlnutzung finden sich keine expliziten Angaben zur Operationalisierung. Im CSR (S. 113) wird unter der Rubrik Alltagsaktivitäten auch die Anzahl der Tage mit Rollstuhlbenutzung als patientenberichteter Endpunkt aus dem SAM Diary dargestellt. (Das SAM Diary war ein Tagebuch für Ereignisse an den Tagen, an denen das Pedometer getragen wurde.) Die zugehörigen Tabellen (14.2.6.1 und 14.2.6.2) verweisen ebenfalls auf das SAM, so dass davon ausgegangen wird, dass die Rollstuhlnutzung nicht aus den Angaben des täglichen Patiententagebuchs, sondern aus dem SAM Diary ermittelt wurde. Demnach wurde der Anteil an Tagen mit Rollstuhlnutzung in Prozent kalkuliert.</p> <p><u>Validität und Patientenrelevanz:</u> Daten zur Validität (mit Ausnahme der Angabe der Genauigkeit der Übereinstimmung der SAM-Messung mit einer manuellen Schrittzählung) der Erfassung von Stürzen oder Rollstuhlnutzung werden nicht vorgelegt, daher kann keine abschließende Einschätzung erfolgen. Insbesondere ist unklar, ob die Angaben zur Rollstuhlnutzung, die an jeweils 7 Tagen nach einer Studienvisite erhoben wurden, für die gesamte Studiendauer als repräsentativ anzusehen sind. Es kann aber davon ausgegangen werden, dass Stürze und Rollstuhlnutzung für Patienten eine unmittelbare Relevanz besitzen.</p>

Tabelle 8: Beschreibung der Endpunkte der Kategorie Lebensqualität

Lebensqualität	
Endpunkt	Änderung in der Lebensqualität gemessen mit dem PedsQL (physikalische, emotionale, soziale, schulische Subskalen), Fatigue-Skala
Operationalisierung und Validierung: Begründung seitens des pU	<p><u>Operationalisierung:</u> Für die Erfassung der Lebensqualität wurde das Generic Core Scales-Paket des PedsQL Inventory verwendet. Das Paket beinhaltet die 4 Domänen physikalische, emotionale, soziale und schulische Funktion. Das Instrument liegt in mehreren Versionen vor, Kleinkinder/Kinder 2-4, 5-7 und 8-12 Jahren, Jugendliche (13-18) und junge Erwachsene (18-25). Mit Ausnahme der Version für Erwachsene sind für alle Fragebogenversionen auch Elternfragebögen verfügbar, die verwendet werden sollen, wenn die Patienten selbst nicht in der Lage sind, den Fragebogen auszufüllen. Für die physikalische Domäne werden 8 Items auf 5-Punkte-Likertskalen erfasst, für die übrigen Domänen jeweils 5 Items (insgesamt 23 Fragen). Es kann ein globaler Score (Werte 0-100) gebildet werden sowie Scores für jede Dimension. Je höher der Score, desto besser ist die Lebensqualität. Die Fragebögen sollen von den jeweils selben Personen ausgefüllt werden.</p> <p>Erhebungszeitpunkte: zum Screening, an Tag 1 (vor Beginn der Studienmedikation), nach 6, 12, 18, 24, 30, 36, 42 und 48 Wochen.</p> <p><u>Begründung des pU gemäß der im Dossier vorhandenen Angaben:</u> Bei Patienten mit DMD kann die Einschränkung der körperlichen Funktionen soziale, erzieherische und psychologische Probleme nach sich ziehen. Zusätzliche Faktoren wie Einnahme von Glukokortikoiden, physische und soziale Isolation, chronische Müdigkeit, Schuldgefühle und Stress der Eltern können zu sozialen und Verhaltensauffälligkeiten führen. Anekdotische Berichte aus der Phase-2a-Studie</p>

Lebensqualität	
	<p>legen nahe, dass Patienten psychologisch, sozial und bezüglich Verhalten (z. B. durch geringere Frustration mit körperlichen Aufgaben, verbesserte Möglichkeiten, mit anderen Kindern zu spielen) in Verbindung mit besserer Ausdauer und reduzierter Fatigue durch Ataluren profitieren.</p> <p>Studien bei Patienten, die bereits ihre Gehfähigkeit verloren hatten, zeigten eine Korrelation von funktioneller Beeinträchtigung mit Lebensqualität einerseits, mit TFTs und Myometrie andererseits. Diese Ergebnisse legen nahe, dass Lebensqualität ein responsiver Endpunkt sein kann.</p> <p>Das PedsQL wurde ausgewählt, weil es sich um ein weithin akzeptiertes Instrument zur Erfassung der Lebensqualität handelt, das sowohl von Eltern wie von Kindern ausgefüllt werden kann. Dabei werden die für Patienten mit DMD relevanten Dimensionen (körperliche, soziale, emotionale und schulische Funktion) erfasst. Die Testgütekriterien sind hoch (Test-Retest-Reliabilität 0,88 für selbstberichtete Scores und 0,90 für elternberichtete Scores), das Instrument wurde extern validiert und ist veränderungssensitiv. Zudem ist es einfach anzuwenden, die Zeit zum Ausfüllen beträgt ca. 4 Minuten. Das PedsQL-Instrument liegt in den jeweiligen Sprachen der Länder vor, in denen die Studienzentren angesiedelt sind.</p> <p>Ergänzend wird das PedsQL-Fatigue-Instrument eingesetzt, das zum Zeitpunkt der Protokollerstellung allerdings nur in Englisch vorlag, aber parallel zur Studiendurchführung in die erforderlichen Sprachen übersetzt und jeweils validiert werden sollte. Der PedsQL-Fatigue enthält 18 Fragen.</p>
Einschätzung zur Qualität u. Patientenrelevanz	<p><u>Definition und Erhebung des Endpunkts:</u></p> <p>Die Operationalisierung des Endpunkts ist nachvollziehbar.</p> <p>Das Inventory wurde in den 1990er Jahren als generisches Instrument zur Erfassung der Lebensqualität bei pädiatrischen Populationen mit chronischen Erkrankungen entwickelt. Seither wurden für zahlreiche Erkrankungen spezifische Module entwickelt, zuletzt auch für DMD. Zum Zeitpunkt der Durchführung der Studie 007 gab es das DMD-Modul des PedsQL noch nicht, dieses wurde erst 2012 publiziert.</p> <p>Die interne Konsistenz wurde an einer umfangreichen Stichprobe mit ca. 8.800 gesunden und ca. 850 chronisch kranken Kindern untersucht. Die Test-Retest-Reliabilität bei Kindern (7-12 Jahre) wird mit 0,86 angegeben. Die interne Validität wurde an einer Stichprobe mit ca. 5.500 Kindern (ca. 570 davon chronisch krank) untersucht. Es konnte gezeigt werden, dass Kinder mit chronischen Erkrankungen konsistent eine niedrigere Lebensqualität berichten als gesunde Kinder. Zur Änderungssensitivität liegen weitere Untersuchungen vor. Die MCID wird mit 4,4 (Änderung im Gesamtscore für selbstberichtete Ergebnisse) bzw. mit 4,5 (Elternfragebogen) angegeben (Quellen: Proquolid, http://www.pedsq.org/).</p> <p>Das PedsQL-Fatigue-Instrument wurde zur Messung der Fatigue bei akuten und chronischen Erkrankungen bei Kindern entwickelt. Es enthält die Domänen Fatigue allgemein, Schlaf/Ruhe und kognitive Fatigue. Zur Testgüte und Änderungssensitivität dieses Moduls wurden keine Unterlagen vorgelegt. In einer Studie wurde bei 837 Kindern und Jugendlichen gezeigt dass die Fragen des Fragebogens in verschiedenen Altersgruppen (5-7, 8-12, 13-18 Jahre), bei beiden Geschlechtern, mit verschiedenen Krankheiten und bei akuten sowie chronischen, milden und schwerwiegenden Erkrankungen in gleicherweise interpretiert werden (Varni et al. 2013).</p> <p><u>Validität und Patientenrelevanz:</u></p> <p>Beim PedsQL-Inventory handelt es sich um ein gut untersuchtes Instrument zur Messung der Lebensqualität bei Kindern und Jugendlichen mit chronischen Erkrankungen mit ausreichender Testgüte. Allerdings liegen keine Daten zur Eignung des Instruments bei Patienten mit DMD vor. Das jüngst publizierte Modul zur DMD weicht bzgl. der erfassten Dimensionen und der Fragen erheblich von der in der</p>

Lebensqualität	
	<p>Studie verwendeten generischen Version ab, so dass unklar ist, ob die in der Studie verwendete Version ausreichend sensitiv ist, um Änderungen der Lebensqualität bei Patienten mit DMD zu erfassen.</p> <p>Für das Fatigue-Modul hat der pU keine Unterlagen vorgelegt, so dass keine abschließende Einschätzung vorgenommen werden kann.</p> <p>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 <u>einer Verbesserung der Lebensqualität.</u>“</p>

Tabelle 9: Beschreibung der Endpunkte der Kategorie Unerwünschte Ereignisse

Unerwünschte Ereignisse	
Endpunkt	Unerwünschte Ereignisse
Operationalisierung und Validierung: Begründung seitens des pU	<p><u>Operationalisierung:</u></p> <p>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 WHODRUG-Dictionary kodiert.</p> <p>Im Einzelnen wurden folgende Maßnahmen durchgeführt:</p> <ul style="list-style-type: none"> • Entnahme von Blutproben für hämatologische und biochemische Analysen: <ul style="list-style-type: none"> ○ Hämatologie: Leukozyten, Hämoglobin, Hämatokrit, Erythrozyten, Thrombozyten ○ Biochemie: Elektrolyte, Harnstoff, Kreatinin, Magnesium, Calcium, Phosphat, Harnsäure, Glukose, Proteine, Bilirubin, Leberwerte (AST, ALT, GGT), CK, LDH, AP, Blutfette, Cystatin C ○ Nebennieren: ACTH, Kortisol, Aldosteron, Renin • Urin: pH-Wert, spezifisches Gewicht, Glukose, Ketone, Blut, Protein, Urobilinogen, Bilirubin, Nitrite, Leukozyten • exploratorisch: neutrophil gelatinase-associated lipocalin (N-GAL), N-acetyl-b-D-glucosaminidase (NAG), kidney injury molecule-1 (KIM-1), clusterin • Vitalzeichen: Puls, Atemfrequenz, Blutdruck, Temperatur während Screening, an Tag 1 (vor Beginn der Studienmedikation), nach jeweils 6 Wochen und zur Visite nach Studienende; 12-Kanal-EKG, Temperatur während Screening, an Tag 1 (vor Beginn der Studienmedikation), nach jeweils 12 Wochen und zur Visite nach Studienende • Ultraschalluntersuchung der Nieren und ableitenden Harnwege: während Screening und nach 48 Wochen • körperliche Untersuchung: während Screening, an Tag 1 (vor Beginn der Studienmedikation), nach jeweils 12 Wochen und zur Visite nach Studienende; Messung von Größe und Gewicht während Screening, an Tag 1 (vor Beginn der Studienmedikation) und nach jeweils 6 Wochen <p>Erhebungszeitpunkte: während Screening, an Tag 1 (vor Beginn der Studienmedikation), nach jeweils 6 Wochen und zur Visite nach Studienende.</p> <p>Wenn ein UE bei einem Patienten mehrmals auftrat, wurde es nur einmal mit dem höchsten CTCAE-Schweregrad dokumentiert.</p> <p>Besonderes Augenmerk wurde auf die Überprüfung der Transaminasen sowie auf</p>

Unerwünschte Ereignisse	
	adrenale und renale Ereignisse gelegt. <u>Begründung des pU gemäß der im Dossier vorhandenen Angaben:</u> Der pU gibt an, dass Nebenwirkungen einen Aspekt der therapiebedingten Morbidität darstellen und daher unmittelbar patientenrelevant sind.
Einschätzung zur Qualität u. Patientenrelevanz	<u>Definition und Erhebung des Endpunkts:</u> Die Operationalisierung des Endpunkts ist nachvollziehbar. <u>Validität und Patientenrelevanz:</u> 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.“

2.6 Ergebnisse zum Zusatznutzen

Da die zulassungsrelevante Dosierung von Ataluren nicht die Hochdosisgruppe umfasst, werden nur die Ergebnisse für die Ataluren-10/10/20mg-Gruppe vs. Placebo dargestellt, mit Ausnahme der Ergebnisse für die unerwünschten Ereignisse (UE). Neuverordnungen von Kortikosteroiden während der Laufzeit der Studie wurden für keinen Patienten in den beiden Gruppen berichtet. Nicht-medikamentöse Begleittherapien (z. B. Physiotherapie, Nutzung von Rollstühlen, Beschäftigungstherapie) war üblich und zwischen den Gruppen vergleichbar (Daten nicht berichtet, vgl. CSR, S. 85).

Compliance mit der Studienmedikation war hoch, lediglich ein Patient in der Hochdosisgruppe brach nach 6 Wochen die Einnahme der Studienmedikation ab. In der Placebo-Gruppe wurde die Dosis bei einem Patienten nach 24 Wochen permanent reduziert. Insgesamt wurden 97,7 % aller geplanten Dosierungen in der Placebo-Gruppe und 97 % in der Ataluren-10/10/20mg-Gruppe eingenommen (patientenberichtete Daten). Nach 24 Wochen wurde bei etwa der Hälfte der Patienten eine Dosisanpassung aufgrund von Gewichtsveränderungen vorgenommen (über das IVR-/IWR-System).

2.6.1 Statistische Methoden

Methoden zur Analyse des primären Endpunkts (Studienprotokoll Version 3.0, S. 75f, SAP Version 3.0, S. 27f)

Für die primäre Analyse (6MWT) wird die ITT-Population herangezogen.

Die statistische Auswertung des primären Endpunkts testet die Änderung im 6MWT nach 48 Wochen im Gruppenvergleich mittels non-parametrischer mixed-model repeated-measure-Kovarianz-Analyse ohne Imputation fehlender Daten. Die Stratifizierungsfaktoren werden in das Modell einbezogen und auf Interaktion geprüft. Die Varianzstruktur wird durch das Akaike Information Criterion bestimmt (compound symmetry, autoregressive, unstructured

covariance). Aus dem Modell werden Mittelwerte und Varianz errechnet und mittels Dunnett-t-Test-Prozedur für die Berücksichtigung multipler Vergleiche auf Signifikanz getestet.

