

**Dossier zur Nutzenbewertung  
gemäß § 35a SGB V**

*Ataluren (Translarna™)*

PTC Therapeutics International Limited

**Modul 4 A**

*Behandlung der Duchenne-Muskeldystrophie infolge einer Nonsense-Mutation im Dystrophin-Gen bei gehfähigen Patienten im Alter ab 5 Jahren*

Medizinischer Nutzen und  
medizinischer Zusatznutzen,  
Patientengruppen mit therapeutisch  
bedeutsamem Zusatznutzen

# Inhaltsverzeichnis

	Seite
<b>Tabellenverzeichnis</b> .....	<b>4</b>
<b>Abbildungsverzeichnis</b> .....	<b>7</b>
<b>Abkürzungsverzeichnis</b> .....	<b>9</b>
<b>4 Modul 4 – allgemeine Informationen</b> .....	<b>13</b>
4.1 Zusammenfassung der Inhalte von Modul 4.....	14
4.2 Methodik .....	23
4.2.1 Fragestellung .....	23
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung .....	25
4.2.3 Informationsbeschaffung .....	25
4.2.3.1 Studien des pharmazeutischen Unternehmers .....	25
4.2.3.2 Bibliografische Literaturrecherche .....	26
4.2.3.3 Suche in Studienregistern .....	26
4.2.3.4 Selektion relevanter Studien .....	27
4.2.4 Bewertung der Aussagekraft der Nachweise.....	28
4.2.5 Informationssynthese und -analyse .....	29
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien .....	29
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	30
4.2.5.3 Meta-Analysen.....	31
4.2.5.4 Sensitivitätsanalysen .....	32
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren .....	32
4.2.5.6 Indirekte Vergleiche .....	34
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen .....	36
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	36
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel .....	36
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	36
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	37
4.3.1.1.3 Studien aus der Suche in Studienregistern .....	40
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	40
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	41
4.3.1.2.1 Studiendesign und Studienpopulationen .....	41
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	53
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	54
4.3.1.3.1 Gehstrecke im 6MWT – RCT .....	54
4.3.1.3.2 Zeitlich gemessene Funktionstests – RCT .....	59
4.3.1.3.3 Unerwünschte Ereignisse – RCT .....	61
4.3.1.3.4 Subgruppenanalysen – RCT.....	64

4.3.1.3.5	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	64
4.3.2	Weitere Unterlagen.....	65
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien .....	65
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche .....	65
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	65
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen .....	66
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT .....	66
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT .....	68
4.3.2.2	Nicht randomisierte vergleichende Studien .....	69
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien .....	69
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	69
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien .....	70
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	70
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien .....	71
4.3.2.3	Weitere Untersuchungen.....	71
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen .....	72
4.3.2.3.2	Charakteristika der weiteren Untersuchungen .....	72
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen .....	72
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen .....	72
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen .....	73
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen.....	73
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	73
4.4.1	Beurteilung der Aussagekraft der Nachweise .....	73
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	74
4.4.2.1	Hintergrund und Fragestellung .....	74
4.4.2.2	Design der Studie PTC124-GD-007-DMD .....	77
4.4.2.3	Validität und Patientenrelevanz der Endpunkte.....	77
4.4.2.4	Ergebnisse der Studie PTC124-GD-007-DMD – RCT .....	85
4.4.2.4.1	Disposition der Patienten und Patientencharakteristika.....	85
4.4.2.4.2	Änderung der 6MWD und Progressor-Analyse - RCT.....	89
4.4.2.4.2.1	Primäranalyse und Sensitivitätsanalysen der 6MWD (ITT).....	89
4.4.2.4.2.2	kITT Analyse und Post-hoc-Analysen des 6MWT .....	91
4.4.2.4.2.3	10%ige Verschlechterung der 6MWD bis Woche 48 (Progressoranalyse) .....	98
4.4.2.4.2.4	Subgruppenanalysen - RCT.....	102
4.4.2.4.3	TFTs (sekundärer Endpunkt) - RCT .....	108
4.4.2.4.4	Andere sekundäre Endpunkte - RCT .....	112
4.4.2.4.5	Sicherheit und Verträglichkeit - RCT .....	114
4.4.2.4.6	Zusammenfassung der Ergebnisse aus RCT .....	122
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht .....	123
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte .....	123
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	123