In der ANCOVA werden die geschätzten Randmittelwerte (least squares mean, LS-Mittelwerte) und die 95%-KI für die Behandlungsunterschiede berechnet. Für die Kovarianzanalyse ist eine Normalverteilung der Residuen Voraussetzung. Die Normalverteilung der Residuen wird mittels Shapiro-Wilk-Test geprüft. Ist dieser statistisch signifikant (d. h. für die Studienpopulation kann keine Normalverteilung angenommen werden), ist eine Transformation der Daten (log oder Rang) geplant.

Präspezifizierte supportive Analysen des primären Endpunkts

Geplant ist die Durchführung eines Permutationstests mit 10.000 möglichen Kombinationen von anderen Randomisierungszuteilungen. Zusätzlich wird eine Per-Protokoll-Analyse vorgenommen. Ergänzend wird eine Sensitivitätsanalyse zum Effekt der Zuteilung von Geschwisterkindern in dieselbe Gruppe vorgenommen. Weiterhin werden alle o. g. Analysen nach allen möglichen alternativen Zuteilungen von Geschwisterpaaren im Rahmen einer weiteren Sensitivitätsanalyse wiederholt.

Es werden explorative Analysen zum Einfluss bestimmter Kovariaten vorgenommen: Körpergröße zu Baseline, Art der Mutation, Gewicht zu Baseline, Medikation zur Prophylaxe von Herzinsuffizienz.

Zusätzlich erfolgt die Berechnung des Anteils der Patienten mit einer Verbesserung $\geq 10\%$ im 6MWT nach 48 Wochen (Responder) bzw. mit einer Verschlechterung von $\geq 10\%$ (Progressoren) in der ITT- und Per-Protokoll-Population mittels Cochran-Mantel-Haenszel- und Chi-Quadrat-Test. Hierfür werden Odds Ratios und 95%-Konfidenzintervalle berechnet. Der Einfluss der Stratifizierungs- und anderer Baseline-Variablen wird mittels multipler logistischer Regression ermittelt.

Weiterhin erfolgt die Auswertung der Zeit bis zur Verbesserung/Verschlechterung $\geq 10\%$ im 6MWT (letzter Zeitpunkt, zu dem noch keine $\geq 10\%$ -Verbesserung/-Verschlechterung im Vergleich zum Baseline-Wert vorlag) mittels Kaplan-Meier-Methode und stratifiziertem Log-Rank-Test. Der Einfluss der Stratifizierungsfaktoren wird mittels Cox-Regressionsmodell untersucht.

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

Sensitivitätsanalysen

Verschiedene Sensitivitätstests sollen durchgeführt werden, um die Robustheit der Ergebnisse zu überprüfen (SAP Version 3.0, S. 29f.): Analyse mit der Per-Protokoll-Population, Reallokation der Geschwisterpaare, Effekt durch vorzeitige Entblindung, Effekt durch fehlende Werte (u. a. verschiedene Imputationsmethoden, LOCF-basierte ANCOVA) sowie Re-Randomisierung mittels Permutationstest.

Subgruppenanalysen

Subgruppenanalysen des primären Endpunkts werden für die folgenden Stratifizierungsfaktoren geplant (vgl. CSR, S. 121): Alter (< 9 / ≥ 9 Jahre), 6MWT (≥ 350 / < 350 m), Kortikosteroidbehandlung zu Baseline (ja / nein), Altersgruppen (5-6, 7-8, 9-11, ≥ 12 Jahre), Lokalisation der Nonsense-Mutation (1-39, 40-79) und Art des Stopp-Codons (UGA, UAG, UAA).

Die Subgruppenanalysen wurden nur für die korrigierte ITT-Population und das post hoc modifizierte statistische Auswertungsmodell berichtet. Die Ergebnisse werden daher unter den post-hoc-Ergebnissen berichtet.

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. Die ITT-Population wurde für alle in der Nutzenbewertung berücksichtigten Endpunkte herangezogen. Die Analyse der UE wurde in der Safety-Population durchgeführt.

Korrigierte Intention-to-Treat (kITT)-Population: Wie ITT-Population, bei zwei Patienten wurden die Baseline-Werte für den 6MWT durch die Screening-Werte vor Studieneinschluss ersetzt, um verletzungsbedingte Verzerrungen der 6MWD zu korrigieren.

As-treated-Population: Alle randomisierten Patienten, die die Studienmedikation erhalten, wobei die Zuordnung zur Behandlungsgruppe der aktuell erhaltenen Medikation entspricht. Diese Population wird für die Evaluation der Sicherheit verwendet.

Per-Protocol-Population: Alle randomisierten Patienten, die die Studienmedikation erhalten, wobei die Zuordnung zur Behandlungsgruppe der aktuell erhaltenen Medikation entspricht. Die Patienten haben an der gesamten Studie teilgenommen und weisen einen gültigen Wert im 6MWT zu Woche 48 auf. Die Population wird für die supportiven Analysen herangezogen. Patienten, die entblindet sind, werden aus dieser Population ausgeschlossen.

2.6.2 Ergebnisse für patientenrelevante Endpunkte

Mortalität

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

Morbidität

Tabelle 10: Ergebnisse zum primären Endpunkt in der Studie 007 (ITT-Population)

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Effektschätzer (95%-KI) p-Wert
Absolute Veränderung im 6MWT zu Woche 48^{1,2,3)}			
Baseline (m), Mittelwert (SD)	359,6 (87,7)	350,0 (97,6)	-
Woche 48 (m), Mittelwert (SD)	317,4 (152,3)	342,7 (142,0)	
Veränderung Woche 48 gegenüber Baseline (m), Mittelwert (SD)	-42,6 (90,1)	-12,9 (72,0)	
Veränderung im 6MWT im Gruppenvergleich zu Woche 48⁴⁾			
Gehstreckendifferenz (m)			26,44 (-4,21;57,09) p=0,0905 (nominal) p=0,1592 (Dunnett)
Gehstreckendifferenz			rangtransformierte Daten: p=0,1490 (nominal) p=0,2539 (Dunnett)
Anteil der Patienten mit einer mindestens 10%igen Verschlechterung nach 48 Wochen (Progressoren)⁵⁾			
N	25	15	p=0,0423
Anteil (%)	43,9	26,3	
Anteil der Patienten mit einer mindestens 10%igen Verbesserung nach 48 Wochen (Responder)⁶⁾			
N	6	12	p=0,297
Anteil (%)	10,5	21,1	
Zeit bis zu einer persistierenden Verschlechterung um mindestens 10 %			
			HR 0,52 (0,28;0,966) p=0,0386 (nominal) p=0,078 (Dunnett)
Zeit bis zu einer persistierenden Verbesserung um mindestens 10 %			
			HR 1,675 (0,656;4,277) p=0,2805

¹⁾ In der Placebo- und der Ataluren-10/10/20mg-Gruppe fehlten zu Woche 48 jeweils zwei valide Ergebnisse im 6MWT aufgrund von Frakturen (N=3) bzw. Angst vor Sturz (N=1).

²⁾ Insgesamt 15 Patienten verloren im Studienverlauf ihre Gehfähigkeit (Placebo N=6, Ataluren 10/10/20 N=4, Ataluren 20/20/40 N=5), die 6MWT-Werte dieser Patienten wurden ab dem Zeitpunkt des Verlusts der Gehfähigkeit mit 0 registriert.

³⁾ Quelle: CSR Tab. 21, S. 92; LOCF-Methode lieferte vergleichbare Werte.

⁴⁾ MMRM-Analyse der nicht transformierten Werte, ITT-Population (unabhängige Variablen sind Baseline-Wert für 6MWT, Behandlung, Visite, Alter und Kortikosteroide als Stratifizierungsfaktoren sowie Interaktion zwischen Behandlung und Visite), unstrukturierte Kovarianzmatrix, Quelle: CSR, S. 94, Tab. 23, Tab. 14.2.1.8.5B, S. 456; Modul 4, S. 90. Der p-Wert der rangtransformierten Daten wurde aufgrund der nicht normalverteilten Daten als primär angesehen.

⁵⁾ Quelle: CSR, S. 95, Tab. 14.2.2.11, S. 1019, S. 101, S. 1008. p-Werte aus Cochran-Mantel-Haenszel-Test. HR aus time-to-event-Analyse im Cox proportional hazards model errechnet mit den Kovariaten Behandlung, Alter (< / ≥ 9 Jahre), Kortikosteroide (ja/nein) und Baseline-Wert des 6MWT.

⁶⁾ Quelle: CSR, S. 96, Tab. 14.2.2.10, S. 101, S. 1006, S. 1016. p-Werte aus Cochran-Mantel-Haenszel-Test. HR aus time-to-event-Analyse im Cox proportional hazards model errechnet mit den Kovariaten Behandlung, Alter (< / ≥ 9 Jahre), Kortikosteroide (ja/nein) und Baseline-Wert des 6MWT.

Das primäre Ziel der Studie 007, der Nachweis der Überlegenheit von Ataluren vs. Placebo in der Behandlung von Patienten mit DMD für den primären Endpunkt 6MWT, wurde nicht erreicht. Der geschätzte Behandlungsunterschied zwischen dem Ataluren-10/10/20mg- und

dem Placebo-Arm von 26,4 m zugunsten von Ataluren war in der primären (präspezifizierten Analyse) statistisch nicht signifikant und blieb unter der vom pU ermittelten klinischen Relevanzschwelle.

Die Varianz der Gehstrecke wurde im Studienverlauf deutlich größer (vgl. CSR, S. 93) und erreichte 90 m in der Placebo- und 72 m in der Ataluren-10/10/20mg-Gruppe, wodurch die Power der Analyse beeinträchtigt wurde. Diese Entwicklung war zum Zeitpunkt der im Studienprotokoll geplanten ersten Interimsanalyse (nach ≥ 45 Patienten bzw. ≥ 12 Wochen Behandlung mit der Fragestellung der Ausweitung der Studienpopulation) noch nicht abzusehen, so dass die Fallzahl nicht erhöht wurde. Die geplante zweite Interimsanalyse (nach 24 Wochen bzw. ≥ 90 Patienten) fand nicht statt, da zu diesem Zeitpunkt der pU und die FDA den SAP diskutierten und auch keine besorgniserregenden Sicherheitssignale vorhanden waren (CSR, S. 121).

Verschiedene Sensitivitätsanalysen (u. a. ANCOVA bzw. MMRM mit LOCF und jeweils nicht transformierten Daten, verschiedene Imputationsmethoden, Reallokation der Geschwisterpaare) und die Wiederholung der Analyse mit der Per-Protocol-Population bestätigten das Ergebnis, allerdings zeigte sich in der ANCOVA mit LOCF an nicht transformierten Daten unerwartet ein signifikanter p-Wert (nominal 0,0445, adjustiert 0,0808), was dazu Anlass gab, post hoc das Auswertungsmodell zu modifizieren (u. a. Berücksichtigung eines Baseline-Visiten-Interaktionsterms; vgl. CSR, S. 73). Der Anteil der Progressoren nach 48 Wochen war im Ataluren-10/10/20mg-Arm geringer als im Placebo-Arm (26 vs. 44 %), entsprechend einer relativen Reduktion von 48 %. Das obere Konfidenzintervall für diese supportive Analyse betrug 0,966, mit einem zugehörigen nominalen p-Wert von 0,0386 und einem adjustierten p-Wert von 0,078. Der Anteil der Responder unterschied sich nicht signifikant zwischen den Gruppen.

Eine weitere Sensitivitätsanalyse unter Ausschluss der beiden Patienten mit ungültigen Baseline-Werten für den 6MWT ergab ähnliche Ergebnisse wie in der Analyse der kITT-Population (Gehstreckendifferenz nach 48 Wochen in der Placebo-Gruppe -45,45 m (SD: 88,279), in der Ataluren-10/10/20mg-Gruppe unverändert). Ergebnisse für den statistischen Vergleich werden allerdings nur für das post hoc modifizierte Modell berichtet (Gehstreckendifferenz 32,7 m, $p=0,0255$ (nominal), $p=0,0504$ (Dunnett-Adjustierung) (CSR, S. 98, 1027).

Post-hoc-Auswertungen zum primären Endpunkt

Nach Entblindung der Daten wurden post hoc weitere Auswertungen vorgenommen. Wie bereits beschrieben, wurden die Baseline-Werte für zwei Patienten (jeweils ein Patient in der Placebo- bzw. Ataluren-20/20/40mg-Gruppe) durch Screening-Werte ersetzt (vgl. Tabelle 3). Retrospektiv gesehen zeigte sich, dass für beide Patienten der Baseline-Wert im Vergleich zur Performance zu allen anderen Erhebungszeitpunkten auffällig niedrig war, was durch sturzbedingte Verletzungen unmittelbar vor der Baseline-Erhebung erklärt wurde (vgl. CSR Abb. 5, S. 98). Durch den Ersatz der Baseline- durch die im Screening erhobenen Werte ergab sich eine korrigierte ITT-Population (kITT). Hierdurch änderten sich die Ergebnisse wie folgt (Quelle: CSR, S. 99, 341 und Tab. 27, S. 101): absolute Änderung der Gehstrecke in der Placebo-Gruppe zu Woche 48 (Mittelwert, SD): -44,14 m (88), in der Ataluren-10/10/20mg-Gruppe keine Veränderung. Effektschätzer für Gehstreckendifferenz im modifizierten MMRM:

Dossierbewertung für Orphan Drugs

Ataluren

31,71 m (95%-KI 5,13;58,28), $p=0,0367$ (Dunnnett-t-Test). Entsprechend wurde die supportive Analyse für die Progressoren und Responder mit kITT-Population wiederholt. Für Progressoren zeigten sich keine Unterschiede in den Anteilen im Vergleich zur Auswertung der ITT-Population, die Hazard Ratios unterschieden sich nur unwesentlich voneinander. Bei den Respondern änderten sich die Anteile, da die beiden Patienten in der präspezifizierten Analyse zu den Respondern zählten, in der kITT-Analyse jedoch zu den Non-Respondern. Der Anteil der Responder in der Placebo-Gruppe änderte sich auf 8,8 % (N=5), in der Ataluren-10/10/20mg-Gruppe blieb der Anteil unverändert. Der Unterschied war weiterhin nicht statistisch signifikant.

Die Ergebnisse der Subgruppenanalyse zeigten eine bessere Wirksamkeit im Ataluren-10/10/20mg-Arm im Vergleich zu Placebo unabhängig von der Subgruppe (CSR, S. 129). Die Wirksamkeit war in der Subgruppe der <9-Jährigen am stärksten ausgeprägt (mittlere Differenz 38 m, 95%-KI 1;76), bei Patienten mit Kortikosteroidtherapie (mittlere Differenz 36 m, 95%-KI 2;70) und bei Patienten mit einem 6MWT zu Baseline <350 m (mittlere Differenz 50 m, 95%-KI 6;94). Der Unterschied zwischen Ataluren-10/10/20 mg und Placebo betrug bei den Patienten in der Altersgruppe 5-6 Jahre 20,5 m, 7-8 Jahre 57,1 m, 9-11 Jahre 7,4 m und bei den ≥ 12 -Jährigen 29,3 m. Die entsprechenden Ergebnisse für die Lokalisation der Nonsense-Mutation sind: 1-39: 24,9 m, 40-79: 30,2 m. Für die Art des Stopp-Codons sind die Ergebnisse: UGA: 33,1 m, UAG: 1,8 m, UAA: 44,7 m.

Tabelle 11: Ergebnisse zu sekundären Endpunkten in der Studie 007 (ITT-Population)

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Effektschätzer (95%-KI) p-Wert
Änderung in der proximalen Muskelfunktion (,Timed Function Tests') zu Woche 48¹⁾			
<i>Aufstehen aus Rückenlage</i>			
Baseline (Sekunden), Mittelwert 48 Wochen	11,5 14,6	10,8 14,0	Differenz (absolut, 95%-KI): 0 (-2,3;2,3) relative Differenz: 0%
<i>Score für die Methode zum Aufstehen aus Rückenlage²⁾</i>			
Baseline, Mittelwert 48 Wochen	3,6 3,3	3,7 3,4	Differenz (absolut, 95%-KI): 0,1 (-0,2;0,5)
<i>10 m laufen/gehen</i>			
Baseline (Sekunden), Mittelwert 48 Wochen	6,9 9,9	7,5 9,1	Differenz (absolut, 95%-KI): -1,3 (-3,7;0,9) relative Differenz: -43%
<i>Score für die Methode um 10 m zu laufen/gehen³⁾</i>			
Baseline, Mittelwert 48 Wochen	4,8 4,3	4,7 4,4	Differenz (absolut, 95%-KI): 0,2 (-0,1;0,6)

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Effektschätzer (95%-KI) p-Wert
<i>benötigte Zeit, um 4 Treppenstufen hinaufzusteigen</i>			
Baseline (Sekunden), Mittelwert 48 Wochen	6,0 10,8	6,9 9,3	Differenz (absolut, 95%-KI): -2,4 (-4,9;0,1) relative Differenz: -50%
<i>Score für die Methode zum Treppenhochsteigen⁴⁾</i>			
Baseline, Mittelwert 48 Wochen	4,0 3,7	3,6 3,6	Differenz (absolut, 95%-KI): 0,5 (0;0,9)
<i>benötigte Zeit, um 4 Treppenstufen hinunterzusteigen</i>			
Baseline (Sekunden), Mittelwert 48 Wochen	5,5 9,6	6,1 8,5	Differenz (absolut, 95%-KI): -1,6 (-4,3;1,0) relative Differenz: -40%
<i>Score für die Methode zum Treppenabsteigen⁵⁾</i>			
Baseline, Mittelwert 48 Wochen	3,7 3,6	3,3 3,4	Differenz (absolut, 95%-KI): 0,3 (-0,2;0,6)
Änderung der Alltagsaktivitäten gemessen mittels ‚Step Activity Monitoring‘ (Pedometer am Knöchel) an \geq 9 Tagen nach jeder Studienvisite⁶⁾			
<i>keine Aktivität</i>			
Baseline, Anteil in Prozent 48 Wochen Δ nach 48 Wochen (%)	48,17 51,33 4,08	51,88 54,83 2,78	Differenz (95%-KI): -1,3 (-5,51; 2,9)
<i>geringe Aktivität</i>			
Baseline, Anteil in Prozent 48 Wochen Δ nach 48 Wochen (%)	32,86 32,1 -1,11	32,38 31,31 -1,12	Differenz (95%-KI): -0,01 (-2,9; 2,88)
<i>mäßige Aktivität</i>			
Baseline, Anteil in Prozent 48 Wochen Δ nach 48 Wochen (%)	11,84 10,33 -1,92	10,0 9,15 -0,69	Differenz (95%-KI): 1,23 (-0,25;2,71)
<i>ausgeprägte Aktivität</i>			
Baseline, Anteil in Prozent 48 Wochen Δ nach 48 Wochen (%)	7,17 6,44 -1,03	5,78 4,83 -0,96	Differenz (95%-KI): 0,07 (-1,18;1,31)
patientenberichtete Tage mit Rollstuhlnutzung⁷⁾			
Baseline, Anteil in Prozent 48 Wochen Δ nach 48 Wochen (%)	13,24 (N=53) 25,06 (N=46) 11,45 (N=44)	13,17 (N=50) 19,04 (N=47) 3,95 (N=44)	Differenz: 7,5 % (nicht signifikant)

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Effektschätzer (95%-KI) p-Wert
patientenberichtete Änderung in der Häufigkeit von Stürzen pro Tag (Patiententagebuch)⁸⁾			
Baseline, Mittelwert (SD)	0,54 (0,943) (N=48)	0,27 (0,48) (N=48)	Relatives Risiko (95%-KI): 0,384 (0,16;0,94) p=0,0356
48 Wochen	0,72 (1,283) (N=53)	0,23 (0,531) (N=55)	
Δ nach 48 Wochen	0,2 (1,282) (N=44)	-0,06 (0,501) (N=47)	

¹⁾ Quelle: CSR, S. 106f, Tab. 28

²⁾ Entsprechend Studienprotokoll (Version 3.0, S. 59) Einteilung in die folgenden 6 Kategorien: 1 = unfähig, aus Rückenlage aufzustehen; 2 = Aufstehen mit abstützen an Möbeln (*assisted Gowers*); 3 = auf Bauch drehen, Aufrichten mittels abstützen an den Beinen (*full Gowers*); 4 = auf Bauch drehen, Aufrichten mittels Abstützen mit einer Hand an den Beinen (*half Gowers*); 5 = auf die Seite drehen und/oder aufstehen nach abstützen auf dem Boden, aber nicht an den Beinen; 6 = Aufstehen ohne drehen oder abstützen. Höhere Werte repräsentieren ein besseres Ergebnis.

³⁾ Entsprechend Studienprotokoll (Version 3.0, S. 59) 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⁾ Entsprechend Studienprotokoll (Version 3.0, S. 60) 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 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.

⁵⁾ Der Score wurde analog zum Treppenhochsteigen gebildet: 1 = unfähig (...) 6 = ohne Hilfe möglich. Höhere Werte repräsentieren ein besseres Ergebnis.

⁶⁾ keine Aktivität = 0 Schritte/Min.; gering = ≤ 15 Schritte/Min.; mäßig = 16-30 Schritte/Min.; ausgeprägt = > 30 Schritte/Min.
Quelle: CSR, S. 113, Tab. 31, Tab. 14.2.5.11.3, S. 1753, Tab. 14.2.5.8.1, S. 1709, Tab. 14.2.5.9.1, S. 1720, Tab. 14.2.5.10.1, S. 1731.

⁷⁾ Quelle: CSR, S. 113, Tab. 14.2.6.1, S. 1772, Tab. 14.2.6.2, S. 1776. Daten wurden jeweils an den Tagen erhoben, an denen das Pedometer getragen wurde (SAM Diary).

⁸⁾ Quelle: CSR, S. 114, Tab. 14.2.9.1, S. 1941 und Tab. 14.2.9.4B, S. 1959. Die Analyse erfolgte mittels GEE-Prozedur unter der Annahme einer negativen Binomialverteilung für die Anzahl der Stürze (vgl. SAP S. 32).

Zu Baseline waren insgesamt 40/174 (23 %) Patienten nicht in der Lage, aus der Rückenlage aufzustehen bzw. innerhalb des vorgegebenen Zeitfensters von 30 Sekunden aufzustehen. Laut CSR (S. 1163) waren zu Baseline 3 Patienten (5,3 %) in der Ataluren-10/10/120mg-Gruppe und 4 Patienten (7 %) in der Placebo-Gruppe der Kategorie 1 zuzuordnen. Bei den übrigen Timed Function Tests (TFT) war das nur bei 3 % der Patienten der Fall. Die Relevanzschwelle von 1,5 Sekunden beim Treppensteigen wurde erreicht, war aber statistisch nicht signifikant.

Aufgrund des hohen Anteils fehlender Werte zu Baseline (N=12 je Gruppe, entsprechend 21 %) wurde eine Sensitivitätsanalyse für den Endpunkt Stürze durchgeführt, bei der die Baseline- durch die 6-Wochen-Werte ergänzt wurden. Dadurch wurde die Anzahl der fehlenden Werte um N=10 reduziert. Dies resultierte in einer Differenz von -0,7038 (95%-KI -1,31;-0,1), p=0,0229 (vgl. CSR Tab. 14.2.9.4D, S. 1961). Das Relative Risiko (relative ratio) betrug 0,4947 (95%-KI 0,27;0,91).

Post-hoc-Auswertungen zur Änderung in der proximalen Muskelfunktion („Timed Function Tests“)

Durch den Ersatz der Baseline- durch die Screening-Werte ergab sich eine korrigierte ITT-Population (kITT). Hierdurch änderten sich die Ergebnisse wie folgt (Quelle: CSR, S. 109 und Tab. 29, S. 111): die Differenz in der Zeit (Sekunden) beim Aufstehen aus der Rückenlage betrug -0,1 (-2,3;2,3), für den Score 0,1 (-0,2;0,5) und war damit praktisch unverändert im Vergleich zur präspezifizierten ITT-Auswertung (Ataluren-10/20/20 mg vs. Placebo). Im 10-Meter-Test betrug die Differenz -1,5 (-3,7;0,7) Sekunden (relativ -47 %) zugunsten Ataluren-10/10/20 mg und die Differenz im Score 0,2 (-0,1;0,6). Die Zeitdifferenz, um 4 Treppenstufen hinaufzusteigen, betrug -2,4 (-4,9;0,0) Sekunden (relativ -50 %), der Score lag bei 0,4 (0,1;0,9). Die Zeitdifferenz, um 4 Treppenstufen herunterzusteigen, betrug -1,6 (-4,3;1,0) Sekunden (-41 %), der Score lag bei 0,2 (-0,2;0,6). Diese Werte waren ebenfalls kaum verändert im Vergleich zur präspezifizierten ITT-Auswertung.

Post-hoc-Auswertungen zu Änderungen der Alltagsaktivitäten, zur Rollstuhlnutzung sowie zu Stürzen wurden weder im CSR noch im Dossier berichtet.

Lebensqualität

Tabelle 12: Ergebnisse zur Lebensqualität in der Studie 007 (ITT-Population)

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	LS-MW-Differenz (95%-KI) p-Wert
Änderung in der Lebensqualität gemessen mit dem PedsQL (physikalische, emotionale, soziale, schulische Subskalen)¹⁾			
<i>physikalische Subskala</i>			
Baseline, Mittelwert (SD)	61,87 (19,411)	59,27 (22,782)	3,56 (-4,31;11,42)
Woche 48	59,53 (23,124)	62,61 (23,07)	p=0,3733 (nominal)
Δ nach 48 Wochen	-1,00 (24,022)	2,37 (25,105)	p=0,5721 (Dunnett)
<i>emotionale Subskala</i>			
Baseline, Mittelwert (SD)	70,13 (19,332)	73,7 (20,223)	-4,2 (-11,5;3,05)
Woche 48	73,8 (22,351)	72,82 (20,018)	p=0,2543 (nominal)
Δ nach 48 Wochen	4,3 (22,315)	-1,83 (23,725)	p=0,4107 (Dunnett)
<i>soziale Subskala</i>			
Baseline, Mittelwert (SD)	63,36 (20,476)	65,09 (18,421)	-2,38 (-9,52;4,76)
Woche 48	69,9 (23,356)	68,64 (21,092)	p=0,5117 (nominal)
Δ nach 48 Wochen	7,75 (18,87)	3,89 (21,841)	p=0,7328 (Dunnett)
<i>schulische Subskala</i>			
Baseline, Mittelwert (SD)	64,65 (17,841)	64,55 (20,396)	2,54 (-4,42;9,5)
Woche 48	68,06 (21,356)	70,82 (18,022)	p=0,4727 (nominal)
Δ nach 48 Wochen	4,06 (23,244)	6,11 (23,765)	p=0,6907 (Dunnett)

Studie 007	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	LS-MW-Differenz (95%-KI) p-Wert
PedsQL Fatigue Scale²⁾			
Baseline, Mittelwert (SD)	69,7 (15,263)	71,62 (16,474)	-2,41 (-8,85;4,03)
Woche 48	72,85 (17,261)	72,62 (20,789)	p=0,4611 (nominal)
Δ nach 48 Wochen	3,92 (16,512)	0,45 (23,068)	p=0,6783 (Dunnett)

¹⁾ Quelle: CSR, S. 114, Tab. 14.2.7.1 S. 1779, Tab. 14.2.7.2B S. 1785, Tab. 14.2.7.4 S. 1801, 14.2.7.6B S. 1807, Tab. 14.2.7.9 S. 1823, Tab. 14.2.7.10B S. 1829, Tab. 14.2.7.13 S. 1845, Tab. 14.2.7.14B S. 1851. MMRM-Analyse der ITT-Population (unabhängige Variablen sind Baseline-Wert für jeweiligen PedsQL-Domain-Score, Behandlung, Visite, Interaktion zwischen Behandlung und Visite), Stratifizierungsfaktoren, p-Wert nominal und aus Dunnett-t-Test.

²⁾ Quelle: CSR, S. 1889, Tab. 14.2.7.21, Tab. 14.2.7.22B S. 1895. MMRM-Analyse der ITT-Population (unabhängige Variablen sind Baseline-Wert für PedsQL-Fatigue-Score, Behandlung, Visite, Interaktion zwischen Behandlung und Visite), Stratifizierungsfaktoren, p-Wert nominal und aus Dunnett-t-Test.

Post-hoc-Auswertungen zur Lebensqualität wurden nicht berichtet.

Unerwünschte Ereignisse

Die Sicherheitspopulation ("as-treated") umfasste alle randomisierten Patienten, die ein Prüfpräparat entsprechend ihrer Randomisierung erhalten hatten. Unerwünschte Ereignisse (UE) werden für die Ataluren-20/20/40mg-Gruppe ergänzend tabellarisch dargestellt, um ein möglichst umfassendes Bild vom Nebenwirkungsprofil von Ataluren zu erhalten.