4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	124
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen .....	124
4.5.4	Verwendung von Surrogatendpunkten .....	124
4.6	Liste der eingeschlossenen Studien.....	125
4.7	Referenzliste.....	126
<b>Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche .....</b>		<b>130</b>
<b>Anhang 4-B : Suchstrategien – Suche in Studienregistern.....</b>		<b>132</b>
<b>Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....</b>		<b>133</b>
<b>Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....</b>		<b>134</b>
<b>Anhang 4-E : Methodik der eingeschlossenen Studien – RCT .....</b>		<b>135</b>
<b>Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten .....</b>		<b>148</b>

**Tabellenverzeichnis**

	<b>Seite</b>
Tabelle 4-1: Darstellung der patientenrelevanten Endpunkte .....	17
Tabelle 4-2: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel .....	37
Tabelle 4-3: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	37
Tabelle 4-4: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel .....	40
Tabelle 4-5: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	41
Tabelle 4-6: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel .....	43
Tabelle 4-7: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	45
Tabelle 4-8: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	45
Tabelle 4-9: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	53
Tabelle 4-10: Matrix der Endpunkte in der eingeschlossenen RCT mit dem zu bewertenden Arzneimittel .....	54
Tabelle 4-11: Operationalisierung von Endpunkt 6MWD.....	56
Tabelle 4-12: Bewertung des Verzerrungspotenzials für den Endpunkt 6MWD (inklusive Progressor-Analyse in RCT mit dem zu bewertenden Arzneimittel.....	57
Tabelle 4-13: Ergebnisse für <Endpunkt xxx> aus RCT mit dem zu bewertenden Arzneimittel.....	58
Tabelle 4-14: Operationalisierung von Endpunkt TFT („timed function tests“) .....	59
Tabelle 4-15: Bewertung des Verzerrungspotenzials für den Endpunkt „TFTs“ in RCT mit dem zu bewertenden Arzneimittel .....	60
Tabelle 4-16: Ergebnisse für <Endpunkt xxx> aus RCT mit dem zu bewertenden Arzneimittel.....	60
Tabelle 4-17: Operationalisierung des Endpunkts „Unerwünschte Ereignisse“ (UEs) .....	62
Tabelle 4-18: Bewertung des Verzerrungspotenzials für den Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel.....	63
Tabelle 4-19: Ergebnisse für den Endpunkt „Unerwünschte Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel .....	63
Tabelle 4-20: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche .....	66
Tabelle 4-21: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	66