Ausmaß der Exposition

Die mediane Dauer der Behandlung war in den Gruppen vergleichbar: Placebo 48,1 (Spanne 46,3-50,7) Wochen, Ataluren-10/10/20 mg 48,1 (47,0-49,1) Wochen, Ataluren-20/20/40 mg 48,1 (6,0-50,3) Wochen. In der Hochdosisgruppe hatte ein Patient die Therapie nach 6 Wochen aufgrund von Protokollverletzung abgebrochen, bei einem weiteren Patienten in der Hochdosisgruppe wurde die Dosis unterbrochen und für 9 Wochen aufgrund eines Hautausschlags reduziert.

Tabelle 13: Ergebnisse der Nebenwirkungen – Zusammenfassung der UE in Studie 007

Patienten mit ...	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
mindestens 1 UE, N (%)	56 (98,2)	55 (96,5)	57 (95)
UE nach Schweregrad, N (%)			
Grad 1 (mild)	21 (36,8)	16 (28,1)	20 (33,3)
Grad 2 (moderat)	26 (45,6)	31 (54,4)	27 (45)
Grad 3 (schwerwiegend)	9 (15,8)	8 (14)	10 (16,7)
Grad 4 (lebensbedrohlich)	0	0	0
UE nach Kausalität, N (%)			
kein Zusammenhang	14 (24,6)	8 (14)	11 (18,3)
Zusammenhang unwahrscheinlich	16 (28,1)	17 (29,8)	13 (21,7)
Zusammenhang möglich	20 (35,1)	25 (43,9)	29 (48,3)
Zusammenhang wahrscheinlich	6 (10,5)	5 (8,8)	4 (6,7)
Therapieabbruch aufgrund von UE	0	0	0
schwerwiegenden UE, N (%)	3 (5,3)	2 (3,5)	2 (3,3)
Todesfälle	0	0	0

Quelle: CSR, S. 133, Tab. 39; für einen Patienten kann mehr als ein Ereignis berichtet werden.

Unerwünschte Ereignisse

Das Nebenwirkungsprofil war zwischen den drei Behandlungsarmen vergleichbar. Therapieabbrüche aufgrund von UE traten nicht 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 Gastrointestinaltrakt, Infektionen, Verletzung, Vergiftung, prozedurale Komplikationen sowie Muskuloskelettal- und Nervensystem beobachtet. Am häufigsten traten Erbrechen (bei insgesamt 46,6 % der Patienten über alle Behandlungsarme), Kopfschmerzen (29,3 %), Diarrhö (24,1 %), Nasopharyngitis (20,7 %), Fieber (19 %), Husten (19 %) und Oberbauchschmerzen (17,8 %) auf. Für diese Symptome zeigten sich keine statistisch signifikanten Unterschiede zwischen den Gruppen. UE, die in den Behandlungsarmen Placebo, Ataluren in der niedrigen und Ataluren in der hohen Dosis mit steigender Häufigkeit auftraten, waren Übelkeit (12,3, 14, 16,7 %), Bauchschmerzen (7, 12,3, 16,7 %), Extremitätenschmerzen (10,5, 12,3, 13,3 %), Flatulenz (7, 8,8, 11,7 %) und verstopfte Nase (7, 8,8, 10 %).

Bei insgesamt 44 Patienten (13 Placebo, 22 Ataluren 10/10/20 mg und 9 Ataluren 20/20/40 mg) wurde die Dosis kurzzeitig (1-2 Tage), in der Regel aufgrund gastrointestinaler Beschwerden, unterbrochen. Bei 4 Patienten wurde die Dosis aufgrund möglicher UE reduziert (2 Placebo, je 1 in den Ataluren-Gruppen).

Tabelle 14: UE in der Studie 007 unabhängig von der Kausalität, die bei > 5% der Patienten auftraten (as-treated-Population)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
Gastrointestinaltrakt	37 (64,9)	42 (73,7)	44 (73,3)
Erbrechen	22 (38,6)	32 (56,1)	27 (45)
Diarrhö	14 (24,6)	11 (19,3)	17 (28,3)
Oberbauchschmerzen	9 (15,8)	9 (15,8)	13 (21,7)
Übelkeit	7 (12,3)	8 (14)	10 (16,7)
Bauchschmerzen	4 (7)	7 (12,3)	10 (16,7)
Flatulenz	4 (7)	5 (8,8)	7 (11,7)
Magenbeschwerden	0	4 (7)	5 (8,3)
Allgemeine Erkrankungen	21 (36,8)	23 (40,4)	20 (33,3)
Fieber	12 (21,1)	14 (24,6)	7 (11,7)
Krankheitsprogression	6 (10,5)	4 (7)	5 (8,3)
Kraftlosigkeit (Asthenie)	2 (3,5)	3 (5,3)	4 (6,7)
Infektionen	43 (75,4)	38 (66,7)	39 (65)
Nasopharyngitis	13 (22,8)	13 (22,8)	10 (16,7)
Infektion der oberen Atemwege	10 (17,5)	9 (15,8)	11 (18,3)
Influenza	8 (14)	6 (10,5)	7 (11,7)
Gastroenteritis	4 (7)	9 (15,8)	3 (5)
Rhinitis	2 (3,5)	6 (10,5)	3 (5)
Ohrinfektion	3 (5,3)	3 (5,3)	4 (6,7)
virale Gastroenteritis	3 (5,3)	4 (7)	3 (5)
Verletzung, Vergiftung, prozedurale Komplikation	26 (45,6)	28 (49,1)	31 (51,7)
Sturz	7 (12,3)	11 (19,3)	6 (10)
verfahrensbedingte Schmerzen	7 (12,3)	6 (10,5)	8 (13,3)
Kontusion	3 (5,3)	6 (10,5)	4 (6,7)
Gelenkverstauchung	1 (1,8)	4 (7)	4 (6,7)
Untersuchungen	4 (7)	10 (17,5)	6 (10)
Gewichtsverlust	1 (1,8)	5 (8,8)	3 (5)
Stoffwechsel und Ernährung	3 (5,3)	7 (12,3)	6 (10)
verminderter Appetit	2 (3,5)	5 (8,8)	5 (8,3)
Muskuloskelettal und Bindegewebe	19 (33,3)	25 (43,9)	28 (46,7)
Extremitätenschmerzen	6 (10,5)	7 (12,3)	8 (13,3)
Rückenschmerzen	5 (8,8)	9 (15,8)	6 (10)
Arthralgie	2 (3,5)	2 (3,5)	6 (10)
Muskelkrämpfe	5 (8,8)	3 (5,3)	1 (1,7)
Muskelschwäche	1 (1,8)	3 (5,3)	5 (8,3)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
Nervensystem	17 (29,8)	25 (43,9)	18 (30)
Kopfschmerzen	14 (24,6)	22 (38,6)	15 (25)
Schwindel	4 (7)	3 (5,3)	3 (5)
Respirationstrakt, Brustkorb, Mediastinum	18 (31,6)	20 (35,1)	22 (36,7)
Husten	11 (19,3)	9 (15,8)	13 (21,7)
verstopfte Nase	4 (7)	5 (8,8)	6 (10)
oropharyngeale Schmerzen	4 (7)	6 (10,5)	4 (6,7)
Rhinorrhö	6 (10,5)	4 (7)	0
Haut, subkutanes Gewebe	18 (31,6)	19 (33,3)	14 (23,3)
Hautausschlag	5 (8,8)	4 (7)	8 (13,3)
Narbe	3 (5,3)	4 (7)	5 (8,3)

Quelle: CSR, S. 135, Tab. 40 und S. 138, Tab. 41

Der Schweregrad der meisten UE war mild oder moderat. Das häufigste als Grad 3 eingestufte UE war eine Krankheitsprogression (Verlust der Gehfähigkeit). Lebensbedrohliche UE oder Todesfälle traten nicht auf (Tabelle 15).

Tabelle 15: UE Grad 3 in der Studie 007 (as-treated-Population)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
Ohr und Innenohr	0	0	1 (1,7)
Ohrenschmerzen	0	0	1 (1,7)
Gastrointestinaltrakt	1 (1,8)	2 (3,5)	1 (1,7)
Bauchschmerzen	0	1 (1,8)	0
Verstopfung	0	1 (1,8)	0
Diarrhö	0	0	1 (1,7)
Ileus	0	1 (1,8)	0
Erbrechen	1 (1,8)	0	1 (1,7)
Allgemeine Erkrankungen	6 (10,5)	4 (7)	5 (8,3)
Krankheitsprogression	6 (10,5)	4 (7)	5 (8,3)
Infektionen	0	0	1 (1,7)
virale Gastroenteritis	0	0	1 (1,7)
Verletzung, Vergiftung, prozedurale Komplikation	1 (1,8)	2 (3,5)	2 (3,3)
Kontusion	0	1 (1,8)	0
Sturz	0	1 (1,8)	0
Femurfraktur	1 (1,8)	0	0
Fußfraktur	0	1 (1,8)	0

SOC / Preferred Term MedDRA, N (%)	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
Beinfraktur	0	0	1 (1,7)
Handgelenkfraktur	0	0	1 (1,7)
Muskuloskelettal und Bindegewebe	0	2 (3,5)	1 (1,7)
muskuloskelettale Thoraxschmerzen	0	1 (1,8)	0
Extremitätenschmerzen	0	1 (1,8)	0
Osteoporose	0	0	1 (1,7)
Nervensystem	1 (1,8)	0	1 (1,7)
Grand-Mal-Anfall	1 (1,8)	0	0
Hypotonie	0	0	1 (1,7)
keine Reaktion auf Stimuli	0	0	1 (1,7)
Chirurgische und medizinische Prozeduren	0	1 (1,8)	0
endodontische Prozedur	0	1 (1,8)	0

Quelle: CSR, S. 141, Tab. 42

Die Auswertung der UE in Abhängigkeit von der gleichzeitigen Einnahme von Kortikosteroiden ergab, dass Kopfschmerzen (35,5 vs. 14 % bezogen auf beide Dosisgruppen) bei diesen Patienten häufiger auftraten, als bei Patienten ohne Kortikosteroide (CSR, S. 145). Häufiger traten auch Rückenschmerzen (15,3 vs. 2 %), Influenza (14,5 vs. 6 %), Extremitätenschmerzen (13,7 vs. 8 %) und Muskelkrämpfe (6,5 vs. 2 %) auf. Alle Fälle von Hypertonie (4 %) traten bei Patienten mit Kortikosteroiden auf. Umgekehrt wurde häufiger über Fieber (26 vs. 16,1 %), Krankheitsprogression (12 vs. 7,3 %), Rhinitis (10 vs. 4,8 %), virale Gastroenteritis (10 vs. 4 %), Rhinorrhoe (8 vs. 4,8 %), Lymphadenopathie (6 vs. 0,8 %), Sinusitis (6 vs. 3,2 %), virale Infektionen (6 vs. 3,2 %), Hautverletzungen (6 vs. 0 %) sowie saisonale Allergien (10 vs. 0 %) bei Patienten ohne Kortikosteroide berichtet (vgl. CSR, S. 146). Relevante Unterschiede zwischen den Behandlungsgruppen waren nicht zu erkennen. Das Nebenwirkungsprofil zeigte keine erkennbare Abhängigkeit vom Alter oder vom 6MWT zu Baseline.

In Zusammenhang mit der Studienmedikation stehende UE

Die Einschätzung, ob ein UE mit der Studienmedikation in Zusammenhang stand, wurde von den Prüfarzten vorgenommen. Diese UE waren in den drei Behandlungsarmen etwa gleich häufig, allerdings traten Erbrechen, Flatulenz, Magenbeschwerden und Fatigue häufiger in den Ataluren-Gruppen als in der Placebo-Gruppe auf, wobei es einen Dosisgradienten gab (Tabelle 16).

Tabelle 16: Mit der Behandlung in möglich oder wahrscheinlich in Zusammenhang stehende UE in der Studie 007, die bei > 1 Patienten auftraten (as-treated-Population)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
Gastrointestinaltrakt	18 (31,6)	21 (36,8)	28 (46,7)
Bauchbeschwerden	1 (1,8)	0	2 (3,3)
Bauchschmerzen	3 (5,3)	1 (1,8)	4 (6,7)
Oberbauchschmerzen	5 (8,8)	5 (8,8)	8 (13,3)
Druckempfindlichkeit des Bauchs	1 (1,8)	0	2 (3,3)
Aerophagie	2 (3,5)	0	1 (1,7)
Verstopfung	1 (1,8)	1 (1,8)	2 (3,3)
Diarrhö	7 (12,3)	3 (5,3)	12 (20)
Flatulenz	3 (5,3)	5 (8,8)	7 (11,7)
häufiger Stuhlgang	1 (1,8)	0	1 (1,7)
Übelkeit	0	7 (12,3)	7 (11,7)
Regurgitation	0	1 (1,8)	1 (1,7)
Magenbeschwerden	0	2 (3,5)	4 (6,7)
Erbrechen	7 (12,3)	11 (19,3)	13 (21,7)
Allgemeine Erkrankungen	2 (3,5)	2 (3,5)	5 (8,3)
Fatigue	0	1 (1,8)	2 (3,3)
Fieber	0	1 (1,8)	1 (1,7)
Untersuchungen	3 (5,3)	6 (10,5)	1 (1,7)
nicht normale Urinfarbe	1 (1,8)	1 (1,8)	0
Gewichtsverlust	1 (1,8)	4 (7)	1 (1,7)
Stoffwechsel und Ernährung	2 (3,5)	5 (8,8)	4 (6,7)
verminderter Appetit	2 (3,5)	5 (8,8)	4 (6,7)
Muskuloskelettal und Bindegewebe	0	3 (5,3)	2 (3,3)
Extremitätenschmerzen	0	2 (3,5)	1 (1,7)
Nervensystem	4 (7)	8 (14)	5 (8,3)
Schwindel	1 (1,8)	1 (1,8)	1 (1,7)
Kopfschmerzen	2 (3,5)	7 (12,3)	4 (6,7)
Niere, harnableitende Wege	2 (3,5)	6 (10,5)	1 (1,7)
Enuresis	0	2 (3,5)	0
häufige Blasenentleerung	1 (1,8)	1 (1,8)	1 (1,7)
renale Zyste	1 (1,8)	1 (1,8)	0
Respirationstrakt, Brustkorb, Mediastinum	0	4 (7)	3 (5)
Husten	0	2 (3,5)	2 (3,3)
Nasenbluten	0	1 (1,8)	1 (1,7)
Haut, subkutanes Gewebe	2 (3,5)	1 (1,8)	2 (3,3)
Erythem	0	1 (1,8)	1 (1,7)
Hautausschlag	1 (1,8)	0	2 (3,3)

SOC / Preferred Term MedDRA, N (%)	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
Gefäßerkrankungen	2 (3,5)	3 (5,3)	0
Hypertonie	1 (1,8)	3 (5,3)	0

Quelle: CSR, S. 148, Tab. 45

Schwere und schwerwiegende UE (SUE)

Schwere UE traten bei insgesamt 9 Patienten auf, bei 5 Patienten in der Placebo-Gruppe und bei jeweils 2 Patienten in den Ataluren-Gruppen (Tabelle 17). Im CSR werden noch zwei weitere Ereignisse, die während der Screening-Phase auftraten, gelistet.