Tabelle 4-22: Operationalisierung von <Endpunkt xxx>.....	67
Tabelle 4-23: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche .....	67
Tabelle 4-24: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	68
Tabelle 4-25: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien .....	70
Tabelle 4-26: Operationalisierung von <Endpunkt xxx>.....	70
Tabelle 4-27: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien .....	71
Tabelle 4-28: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	72
Tabelle 4-29: Darstellung der berücksichtigten patientenrelevanten Endpunkte.....	76
Tabelle 4-30: Änderung des 6MWT stratifiziert nach PODCI-Ausgangswert, die bei Jungen mit DMD mit einer bedeutsamen Änderung des PODCI um 4.5 Punkte (MCID, Transfer und basale Mobilität) innerhalb eines Jahres einhergehen (13).....	80
Tabelle 4-31: Kortikosteroidanwendung zum Zeitpunkt der Randomisierung („as treated“-Population) .....	86
Tabelle 4-32: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	87
Tabelle 4-33: Änderung der 6MWD von BL bis Woche 48 in der prä-spezifizierten MMRM-Analyse unter Anwendung einer nicht strukturierten Kovarianzstruktur (ITT).....	90
Tabelle 4-34: Prä-spezifizierte Sensitivitätsanalysen der nicht transformierten 6MWD-Ergebnisse .....	91
Tabelle 4-35: Änderung der mittleren 6MWD über den Studienverlauf .....	93
Tabelle 4-36: Änderung der mittleren 6MWD über den Studienverlauf (augmentiertes gemischtes Modell mit unstrukturierter Kovarianzstruktur, kITT).....	94
Tabelle 4-37: MMRM-Analyse der Änderung der untransformierten 6MWD und Permutationstest (kITT) .....	96
Tabelle 4-38: Sensitivitätsanalysen Änderung der mittleren 6MWD (augmentiertes gemischtes Modell mit unstrukturierter Kovarianzstruktur).....	98
Tabelle 4-39: Anteil der Patienten, deren 6MWD sich dauerhaft um mindestens 10% bis Woche 48 verschlechterte .....	98
Tabelle 4-40: Zeit bis zu einer mindestens 10%igen Verschlechterung der 6MWD in dem Cox-proportionalen Hazard-Modell.....	100
Tabelle 4-41: Progressor Analyse (Verschlechterung der 6MWD um $\geq 0.5 \times SD$ bis Woche 48, beobachtete Werte, kITT) .....	101
Tabelle 4-42: Progressor Analyse (Verschlechterung der 6MWD um $\geq 0.5 \times SD$ bis Woche 48, fehlende Werte imputiert, kITT).....	102
Tabelle 4-43: Änderung der mittleren 6MWD über 48 Wochen in Abhängigkeit von den Stratifizierungsfaktoren ( <b>kITT</b> ) .....	103
Tabelle 4-44: Änderung der Dauer der TFTs von BL bis Woche 48 (kITT).....	108

Tabelle 4-45: Überblick der behandlungsspezifischen UEs aus RCT (Sicherheitspopulation) .....	115
Tabelle 4-46: Behandlungsspezifische UEs stratifiziert nach Organsystem aus RCT, die bei mehr als 5% der Patienten der Gesamtpopulation über 48 Wochen auftraten (Sicherheitspopulation) .....	116
Tabelle 4-47: Inzidenz der behandlungsspezifischen mit der Behandlung in Zusammenhang stehenden UEs aus RCT, die bei mehr als einem Patienten auftraten (Sicherheitspopulation) .....	117
Tabelle 4-48: Inzidenz der behandlungsspezifischen schweren UEs aus RCT (Sicherheitspopulation) .....	119
Tabelle 4-49: Inzidenz der behandlungsspezifischen schwerwiegenden UEs aus RCT (Sicherheitspopulation) .....	120
Tabelle 4-50: Übersicht der Ergebnisse der patientenrelevanten Endpunkte aus RCT: Ataluren 10/10/20 vs. Placebo (kITT).....	122
Tabelle 4-51: Überblick zum patientenrelevanten Zusatznutzen von Ataluren bei Patienten mit nmDMD im Alter von $\geq 5$ Jahren .....	123
Tabelle 4-52: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	123
Tabelle 4-53 (Anhang): Studiendesign und -methodik für Studie PTC124-GD-007-DMD..	136
Tabelle 4-54 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PTC124-GD-007-DMD .....	149