Tabelle 17: Schwerwiegende UE in der Studie 007 (as-treated-Population)

Preferred Term MedDRA, N (%)	Placebo (N=57)	Ataluren 10/10/20 mg (N=57)	Ataluren 20/20/40 mg (N=60)
Patienten insgesamt	3 (5,3)	2 (3,5)	2 (3,3)
Ereignisse insgesamt	5 (8,8)	2 (3,5)	2 (3,3)
Bauchschmerzen	1 (1,8)c	0	0
Appendizitis	0	1 (1,8)a	0
Dehydratation	0	1 (1,8)a	0
Femurfraktur	1 (1,8)b	0	0
Grand Mal-Anfall	1 (1,8)1)b	0	0
Influenza	1 (1,8)b	0	0
Beinfraktur	0	0	1 (1,7)a
supraventrikuläre Tachykardie	0	0	1 (1,7)b
Varizellen	1 (1,8)a	0	0

Quelle: CSR, S. 152, Tab. 46 und Tab. 14.3.2.2, S. 2556. Patienten können mehrfach gezählt werden

¹⁾ initial als möglicherweise durch Medikament verursacht eingestuft

a=kein kausaler Zusammenhang; b=Zusammenhang unwahrscheinlich; c=Zusammenhang möglich

Vom pU beschriebene UE von besonderem Interesse

Bei einem Patienten in der Ataluren-20/20/40mg-Gruppe wurde die Dosis aufgrund eines Hautausschlags (Grad 2), der 3 Monate nach Behandlungsbeginn auftrat, für 3 Wochen unterbrochen. Nach 2 Wochen verschwand der Ausschlag, nach einer weiteren Woche wurde Ataluren in der Dosierung 10/10/20 mg verabreicht, nach weiteren 2 Monaten wurde wieder die ursprüngliche Dosis gegeben. Weitere Hautausschläge traten bei diesem Patienten nicht auf.

Weitere sicherheitsrelevante Aspekte

Aufgrund von Signalen aus toxikologischen und Phase-I-Studien wurde eine Reihe von Laborwerten erfasst und ausgewertet, insbesondere zu Leberenzymen, Serum-Cholesterin und Triglyceriden.

In Studie 007 fanden sich keine Hinweise auf hämatologische oder renale Toxizität. Bei 5 Patienten (1 Placebo, 3 Ataluren 10/10/20 mg, 2 Ataluren 20/20/40 mg) fand sich jeweils einmalig eine leichte Erhöhung (Grad 1) der Bilirubinwerte, bei weiteren 5 Patienten (2 Ataluren 10/10/20 mg, 3 Ataluren 20/20/40 mg) war die Gamma-Glutamyltransferase einmalig leicht (Grad 1) erhöht (CSR, S. 161). Cholesterin- und Triglyceridwerte waren zu Baseline in allen drei Gruppen erhöht und stiegen im Behandlungszeitraum weiter an (bis etwa Woche 9), wobei eine Interaktion mit Kortikosteroidtherapie vorhanden war, aber auch ein Dosisgradient von Placebo hin zu Ataluren 20/20/40 mg, d. h. die Blutfettwerte waren in den Ataluren-Armen stärker erhöht (CSR, S. 157). Dieses Muster fand sich auch bei den (diastolischen und systolischen) Blutdruckwerten; 6 Patienten (1 Placebo, 5 Ataluren 10/10/20 mg) wiesen Blutdruckwerte oberhalb der altersspezifischen Norm auf, alle nahmen jedoch Kortikosteroide (CSR, S. 164). Vierzehn Patienten wiesen vorübergehend erhöhte ACTH-Werte auf, wobei 10 ebenfalls Kortikosteroide erhielten. Bei allen Patienten normalisierten sich die Werte im Studienverlauf wieder (CSR, S. 162). Bei 25 Patienten fanden sich vorübergehend erhöhte Kreatininwerte (4 Placebo, 13 Ataluren 10/10/20 mg, 8 Ataluren 20/20/40 mg) und Hinweise auf Blut (4 Placebo, 1 Ataluren 10/10/20 mg, 2 Ataluren 20/20/40 mg) in der Urinanalyse.

Die EMA weist noch auf die Interaktion mit potentiell nephrotoxischen Substanzen hin, insbesondere auf intravenös applizierte Aminoglykoside, so dass diese Medikamente im Rahmen der Studie nicht erlaubt waren (EPAR, 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, S. 59).

2.7 Einschätzung zum Ausmaß des Zusatznutzens durch den pU

Der pU leitet aus den Daten der Studie 007 einen beträchtlichen Zusatznutzen für die Zielpopulationen ab. Der pU begründet seine Einschätzung zusammengefasst wie folgt:

„Bei der DMD handelt es sich um eine sehr schwerwiegende, progredient verlaufende Erkrankung mit hohem Leidensdruck für die Patienten sowie deren familiäres Umfeld. Die Behandlung mit Ataluren führt verglichen mit Placebo (SOC) zu einer Abschwächung schwerwiegender (bzw. schwerer) Symptome (bzw. Folgekomplikationen), in diesem Falle eines für die Patienten klinisch bedeutsamen Effekts auf die Gehfähigkeit, der Einfluss auf die Krankheitsprogression hat. Der Therapieeffekt ist klinisch relevant, was sich auch im EPAR der CHMP bestätigt findet (9). Entsprechend § 5 Absatz 7 der VerfO ergibt sich für Ataluren 10/10/20 in der zugelassenen Indikation daher ein beträchtlicher Zusatznutzen in der Zielgrößenkategorie schwerwiegende bzw. schwere Symptome (Morbidity), d.h. eine deutliche Verbesserung des therapierelevanten Nutzens. In Tabelle 4-50 sind die Ergebnisse der patientenrelevanten Endpunkte zusammenfassend dargestellt und Tabelle 4-51 gibt einen Überblick zur Einschätzung des Zusatznutzens von Ataluren 10/10/20 in der zugelassenen Indikation.“ (Dossier S. 122)

Tabelle 4-51: Überblick zum patientenrelevanten Zusatznutzen von Ataluren bei Patienten mit nmDMD im Alter von ≥ 5 Jahren

Nutzendimension	Einstufung des Zusatznutzens
Mortalität	Keine Daten vorhanden (nicht anwendbar in einer 1-Jahres-Studie)
Morbidity 6MWD (Verzögerung des Verlusts der Gehfähigkeit bzw. Krankheitsprogression)	Beträchtlich
Gesundheitsbezogene Lebensqualität	Kein Zusatznutzen bekannt
UEs inklusive schwerwiegender UEs	Kein relevantes Schadenspotential

Quelle: Dossier, Modul 4, S. 123

3 Methodische Anmerkungen und Bewertung der Unterlagen

Eine Marktzulassung unter „besonderen Bedingungen“ (conditional approval) für Ataluren durch die EMA erfolgte 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*“. Die empfohlene tägliche Dosis ist 40 mg/kg KG aufgeteilt auf drei Einzeldosen (10, 10, 20 mg).

Die Zulassung unter “besonderen Bedingungen” ist 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)”*. Die Studienergebnisse werden für das vierte Quartal 2015 erwartet (EMA Summary of product characteristics). Bei der Studie 020 handelt es sich um eine Phase-3-Studie mit 220 Patienten, in der Ataluren 10/10/20 mg 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 festgelegt. Die Studie soll Patienten in der sogenannten decline phase erfassen, bei denen eine bessere Wirksamkeit von Ataluren erwartet wird. Die Studie wird an 55 Zentren in den USA, Kanada, Australien, Südamerika, Europa, Südostasien und Israel durchgeführt. Die Patientenrekrutierung war im September 2014 abgeschlossen (<https://www.clinicaltrials.gov/ct2/show/study/NCT01826487>).

Der (bedingten) Zulassungsentscheidung war ein negatives Votum des CHMP vom Januar 2014 vorausgegangen. Zu diesem Zeitpunkt kam die EMA bei der Bewertung der eingereichten Unterlagen zu dem Schluss: *„While the effects observed in the pivotal study were considered generally encouraging, the CHMP considered that the clinical efficacy data submitted were not adequate and did not provide sufficient evidence to support the indication of ataluren for the treatment of patients with Duchenne muscular dystrophy.“* Eine abschließende Beurteilung sei erst nach Vorliegen der Ergebnisse aus der o. g. Phase-3-Studie möglich (EPAR, S. 51). Obwohl es sich bei DMD um eine lebensbedrohliche Erkrankung handle und Ataluren einen bisher ungedeckten Bedarf adressiere sei die Datenlage zur Wirksamkeit derzeit als unzureichend einzustufen. Zudem könne eine Zulassung die Durchführung der Studie 020 gefährden (EPAR, S. 75). Der pU reichte daraufhin weitere Unterlagen ein und verlangte eine erneute Prüfung, die in einem positiven Votum des CHMP resultierte. Insbesondere wurde die zuvor angezweifelte Argumentation des pU, dass Ataluren eine konzentrationsabhängige (glockenkurvenförmige) Wirksamkeit besitzt, die bei hohen Dosen wieder abnimmt (in der Hochdosisgruppe hatte sich kein Effekt gegenüber Placebo gezeigt), nunmehr als plausibel angesehen. Das CHMP revidierte daher seine initiale Auffassung, dass der positive Effekt in der Ataluren-10/10/20mg-Gruppe ein Zufallseffekt gewesen sein könnte. Allerdings sah das CHMP seine Zweifel an den Wirksamkeitsdaten nicht ausgeräumt, stellte aber ein positives Verhältnis von Wirksamkeit und Sicherheit vor allem aufgrund des Sicherheitsprofils von Ataluren fest, das als unbedenklich eingestuft wurde (EPAR, S. 81). Beide Entscheidungen des CHMP waren umstritten und es sind entsprechende Minderheitenvoten im EPAR dokumentiert.

Die folgenden Ausführungen beziehen sich auf die pivotale Studie 007, deren Ergebnisse der pU im Dossier als Grundlage für seine Ausführungen zum Ausmaß des Zusatznutzens heranzieht.

3.1 Design und Methodik

Bei der Studie 007 handelt es sich um eine multizentrische, randomisierte, placebokontrollierte, verblindete Phase-2b-Dosisfindungsstudie im Parallelgruppendesign (1:1:1) mit drei Behandlungsarmen (Placebo, Ataluren 10/10/20 mg/kg KG, Ataluren 20/20/40 mg/kg KG). Das primäre Ziel der Studie war der Nachweis der Verbesserung der Gehfähigkeit im Vergleich zu Placebo anhand der 6-Minuten-Gehstrecke (6MWT).

Bestandteil der Studie war eine 6-wöchige Screening-Phase, eine 48-wöchige verblindete Behandlungsphase und eine 6-wöchige Nachbeobachtungsphase. Die Patienten wurden stratifiziert für die Faktoren Alter (< 9 / ≥ 9 Jahre), Einnahme von Kortikosteroiden (ja / nein) und 6MWT-Wert zu Baseline (< 350 / ≥ 350 m). Nach Abschluss der 48-wöchigen randomisierten Studienphase war eine offene Extensionsphase mit der höheren Ataluren-Dosis vorgesehen.

Eingeschlossen wurden gehfähige Patienten (≥ 75 m ohne Hilfe) im Alter ab 5 Jahren mit bestätigter Diagnose einer DMD. Die Studie wurde in 37 Zentren in 11 Ländern durchgeführt. Die Randomisierung erfolgte durch ein IVR-/IWR-System. Entsprechend der vorgesehenen Stratifizierung wurde eine adaptive Randomisierung vorgenommen, um balancierte Gruppen zu gewährleisten. Die Zuteilung erfolgte nach Eingabe der Patientencharakteristika nach einem zuvor festgelegten Algorithmus. Die Prüfarzte hatten während des Rekrutierungszeitraums keinen Zugriff auf die Randomisierungsliste. Der erste Patient wurde am 28.02.2008 eingeschlossen, die letzte Studiervisite fand am 17.12.2009 statt. Aufgrund des Designs und der Methodik der Studie 007 wird von einem niedrigen Verzerrungspotential auf Studienebene ausgegangen.

Post hoc wurden allerdings mehrere relevante Änderungen vorgenommen. Zum einen wurde bei zwei Patienten der Baseline-Wert für den 6MWT durch den Wert aus dem Screening ersetzt. Zum anderen wurde das statistische Auswertungsmodell für den primären Endpunkt modifiziert. Drittens wurde nachträglich eine nicht präspezifizierte Subgruppe aus den Patienten gebildet, die sich in der sogenannten decline phase befinden. Der pU leitet den Zusatznutzen von Ataluren aus der post-hoc-Analyse der Studiendaten ab.

Ersetzung von Baseline- durch Screening-Werte bei zwei Patienten für den 6MWT

Bei je einem Patienten in der Placebo- bzw. in der Ataluren-20/20/40mg-Gruppe wurde der Baseline-Wert für den 6MWT durch den Screening-Wert ersetzt. Grund hierfür waren akute Verletzungen (Knie bzw. Sprunggelenk) wenige Tage vor der Erhebung der Baseline-Werte der unteren Extremität, die zu einem falsch-niedrigen, d. h. nicht repräsentativen Ergebnis zu Baseline führten (309 statt 395 m bei der Screening-Visite bzw. 125 statt 303 m bei der Screening-Visite). Die Werte wurden zunächst als gültig eingestuft und nachträglich durch die

Screening-Werte ersetzt, was zu einer korrigierten ITT-Population (kITT) führte. Der pU legt plausibel dar, dass die tatsächliche Performance der beiden Patienten durch die Verletzung vorübergehend beeinträchtigt war, so dass diesem Vorgehen gefolgt werden kann (vgl. CSR, S. 96).