**Abbildungsverzeichnis**

	<b>Seite</b>
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel .....	39
Abbildung 2: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	58
Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	61
Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	64
Abbildung 5: Anteil der Jungen mit DMD stratifiziert nach BL 6MWD (30-m-Intervalle), die über zwei Jahre ihre Gehfähigkeit verloren (Mercuri et al. zitiert nach (30)).....	79
Abbildung 6: Verlauf der 6MWD in Abhängigkeit von BL-Werten .....	81
Abbildung 7: Änderung der 6MWD über 48 Wochen bei 57 DMD-Patienten in der Placebo-Gruppe der Phase 2b Ataluren-Studie (12) .....	82
Abbildung 8: Mittlere Änderung der 6MWD bis Woche 48 stratifiziert nach BL 6MWD (<350 vs. ≥350 m) in der Placebo-Gruppe der Phase 2b Ataluren-Studie (31) .....	82
Abbildung 9: Ergebnisse der 6MWD bei zwei Patienten mit Verletzungen der unteren Extremitäten, die inkorrekt Weise in die ITT-Population eingeschlossen worden waren ...	92
Abbildung 10: Mittlere Änderung der 6MWD bis Woche 48 (kITT).....	95
Abbildung 11: MMRM Monte Carlo-Analysen der 6MWD in der Gesamtpopulation (oben) und der „ambulatory decline“-Gruppe (unten) (31) .....	97
Abbildung 12: Zeit bis zu einer dauerhaften 10%igen Verschlechterung der 6MWD (kITT) .....	99
Abbildung 13: Mittlere Änderung der beobachteten 6MWD über den Studienverlauf in der Subgruppe BL 6MWD <350 m (kITT) .....	104
Abbildung 14: Mittlere Änderung der beobachteten 6MWD über den Studienverlauf in der Gesamtpopulation, "ambulatory decline phase"-Gruppe, <350 m BL-Gruppe (kITT) ...	105
Abbildung 15: Änderung der mittleren 6MWD bis Woche 48 stratifiziert nach prozentual erwarteter 6MWD zu Beginn der Studie (31) .....	106
Abbildung 16: MMRM Monte Carlo-Analysen des TFTs (vier Treppenstufen ersteigen) in der Gesamtpopulation (31).....	110
Abbildung 17: Ergebnisse der TFTs in der Gesamtpopulation, in der "ambulatory decline phase"-Gruppe und bei Patienten mit einer BL 6MWD <350 m (31). Werte sind Mittelwert-Unterschiede von Baseline zu Woche 48.....	111
Abbildung 18: Mittelwertsveränderungen in der physischen Domäne des pedsQL in der Gesamtpopulation, in der "ambulatory decline phase"-Gruppe, und bei Patienten mit einer BL 6MWD <350 m .....	113
Abbildung 19: Patientenfluss in der Studie PTC124-GD-007-DMD .....	147





**Abkürzungsverzeichnis**

<b>Abkürzung</b>	<b>Bedeutung</b>
$\Delta$	Delta
6MWD	6-minute walk distance
6MWT	6-minute walk test
10MWT	10-minute walk test
ACC	American College of Cardiology
ACTH	Adrenocorticotropes Hormon
AHA	American Heart Association
ANCOVA	analysis of covariance
ATS	American Thoracic Society
B	best
BL	Baseline
BMI	body mass index
CE	clinical evaluator
CHMP	Committee for Medicinal Products for Human Use
cm	centimetre(s)
CMH	Cochran-Mantel-Haenszel
CONSORT	Consolidated Standards of Reporting Trials
CRO	contract research organisation
CTCAE	Common Terminology Criteria for Adverse Events
d	day(s)
d.h.	das heißt
DBP	diastolischer Blutdruck
DIMDI	Deutsches Institut für Medizinische Dokumentation
DMC	data monitoring committee
DMD (nmDMD)	Duchenne Muskeldystrophie (Duchenne Muskeldystrophie infolge einer Nonsense-Mutation)
EG	Europäische Gemeinschaft
EKG	Elektrokardiogramm
EMA	European Medicines Agency
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice

h	hour(s)
HR	Hazard-Ratio
ICC	intra-class correlation coefficient
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
IVR/IWR	interactive voice response/interactive web response
kg	Kilogramm
KG	Körpergewicht
KI	Konfidenzintervall
kITT	korrigierte intention to treat
l	Liter
LOCF	last observation carried forward
LSM	least square mean
M	Mittelwert
m <sup>2</sup>	Quadratmeter
MCID	minimal clinically important difference
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligramm
min	Minute(n)
ml	Milliliter
MMRM	mixed-model repeated-measures
MPS	Mukopolysaccharidose
mRNA	Boten („messenger“)-Ribonukleinsäure
MTC	Mixed Treatment Comparison
n	Anzahl
n.a.	nicht anwendbar
nm	Nonsense-Mutation
nmDMD	Nonsense-Mutation Muskeldystrophie Typ Duchenne
OR	Odds-Ratio
p	probability
p.o.	<i>per os</i> (oral)
PD	Pharmakodynamik
PedsQL	Pediatric Quality of Life Inventory