Für die Bewertung der Größenordnung des Effekts ist nur der ersetzte Wert in der Placebo-Gruppe relevant, da der zweite Wert bei einem Patienten in der hier nicht relevanten Hochdosisgruppe ersetzt wurde. Der Mittelwert für die Gehstrecke zu Baseline änderte sich in der Placebo-Gruppe dadurch von 359,6 m (SD 87,7) auf 361,1 m (87,5). Die absolute Veränderung der Gehstrecke betrug nach Korrektur in der Placebo-Gruppe zu Woche 48 (Mittelwert, SD) -44,14 (88) im Vergleich zu -42,56 (90) Meter in der ursprünglichen ITT-Population (vgl. CSR, S. 91). Die entsprechenden Werte unter Ausschluss des Patienten in der Placebo-Gruppe waren 360,51 m (88,2) zu Baseline und -45,45 m (88,3) Differenz nach 48 Wochen. Die Differenz zwischen dem Placebo- und dem Ataluren-10/10/20mg-Arm betrug in der kITT-Population 31,3 m (CSR, S. 99) im Vergleich zu 29,7 m in der ITT-Population (CSR, S. 91).

Änderung des statistischen Auswertungsmodells

Weitere Änderungen betreffen das statistische Auswertungsmodell. Im SAP war ursprünglich eine non-parametrische mixed-model repeated-measure analysis of variance (ANOVA) vorgesehen (SAP Version 1.0, 03.12.2007). In der Version 3.0 (04.03.2009) wurde das Modell in ein mixed-model, repeated-measures analysis of covariance (ANCOVA) geändert. Diese Änderung wurde damit begründet, dass angenommen wurde, die Größenordnung der Änderung im 6MWT werde mit dem Baseline-Wert korrelieren. Außerdem wurde davon ausgegangen, dass die Power der ANCOVA größer sei (Studienprotokoll, S. 141). Diese Änderung erfolgte vor der Entblindung der Daten bzw. Datenbankschluss.

Die Analyse der Daten wurde zunächst mit diesem Modell an der ITT-Population durchgeführt und berichtet. Im Rahmen der im SAP vorgesehenen Auswertung zeigten sich Inkonsistenzen in den Ergebnissen:

- Die Standardabweichung in den Mittelwerten nach 48 Wochen war erheblich höher als die erste (geplante) Zwischenauswertung vermuten ließ (SD 72 bis 90 m) bzw. in der Fallzahlkalkulation angenommen wurde, so dass die Detektion eines signifikanten Effekts erschwert war (vgl. CSR, S. 92).
- Im Modul 4 des Herstellerdossiers (S. 51) wird zudem festgehalten: *„Der Vergleich von Ataluren 10/10/20 und Placebo zeigte einen substantiellen Unterschied zwischen der prä-spezifizierten MMRM-Analyse der nicht transformierten 6MWD ($p=0.0905$) verglichen mit dem p -Wert der prä-spezifizierten ANCOVA-Analyse ($p=0.0445$). Da nur wenige Werte in Woche 48 fehlten, deutete diese Diskrepanz an, dass das protokollgemäße MMRM-Modell möglicherweise ungeeignet sein könnte.“*

Nach Konsultation mit einem Biostatistiker (Prof. Gary Koch, North Carolina) wurde ein Baseline-Visite-Interaktionsterm im Modell ergänzt, um eine Adjustierung für jede Visite zu ermöglichen. Außerdem wurde die primäre Analyse statt mit rangtransformierten Daten mit den untransformierten Daten gerechnet und statt der Rangtransformation wurde ein

Permutationstest durchgeführt, um die Abweichung von der Vermutung der Normalverteilung der Residuen zu adressieren (CSR, S. 72, 100).

Insgesamt wurden also drei Änderungen im statistischen Auswertungsmodell des primären Endpunkts vorgenommen, wovon zwei post hoc nach Entblindung und initialer Datenanalyse stattfanden. Zudem wurde die post-hoc-Auswertung an der korrigierten ITT-Population durchgeführt. Während die Änderung auf Patientenebene, die zur korrigierten ITT-Population führte, als plausibel eingestuft werden kann, sind die Änderungen im statistischen Auswertungsmodell nicht eindeutig zu bewerten. Beispielsweise ist unklar, warum kein Interaktionsterm für Alter oder andere Parameter in das Modell aufgenommen wurde. Die für die Studienergebnisse relevante Konsequenz liegt darin, dass mit dem modifizierten Modell nun ein signifikanter Unterschied zwischen Placebo und Ataluren 10/10/20 mg resultiert ($p=0,0367$ bzw. $p=0,0561$ nach Permutationstest) bei einer absoluten Differenz der Gehstrecke von 31,7 m (95%-KI 5,1;58,3). Die post hoc gewonnenen Ergebnisse werden daher als potentiell hoch verzerrt eingestuft, da nicht ausgeschlossen werden kann, dass sie ergebnisgesteuert sind.

Post-hoc-Definition einer ‚ambulatory decline phase‘-Subgruppe

Neben der nachträglichen Korrektur der ITT-Population und der Änderung des statistischen Auswertungsmodells wurde post hoc auch eine neue Subgruppe gebildet, nämlich Patienten in der sogenannten ambulatory decline phase. Die Subgruppe wurde aufgrund einer separaten Auswertung des Verlaufs im Placebo-Arm der Studie 007 identifiziert (NB: diese Auswertung ist nicht im CSR enthalten). Dabei handelt es sich um Patienten, die über den Studienzeitraum von 48 Wochen einen Verlust an Gehstrecke im 6MWT erlitten hatten, sie waren älter als 7 Jahre, wiesen eine initiale Gehstrecke von < 350 m auf und erhielten Kortikosteroide. Der pU leitete aus den Daten ab, dass Patienten im Alter von 7-16 Jahren, einer initialen Gehstrecke von mehr als 150 m und < 80 % predicted 6MWD eher von Ataluren profitieren würden (vgl. Einschlusskriterien für die noch laufende Phase-3-Studie). In diese Subgruppe wurden 32 Patienten aus der Ataluren-10/10/20mg- und 31 Patienten aus der Placebo-Gruppe klassifiziert. Für den 6MWT ergab sich für diese Subgruppe eine Gehstreckendifferenz von 49,9 m (nominaler p-Wert 0,0096) (vgl. EPAR, S. 42). Auch hier ist von einem hohen Verzerrungspotential auszugehen.

3.2 Endpunkte zur Wirksamkeit

Morbidität

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

Der 6MWT wurde standardisiert nach der Methode der Universität von Kalifornien erhoben, wobei es sich um eine gegenüber der Version der American Thoracic Society modifizierte bzw. adaptierte Version handelt (McDonald et al. 2010), die u. a. folgende Änderungen enthält: 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*“).

Der pU begründet die Wahl des Endpunkts damit, dass Änderungen im 6MWT als direktes Maß zur Einschätzung der Gehfunktion bzw. als sensitives Maß für die Krankheitsprogression von DMD-Patienten angesehen werden. Der Verlust der Gehfähigkeit wird als ernsthafteste Komplikation im Krankheitsverlauf von DMD angesehen. Der Erhalt der Gehfähigkeit ist daher ein zentrales Ziel der Therapie. Die *National Institutes of Health* stufen den 6MWT als „globales Instrument der Gehfähigkeit und Ausdauer“ ein. Der Erhalt der Muskelfunktion wurde in einer Umfrage für DMD-Patienten als wichtiger eingeschätzt als die Verlängerung der Lebenserwartung (Modul 4, S. 78).

Für Patienten ist die Gehfähigkeit als relevanter Endpunkt einzustufen und es stellt sich die Frage, inwieweit der 6MWT als valides Surrogat für den Erhalt der Gehfähigkeit angesehen werden kann bzw. inwieweit eine Differenz in der Gehstrecke mit der Zeit bis zum Verlust der Gehfähigkeit korreliert. In mehreren Studien wurde gezeigt, dass DMD-Patienten eine durchschnittlich geringere Gehstrecke im 6MWT aufweisen als gesunde Gleichaltrige (bspw. in einer Studie 268 ± 92 m vs. 468 ± 42 m; unpublizierte Daten von 52 Kindern, zitiert nach Studienprotokoll Version 3.0, S. 87). In Modul 4 des Dossiers verweist der pU auf eine italienische multizentrische Kohortenstudie mit einer Laufzeit von 24 Monaten, mit dem Ziel, das Spektrum von krankheitsbedingten Veränderungen bei DMD-Patienten mit mind. 75 m Gehstrecke mit verschiedenen Erhebungsinstrumenten zu erfassen. Insgesamt wurden 113 Kinder (Durchschnittsalter 8,2 Jahre, Spanne 4,1-17) eingeschlossen. Der Gehstreckenverlust betrug im Durchschnitt 87,3 m über 2 Jahre, bei beträchtlicher Variabilität (SD 123,1 m), besonders ausgeprägt bei Kindern über 7 Jahren (vs. < 7). Demnach liegt das Risiko, innerhalb von 2 Jahren die Gehfähigkeit zu verlieren, bei Kindern mit einer Gehstrecke im 6MWT von < 330 m bei 30 % (Mazzone et al. 2013). Außerdem verweist der pU auf Studien von Henricson et al. (2013) und Pane et al. (2014), um die prognostische Relevanz der Gehstrecke über/unter 350 m im 6MWT aufzuzeigen. Die Operationalisierung und Stratifizierung des Endpunkts sind insgesamt nachvollziehbar.

Vorerfahrungen bzgl. einer MCID für den 6MWT bei DMD-Patienten existieren bisher nicht. Hinsichtlich der MCID wird auf Untersuchungen von McDonald et al. (2010) verwiesen, die zeigen, dass eine Gehstreckenverkürzung um 30 m das Risiko, die Gehfähigkeit innerhalb von 2 Jahren zu verlieren, erheblich erhöht. Erfahrungen bei Patienten mit pulmonaler Hypertonie und Mukopolysaccharidose legen nahe, dass eine Verbesserung um 10-15 % als klinisch relevant angesehen wird. Bezogen auf eine durchschnittliche Gehstrecke von ca. 270 m ergab sich damit eine anzustrebende absolute Verbesserung um 30 m im Vergleich zu Placebo. Die Patientenrelevanz des 6MWT als Surrogat für „Meilensteine“ des Funktionsverlusts ist entsprechend der Expertenempfehlungen nachvollziehbar (Bushby 2011), allerdings ist die Begründung für die gewählte Relevanzschwelle von 30 m aufgrund der methodischen Einschränkungen der Assoziationsstudie mit dem PODCI-Instrument derzeit nur eingeschränkt als valide zu betrachten, insbesondere weil der Zusammenhang bisher lediglich indirekt untersucht wurde. Zudem wurde das PODCI-Instrument in der Studie 007 nicht eingesetzt.

Der Nachweis der Überlegenheit für den primären Endpunkt wurde nicht erreicht. Der geschätzte Behandlungsunterschied zwischen dem Ataluren-10/10/20mg- und dem Placebo-

Arm von 26,4 m zugunsten von Ataluren 10/10/20 mg war in der primären (präspezifizierten Analyse) statistisch nicht signifikant. Erst in der bereits diskutierten post-hoc-Analyse wurde ein statistisch signifikanter und klinisch als relevant eingeschätzter Effekt gezeigt.

Für den primären Endpunkt wurden verschiedene Sensitivitätstests durchgeführt (u. a. ANCOVA bzw. MMRM mit LOCF und jeweils nicht transformierten Daten, verschiedene Imputationsmethoden, Reallokation der Geschwisterpaare) und die Analyse wurde mit der Per-Protokoll-Population wiederholt. Diese Analysen bestätigten das Ergebnis, allerdings zeigte sich in der ANCOVA mit LOCF an nicht transformierten Daten ein signifikanter p-Wert (nominal 0,0445, adjustiert 0,0808), was dazu Anlass gab, wie bereits oben diskutiert, *post hoc* das Auswertungsmodell zu modifizieren. Die Sensitivitätsanalyse unter Ausschluss der beiden Patienten mit ungültigen Baseline-Werten für den 6MWT ergab ähnliche Ergebnisse wie in der Analyse der kITT-Population (Gehstreckendifferenz nach 48 Wochen in der Placebo-Gruppe -45,45 m (88,279), in der Ataluren-10/10/20mg-Gruppe unverändert). Ergebnisse für den statistischen Vergleich werden allerdings nur für das *post hoc* modifizierte Modell berichtet (Gehstreckendifferenz 32,7 m, $p=0,0255$ (nominal), $p=0,0504$ (Dunnett-Adjustierung)) (CSR, S. 98, 1027).

Weitere supportive Analysen bezogen sich auf den Anteil der Progressoren, der nach 48 Wochen im Ataluren-10/10/20mg-Arm geringer war als im Placebo-Arm (26 vs. 44 %), eine relative Reduktion von 48 %. Das obere Konfidenzintervall für diese supportive Analyse betrug 0,966, der nominale p-Wert 0,0386, der adjustierte p-Wert 0,078. Der Anteil der Patienten mit einer mindestens 10%igen Verbesserung nach 48 Wochen (Responder) unterschied sich nicht signifikant zwischen den Gruppen.

Subgruppenanalysen des primären Endpunkts wurden für die folgenden Variablen geplant (vgl. CSR, S. 121): Alter (< 9 / ≥ 9 Jahre), 6MWT (≥ 350 / < 350 m), Kortikosteroide zu Baseline (ja / nein), Altersverteilung (5-6, 7-8, 9-11, ≥ 12 Jahre), Lokalisation der Nonsense-Mutation (1-39, 40-79), Art des Stopp-Codons (UGA, UAG, UAA), Die Subgruppenanalysen wurden an der korrigierten ITT-Population durchgeführt und mittels des *post hoc* modifizierten statistischen Auswertungsmodells gerechnet. Die Ergebnisse zeigten eine bessere Wirksamkeit im Ataluren-10/10/20mg-Arm im Vergleich zu Placebo unabhängig von der Subgruppe. Die Wirksamkeit war am stärksten ausgeprägt in der Subgruppe der < 9 -Jährigen, bei Patienten mit Kortikosteroidtherapie und bei Patienten mit einem 6MWT zu Baseline < 350 m. Da es sich um Ergebnisse mit dem post-hoc-Analysemodell handelt, sind diese ebenfalls als potentiell verzerrt einzustufen. Dies gilt auch für die Auswertung der *decline-phase*-Subgruppe, zumal auch keine Daten zur Vergleichbarkeit der *post hoc* gebildeten Gruppen vorliegen.

Sekundäre Endpunkte

Zu den sekundären Endpunkten gehören die sogenannten Timed Function Tests (TFT), die die Messung der Muskelfunktion ermöglichen, die Erfassung von Änderungen in den Alltagsaktivitäten mittels Schrittzählern und Patiententagebuch sowie patientenberichtete Stürze und Rollstuhlnutzung (Patiententagebuch). Endpunkte zur Messung der Muskelkraft wurden nicht berücksichtigt, da die Muskelfunktion bei DMD-Patienten im Vordergrund steht

und das Verhältnis von Muskelkraft und Muskelfunktion offenbar nicht linear ist; Änderungen der Muskelkraft sind daher schwer zu interpretieren. Geringe Änderungen in der Dystrophin-Produktion in den Muskelzellen können die Funktion offenbar beeinflussen, ohne dass dies als Änderung der Muskelkraft messbar ist. Die Muskelkraft wird deshalb nicht als direkt patientenrelevant eingestuft. Auch die EMA kommt zu dieser Einschätzung (vgl. EPAR, S. 49f.).

Änderung in der proximalen Muskelfunktion (,Timed Function Tests') zu Woche 48

Die Messung der proximalen Muskelfunktion in TFTs, wurde in vier Testsituationen aufgeteilt, d. h. die Patienten sollten so schnell wie möglich (innerhalb von max. 30 Sekunden) aus der Rückenlage aufstehen, 10 m laufen/gehen und vier Treppenstufen hinauf bzw. hinuntersteigen. Neben dem Zeitbedarf wurde auch ein Score aus jeweils 6 Kategorien aus dem Hilfe-/Unterstützungsbedarf für die Bewältigung der Aufgaben gebildet.

TFTs werden als supportiv zum primären Endpunkt erfasst. Die bereits erwähnte DMD-Expertengruppe weist darauf hin, dass der Erhalt von Fähigkeiten wie Treppensteigen, selbständiges Gehen, Essen usw. als relevant für die Krankheitsprogression angesehen werden kann. Die in der Studie 007 eingesetzten Tests werden in Studien bei DMD-Patienten üblicherweise verwendet. Sie sind reproduzierbar und konnten den Effekt von Steroiden auf die DMD erfassen. Der 10-Meter-Test ist breit akzeptiert und wird häufig eingesetzt, um die physische Spitzenaktivität zu messen. Bei einer Zeit von mehr als 12 Sekunden (in einer anderen Studie wurden 7 Sekunden ermittelt) ist das Risiko deutlich erhöht, innerhalb von 12 Monaten die Gehfähigkeit zu verlieren. Patienten, die weniger als 6 Sekunden benötigen, haben ein signifikant niedrigeres Risiko die Gehfähigkeit zu verlieren. Auch die Fähigkeit Treppen zu gehen hat sich als prädiktiv für den Verlust der Gehfähigkeit erwiesen (in einer Studie verloren 40 % der Patienten, die länger als 8 Sekunden benötigten, innerhalb von 12 Monaten ihre Gehfähigkeit). Eine MCID wurde mit 0,19 Meter/Sekunde ermittelt (Henricson et al. 2013).

Die Ergebnisse zeigten keine statistisch signifikanten Effekte zugunsten von Ataluren 10/10/20 mg, auch nicht in der post-hoc-Auswertung mit der KITT-Population. Im 10-Meter-Test (relative Differenz -43 %), beim Treppenhinaufsteigen (-50 %) und beim Treppenhinabsteigen (-40 %) zeigten sich Trends zugunsten von Ataluren 10/10/20 mg. Beim Aufstehen aus Rückenlage konnte kein Trend festgestellt werden. Der pU weist darauf hin, dass zu Baseline bereits ein Viertel der Patienten dies nicht mehr konnte, so dass hier möglicherweise der Nachweis eines Effekts erschwert war (bei den anderen TFTs waren etwa 3 % der Patienten nicht in der Lage, die Übungen auszuführen).

Änderung der Alltagsaktivitäten gemessen mittels ,Step Activity Monitoring' (Pedometer am Knöchel) an ≥ 9 Tagen nach jeder Studienvisite

Die Änderung der Alltagsaktivitäten wurde mittels ,Step Activity Monitoring' (SAM) durch ein Pedometer (Schrittzähler) in Kombination mit einem Patiententagebuch (SAM Diary) für jeweils vier bis sieben Tage nach jeder Studienvisite erfasst. Der Aktivitätslevel wurde in 4 Kategorien eingeteilt, von ,keine Aktivität' (0 Schritte) bis ,ausgeprägte Aktivität' (≥ 30 Schritte/Minute). Der pU wertet die Erfassung der Alltagsaktivitäten als supportiv zum primären Endpunkt. In Modul 4 wird der Endpunkt jedoch nicht als patientenrelevant eingestuft

und daher nicht zur Herleitung des Zusatznutzens aufgeführt. Von der EMA wird die Erfassung der Aktivitäten des täglichen Lebens empfohlen (vgl. EMA 2013). Da die Partizipation im Alltag für die Betroffenen von großer Bedeutung ist, wird dieser Endpunkt als relevant für die Nutzenbewertung eingestuft.

Für diesen Endpunkt zeigten sich keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen und auch keine konsistenten Trends. Der Anteil mit Patienten ohne Aktivität lag in beiden Gruppen bei etwa 50 % zu Beginn der Studie und stieg im Studienverlauf um ca. 3-4 % an. Eine geringe Aktivität wies rund ein Drittel der Patienten auf, mäßige Aktivität rund 11 % und eine ausgeprägte Aktivität 6-7 % der Patienten. In den Kategorien mit geringer und ausgeprägter Aktivität sank der Anteil nach 48 Wochen um jeweils ca. 1 %, in der Kategorie mäßige Aktivität sank der Anteil in der Placebo-Gruppe stärker als in der Ataluren-10/10/20mg-Gruppe (-1,92 vs. -0,69 %), der Unterschied war jedoch nicht statistisch signifikant, die klinische Relevanz einer Änderung in dieser Größenordnung ist zudem unklar. Post-hoc-Auswertungen wurden für diesen Endpunkt nicht berichtet.

Patientenberichtete Änderung der Häufigkeit von Stürzen und der Rollstuhlnutzung (Patiententagebuch)

Für DMD-Patienten relevante Ereignisse, wie ungewollte Stürze und die Häufigkeit der Nutzung von Rollstühlen, wurden mit Hilfe eines täglichen Patiententagebuchs erfasst, in dem auch die Einnahme von Medikamenten dokumentiert wurde. Frühere Studien haben gezeigt, dass Stürze mittels Patiententagebüchern zuverlässig erfasst werden können. Studien zum Effekt von Glukokortikoiden bei DMD-Patienten haben zudem gezeigt, dass die Sturzfrequenz mittels medikamentöser Therapie reduziert werden kann. Die Reduktion von Stürzen und damit assoziierten Verletzungen (Frakturen) und Schmerzen wird daher als patientenrelevanter Endpunkt angesehen, insbesondere bei Patienten mit glukokortikoidinduzierter Osteoporose. Es ist bei diesem Endpunkt nicht eindeutig aus den Unterlagen des pU ersichtlich, wie die Rollstuhlnutzung erfasst wurde. Aus dem Studienprotokoll und dem CSR kann gefolgert werden, dass die Rollstuhlnutzung nicht aus den Angaben des täglichen Patiententagebuchs, sondern aus dem *SAM Diary* ermittelt wurde. Das *SAM Diary* wurde aber nur an den Tagen unmittelbar nach jeder Studienvisite erfasst, so dass unklar ist, ob die Angaben für die gesamte Studiendauer repräsentativ sind. Das Verzerrungspotential für den Endpunkt Rollstuhlnutzung wird daher als unklar eingeschätzt. Post-hoc-Analysen für die beiden Endpunkte wurden nicht berichtet.

Die Sturzfrequenz (im Mittel pro Tag) war zu Baseline in der Placebo-Gruppe etwa zweimal so hoch wie in der Ataluren-10/10/20mg-Gruppe (0,54 vs. 0,27). Nach 48 Wochen lag sie bei 0,72 in der Placebo-Gruppe und bei 0,23 in der Ataluren-Gruppe. Das Relative Risiko (*relative ratio*) betrug 0,384 (95%-KI 0,16;0,94, $p=0,0356$) zugunsten der Ataluren-Gruppe und zeigte damit einen Vorteil für die mit Ataluren 10/10/20 mg behandelte Gruppe. [Für diesen Endpunkt zeigte sich auch ein signifikanter Effekt zugunsten der Hochdosisgruppe.] Aufgrund des hohen Anteils fehlender Werte zu Baseline (21 % je Gruppe) wurde eine Sensitivitätsanalyse durchgeführt, bei der die Baseline- durch die 6-Wochen-Werte ergänzt wurden. Dadurch wurde die Anzahl der fehlenden Werte um $N=10$ reduziert. Dies resultierte in einem ebenfalls

statistisch signifikanten Relativen Risiko von 0,4947 (95%-KI 0,27;0,91, $p=0,0229$) zugunsten der Ataluren-Gruppe.

Die Häufigkeit der Rollstuhlnutzung war zu Baseline in den Behandlungsgruppen mit etwa 13 % vergleichbar, nach 48 Wochen nahm die Nutzung in der Placebo-Gruppe auf 25 %, in der Ataluren-Gruppe auf 19 % zu, eine Differenz von 7,5 % zugunsten von Ataluren 10/10/20 mg. Der Unterschied war statistisch nicht signifikant.

3.2.1 Zusammenfassende Einschätzung zur Wirksamkeit

Das Verzerrungspotential für die Erhebung der Endpunkte ist (bezogen auf die präspezifizierten Auswertungen) insgesamt als niedrig einzuschätzen, mit Ausnahme des Endpunkts Rollstuhlnutzung (Verzerrungspotential unklar).

Der MCID für den 6MWT von 30 m kann grundsätzlich gefolgt werden. Für den primären Endpunkt ergibt sich in der präspezifizierten ITT-Auswertung kein statistisch signifikanter Effekt. Für die supportive Auswertung zum primären Endpunkt, den Anteil der Progressoren (= Anteil der Patienten mit einer mindestens 10%igen Verschlechterung nach 48 Wochen), zeigte sich ein statistisch signifikanter Vorteil. In der post-hoc-Analyse wurde ein statistisch signifikanter und klinisch relevanter Effekt für den 6MWT gefunden.

Bei den sekundären Endpunkten zeigte sich lediglich für die patientenberichtete Sturzhäufigkeit ein Vorteil zugunsten von Ataluren 10/10/20 mg. Die EMA stufte die Ergebnisse der sekundären Endpunkte insgesamt als nicht supportiv für den primären Endpunkt ein (vgl. EPAR, S. 49).

Letztlich ist nicht auszuschließen, dass die Power der Studie nicht ausreichend war, um einen statistisch signifikanten Effekt in der präspezifizierten Analyse nachzuweisen. Die Varianz der Gehstrecke wurde im Studienverlauf deutlich größer als im SAP angenommen (statt 50 m waren es nach 48 Wochen 90 m in der Placebo- und 72 m in der Ataluren-10/10/20mg-Gruppe), wodurch die Power der Analyse beeinträchtigt wurde. Diese Entwicklung war zum Zeitpunkt der geplanten ersten Interimsanalyse noch nicht abzusehen. Die geplante zweite Interimsanalyse (nach 24 Wochen bzw. ≥ 90 Patienten) wurde nicht durchgeführt, da zu diesem Zeitpunkt der pU und die FDA den SAP diskutierten und auch keine besorgniserregenden Sicherheitssignale vorhanden waren (vgl. CSR, S. 121).

3.3 Lebensqualität

Für die Erfassung der Lebensqualität wurde das *Generic Core Scales*-Paket des PedsQL Inventory verwendet mit den Domänen physikalische, emotionale, soziale und schulische Funktion. Es handelt sich um ein ausreichend validiertes generisches Instrument zur Messung der Lebensqualität bei Kindern mit chronischen Erkrankungen. Ergebnisse anderer Studien legen nahe, dass Lebensqualität bei Patienten mit DMD ein aussagekräftiger Endpunkt sein kann. Ergänzend zum PedsQL wurde das PedsQL-Fatigue-Instrument eingesetzt, das zum Zeitpunkt der Protokollerstellung allerdings nur in Englisch vorlag, aber parallel zur Studien-

durchführung in die erforderlichen Sprachen übersetzt und jeweils validiert werden sollte. Für dieses Modul hat der pU keine Unterlagen vorgelegt, so dass keine abschließende Einschätzung zur Validität und Patientenrelevanz vorgenommen werden kann.

Für keine der ausgewerteten Domänen fand sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen, auch nicht für die Fatigue-Scale. Allerdings zeigte sich ein Trend zugunsten von Ataluren 10/10/20 mg in der physikalischen und in der schulischen Domäne, in der Domäne physikalische Funktion allerdings unterhalb der vom pU angegebenen MCID von 4,5 Punkten. In den anderen Domänen waren die Ergebnisse heterogen.

In jüngster Zeit wurde auch ein spezifisches Modul des PedsQL für Patienten mit DMD entwickelt, aber erst nach Abschluss der Studie 007 publiziert. Die erfassten Dimensionen und die Fragen weichen erheblich von der in der Studie verwendeten generischen Version ab, so dass unklar ist, ob die in der Studie verwendete generische Version ausreichend sensitiv ist, um Änderungen der Lebensqualität bei Patienten mit DMD zu erfassen.

3.4 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 WHODRUG Dictionary kodiert. Besonderes Augenmerk wurde auf die Überprüfung der Transaminasen, sowie auf adrenale und renale Ereignisse gelegt. Für die Einschätzung der Sicherheit wurden Daten aller drei Behandlungsarme dargestellt. Das Verzerrungspotential des patientenrelevanten Endpunkts UE wird als niedrig eingeschätzt.

Das Nebenwirkungsprofil war zwischen den drei Behandlungsarmen vergleichbar. Therapieabbrüche aufgrund von UE traten nicht auf, es gab keine Todesfälle. Die häufigsten UE waren Erbrechen (bei insgesamt 46,6 % der Patienten über alle Behandlungsarme), Kopfschmerzen (29,3 %), Diarrhö (24,1 %), Nasopharyngitis (20,7 %), Fieber (19 %), Husten (19 %) und Oberbauchschmerzen (17,8 %) ohne signifikante Unterschiede zwischen den Gruppen. Hinsichtlich Dosisunterbrechungen oder Dosisreduzierungen zeigten sich keine auffälligen Gruppenunterschiede. Die Subgruppenanalyse ergab, dass Patienten, die Kortikosteroide einnehmen, häufiger Kopfschmerzen, Rückenschmerzen, Influenza, Extremitätenschmerzen und Hypertonie hatten, dagegen seltener Fieber, Krankheitsprogression, Rhinitis, virale Gastroenteritis, Rhinorrhö, Lymphadenopathie, Sinusitis, virale Infektionen, Hautverletzungen sowie saisonale Allergien. Relevante Unterschiede zwischen den Behandlungsgruppen waren nicht zu erkennen. Das Nebenwirkungsprofil zeigte keine erkennbare Abhängigkeit vom Alter oder vom 6MWT zu Baseline.

UE, die nach Einschätzung der Prüfarzte mit der Studienmedikation in Zusammenhang standen, waren in den drei Behandlungsarmen etwa gleich häufig, allerdings traten Erbrechen, Flatulenz, Magenbeschwerden und Fatigue häufiger in den Ataluren-Ggruppen als in der Placebo-Gruppe auf, wobei es einen Dosisgradienten gab hin zur Hochdosisgruppe (vgl. Tabelle 16). Schwere UE traten bei insgesamt 9 Patienten auf, bei 5 Patienten in der Placebo-

Gruppe und bei jeweils 2 Patienten in den Ataluren-Gruppen (Tabelle 17). Ein SUE (Grand-Mal-Anfall in der Placebo-Gruppe) wurde initial als möglicherweise mit der Studienmedikation assoziiert eingestuft, ein weiteres SUE (1 Patient mit Bauchschmerzen in der Placebo-Gruppe) wurde als möglicherweise durch das Medikament verursacht klassifiziert. Bei einem Patienten in der Hochdosisgruppe trat vorübergehend ein Hautausschlag auf, der nach Therapieunterbrechung verschwand.

UE von besonderem Interesse waren Leberenzyme, Serum-Cholesterin und Triglyceride. Cholesterin- und Triglyceridwerte waren zu Baseline in allen drei Gruppen erhöht und stiegen im Behandlungszeitraum weiter an, wobei eine Interaktion mit Kortikosteroidtherapie vorhanden war, aber auch ein Dosisgradient von Placebo hin zu Ataluren 20/20/40 mg, d. h. die Blutfettwerte waren in den Ataluren-Armen stärker erhöht. Dieses Muster fand sich auch bei den (diastolischen und systolischen) Blutdruckwerten. Weitere Auffälligkeiten in den Laboranalysen waren vorübergehend und klinisch nicht bedrohlich.

Die EMA weist auf die Interaktion mit potentiell nephrotoxischen Substanzen hin, insbesondere auf intravenös applizierte Aminoglykoside, so dass diese Medikamente im Rahmen der Studie nicht erlaubt waren. Zudem liegen keine Studien an Patienten mit renaler oder hepatischer Funktionseinschränkung vor, so dass die Sicherheit bei diesen Patientenpopulationen nicht untersucht ist. Da Ataluren über diese Organe ausgeschieden wird, kann es zur Akkumulation von Wirkstoff und Abbauprodukten kommen, wenn entsprechende Funktionseinschränkungen vorliegen (EPAR, S. 58f).

3.4.1 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 Behandlung mit Ataluren sollte nur durch Fachärzte initiiert werden, die Erfahrung mit der Behandlung der Duchenne- / Becker-Muskeldystrophie haben.

Dieses Arzneimittel wurde unter „Besonderen Bedingungen“ zugelassen. Das bedeutet, dass weitere Nachweise für den Nutzen des Arzneimittels erwartet werden. Die EMA wird neue Informationen zu diesem Arzneimittel mindestens jährlich bewerten und, falls erforderlich, die Zusammenfassung der Merkmale des Arzneimittels aktualisieren.

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 von 40 mg pro kg Körpergewicht pro Tag, aufgeteilt auf drei Dosen (10/10/20 mg/kgKG) im Abstand von 6, 6 und 12 Stunden.

Die Nutzenbewertung von Ataluren basiert auf der zulassungsbegründenden Studie PTC124-GD-007-DMD. Bei dieser Studie handelt es sich um eine multizentrische, randomisierte placebokontrollierte, verblindete Dosisfindungsstudie der Phase 2b im Parallelgruppendesign mit drei Behandlungsarmen. In die Studie wurden Patienten ab einem Alter von 5 Jahren mit DMD und einer Gehstrecke von ≥ 75 m ohne Hilfe 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 007 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 18: Zusammenfassende Darstellung der Studienergebnisse (ITT-Population, präspezifizierte Auswertung)

Endpunkt	Ataluren 10/10/20 mg vs. Placebo	Effekt
Mortalität		
Gesamtmortalität	nicht erhoben	
Morbidität		
Primärer Endpunkt¹⁾		
Veränderung im 6MWT im Vergleich zu Baseline nach 48 Wochen (m) ²⁾	26,44 (95%-KI: -4,21;57,09) p=0,0905 (nominal) p=0,1592 (Dunnett)	\leftrightarrow
<i>supportiv:</i>		
Anteil der Patienten mit einer mindestens 10%igen Verschlechterung nach 48 Wochen (Progressoren) (%)	26,3 vs. 43,9 p=0,0423	$\uparrow\uparrow$
Anteil der Patienten mit einer mindestens 10%igen Verbesserung nach 48 Wochen (Responder) (%)	21,1 vs. 10,5 p=0,297	\leftrightarrow

Endpunkt	Ataluren 10/10/20 mg vs. Placebo	Effekt
Zeit bis zu einer persistierenden Verschlechterung um mindestens 10 %	HR 0,52 (0,28;0,966) p=0,0386 (nominal) p=0,078 (Dunnett)	↑↑
Zeit bis zu einer persistierenden Verbesserung um mindestens 10 %	HR 1,675 (0,656;4,277) p=0,2805	↔
Sekundäre Endpunkte		
Änderung in der proximalen Muskelfunktion (,Timed Function Tests') zu Woche 48 Aufstehen aus Rückenlage (Sekunden, Mittelwert)	Differenz (absolut, 95%-KI): 0 (-2,3;2,3) relative Differenz: 0%	↔
Score für die Methode zum Aufstehen aus Rückenlage	Differenz (absolut, 95%-KI): 0,1 (-0,2;0,5)	↔
10 m laufen/gehen (Sekunden, Mittelwert)	Differenz (absolut, 95%-KI): -1,3 (-3,7;0,9) relative Differenz: -43%	↔
Score für die Methode, um 10 m zu laufen/gehen	Differenz (absolut, 95%-KI): 0,2 (-0,1;0,6)	↔
benötigte Zeit, um 4 Treppenstufen hinaufzusteigen (Sekunden, Mittelwert)	Differenz (absolut, 95%-KI): -2,4 (-4,9;0,1) relative Differenz: -50%	↔
Score für die Methode zum Treppenhochsteigen	Differenz (absolut, 95%-KI): 0,5 (0; 0,9)	↔
benötigte Zeit, um 4 Treppenstufen hinunterzusteigen (Sekunden, Mittelwert)	Differenz (absolut, 95%-KI): -1,6 (-4,3;1,0) relative Differenz: -40%	↔
Score für die Methode zum Treppenabsteigen	Differenz (absolut, 95%-KI): 0,3 (-0,2;0,6)	↔
Änderung der Alltagsaktivitäten gemessen mittels ,Step Activity Monitoring' (Pedometer am Knöchel) an ≥ 9 Tagen nach jeder Studiervisite (%)		
keine Aktivität	Differenz (95%-KI): -1,3 (-5,51;2,9)	↔
geringe Aktivität	Differenz (95%-KI): -0,01 (-2,9;2,88)	↔
mäßige Aktivität	Differenz (95%-KI): 1,23 (-0,25;2,71)	↔
ausgeprägte Aktivität	Differenz (95%-KI): 0,07 (-1,18;1,31)	↔
Änderung in der Häufigkeit von Rollstuhlnutzung (Patiententagebuch)	Differenz: 7,5 %	↔ ³⁾

Endpunkt	Ataluren 10/10/20 mg vs. Placebo	Effekt
Änderung in der Häufigkeit von Stürzen pro Tag (Patiententagebuch)	Relatives Risiko (95%-KI): 0,384 (0,16;0,94) p=0,0356	↑↑
Lebensqualität		
Änderung in der Lebensqualität gemessen mit dem PedsQL		
physikalische Subskala	3,56 (-4,31;11,42) p=0,3733 (nominal) p=0,5721 (Dunnett)	↔
emotionale Subskala	-4,2 (-11,5;3,05) p=0,2543 (nominal) p=0,4107 (Dunnett)	↔
soziale Subskala	-2,38 (-9,52;4,76) p=0,5117 (nominal) p=0,7328 (Dunnett)	↔
schulische Subskala	2,54 (-4,42;9,5) p=0,4727 (nominal) p=0,6907 (Dunnett)	↔
PedsQL Fatigue Scale	-2,41 (-8,85;4,03) p=0,4611 (nominal) p=0,6783 (Dunnett)	↔
Unerwünschte Ereignisse	Placebo vs. Ataluren 10/10/20 mg vs. Ataluren 20/20/40 mg	
mindestens 1 UE, N (%)	56 (98,2) vs. 55 (96,5) vs. 57 (95)	↔
mindestens 1 SUE, N (%)	3 (5,3) vs. 2 (3,5) vs. 2 (3,3)	↔
Zu einer Unterbrechung der Behandlung führendes UE, N (%)	13 (22,8) vs. 22 (38,6) vs. 9 (15)	↔

¹⁾ In der post-hoc-Analyse wurde folgender Effektschätzer für die kITT-Population ermittelt: Gehstreckendifferenz im modifizierten MMRM: 31,71 m (95%-KI 5,13;58,28), p=0,0367 (Dunnett-t-Test).

²⁾ p-Werte für die rangtransformierten Daten: p=0,1490 (nominal), p=0,2539 (Dunnett).

³⁾ Laut CSR (S. 113) nicht statistisch signifikant, Verzerrungspotential als „unklar“ eingestuft.

Referenzen

Bushby K, Finkel R, Birnkrant DJ, Case L, Clemens PR, Cripe L, Kaul A, Kinnett K, McDonald C, Pandya S, Poysky J, Shapiro F, Tomezsko J, Constantin C, DMD Care Considerations Working Group. Diagnosis and management of Duchenne muscular dystrophy. Part 1: diagnosis and pharmacological and psychosocial management. *Lancet Neurol* 2010; 9 (1): 77-93.

Bushby K, Finkel R, Wong B, Barohn R, Campbell C, Comi GP, Connolly AM, Day JW, Flanigan KM, Goemans N, Jones KJ, Mercuri E, Quinlivan R, Renfroe JB, Russman B, Ryan MM, Tulinius M, Voit T, Moore SA, Lee SH, Abresch RT, Coleman KL, Eagle M, Florence J, Gappmaier E, Glanzman AM, Henricson E, Barth J, Elfring GL, Reha A, Spiegel RJ, O'donnell MW, Peltz SW, McDonald CM. Ataluren treatment of patients with nonsense mutation dystrophinopathy. *Muscle Nerve* 2014; 50 (4): 477-87.

Committee for Medicinal Products for Human Use (CHMP). Translarna (ataluren). Assessment report for initial marketing authorisation application. London (UK): European Medicines Agency (EMA) 2014; (EMA/369266-2014). http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002720/WC500171816.pdf [Zugriff am 20.02.2015].

European Medicines Agency (EMA). Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. London (UK): European Medicines Agency (EMA) 2013; (EMA/CHMP/236981-2011). http://www.ema.europa.eu/ema/pages/includes/document/open_document.jsp?webContentId=WC500139508 [Zugriff am 20.02.2015].

Henricson E, Abresch R, Han JJ, Nicorici A, Goude KE, de BE, McDonald CM. 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.

Rubin M, The Merck Manuals. Duchenne Muscular Dystrophy and Becker Muscular Dystrophy [Online]. Whitehouse Station (USA): Merck Sharp & Dohme Corp 2014; http://www.merckmanuals.com/professional/pediatrics/inherited_muscular_disorders/duchenne_muscular_dystrophy_and_becker_muscular_dystrophy.html [Zugriff am 20.02.2015].

Kliegman RM, Behrman RE, Jenson HB, Stanton B. Chapter 608. In: Nelson textbook of pediatrics, 18th Edition. Saunders 2007.

Mazzone ES, Pane M, Sormani MP, Scalise R, Berardinelli A, Messina S, Torrente Y, D'Amico A, Doglio L, Viggiano E, D'Ambrosio P, Cavallaro F, Frosini S, Bello L, Bonfiglio S, De SR, Rolle E, Bianco F, Magri F, Rossi F, Vasco G, Vita G, Motta MC, Donati MA, Sacchini M, Mongini T, Pini A, Battini R, Pegoraro E, Previtali S, Napolitano S, Bruno C, Politano L, Comi GP, Bertini E, Mercuri E. 24 month longitudinal data in ambulant boys with Duchenne muscular dystrophy. *PLoS One* 2013; 8 (1): e52512. (Erratum in: *PLoS One*. 2013;8(11). doi:10.1371/annotation/cbe611fe-cda9-4d98-9574-0ac18e109daa).

McDonald CM, Henricson EK, Han JJ, Abresch RT, Nicorici A, Elfring GL, Atkinson L, Reha A, Hirawat S, Miller LL. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. *Muscle Nerve* 2010; 41 (4): 500-10.

McDonald CM. Natural history data argues against strength and function as co-primary endpoints in Duchenne muscular dystrophy [Presentation]. Treat-NMD Meeting Europe 2013; http://www.treat-nmd.eu/downloads/file/meetings/2013/workshop/Session2/McDonald_Strength_Function.pdf [Zugriff am 10.02.2015].

PTC Therapeutics. Statistical Analysis Plan. Protocol Number PTC124-GD-007-DMD 2010.

PTC Therapeutics. Clinical Study Report PTC124-GD-007-DMD 2012.

PTC Therapeutics. Ataluren (Translarna™). Dossier zur Nutzenbewertung gemäß § 35a SGB V (Modul 4A, 24.11.2014). PTC Therapeutics International Limited 2014.

Varni JW, Beaujean AA, Limbers CA. Factorial invariance of pediatric patient self-reported fatigue across age and gender: a multigroup confirmatory factor analysis approach utilizing the PedsQL Multidimensional Fatigue Scale. *Qual Life Res* 2013; 22 (9): 2581-94.