PODCI	Pediatric Outcomes Data Collection Instrument
POSNA	Pediatric Orthopaedic Society of North America
PP	per Protokoll
PRO	patient-reported outcome
Prof.	Professor
PT	preferred term
QoL	Quality of Life
r	Reliabilität
RANCOVA	repeated-measures analysis of covariance
RANOVA	repeated-measures analysis of variance
RCT	Randomized Controlled Trial
SAM	Stepwatch <sup>®</sup> Activity Monitor
SAP	statistischer Analyseplan
SAS	statistisches Analysesystem
SBP	systolischer Blutdruck
SD	standard deviation
SE	standard error
sec	second(s)
SGB	Sozialgesetzbuch
SmPC	Summary of Product Charakteristiks
SOC	Standard of care
SOK	System-Organklasse
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	schwerwiegendes unerwünschtes Ereignis
TFT	timed function test
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
TSQM	Treatment Satisfaction Questionnaire for Medication
U	unit(s)
u.U.	unter Umständen
UAA	Uridin-Adenosin-Adenosin
UAG	Uridin-Adenosin-Guanosin

UE	unerwünschtes Ereignis
UGA	Uridin-Guanosin-Adenosin
UK	United Kingdom
USA	United States of America
VerfO	Verfahrensordnung
vs.	versus
W	worst
WHO	World Health Organization
z.B.	zum Beispiel
z.T.	zum Teil

#### 4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

## 4.1 Zusammenfassung der Inhalte von Modul 4

*Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.*

### Fragestellung

Translarna (Ataluren) ist zur Behandlung der Muskeldystrophie Typ Duchenne (DMD) aufgrund einer Nonsense-Mutation im Dystrophin-Gen bei gehfähigen Patienten im Alter von 5 Jahren und älter zugelassen. Die Nonsense-Mutation des Dystrophin-Gens muss mittels genetischer Methoden nachgewiesen werden. Die Muskeldystrophie Typ Duchenne ist eine sehr schwerwiegende X-chromosomal vererbte Erkrankung vorwiegend bei Jungen, die zu fortschreitendem Muskelabbau, Behinderung, dem frühen Verlust der Gehfähigkeit bis zur Rollstuhlabhängigkeit, Pflegebedürftigkeit und Tod im zweiten bzw. dritten Lebensjahrzehnt führt (1,2). Es handelt sich um eine sehr seltene Erkrankung. Bei einem Teil der Duchenne-Population enthält die mRNA des Dystrophin-Gens ein zusätzliches Stop-Codon, wodurch die Translation vorzeitig abgebrochen und ein verkürztes, nicht funktionales Protein gebildet wird. Durch Translarna wird das vorzeitige Stop-Codon überlesen, sodass die Translation bis zu dem normalen Stop-Codon fortschreiten kann und funktionales Dystrophin entsteht. Translarna ist das erste zugelassene kausal wirksame Medikament bei DMD aufgrund einer Nonsense-Mutation (nmDMD).

In diesem Dossier wird das Ausmaß des Zusatznutzens einer Therapie mit Translarna bei gehfähigen Patienten mit DMD im Alter von mindestens fünf Jahren dargestellt. Bei der DMD handelt es sich um ein seltenes Leiden. Translarna erhielt am 27. Mai 2005 die Designation als Arzneimittel für seltene Leiden zur Behandlung der DMD gemäß der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 (3). Für Arzneimittel zur Behandlung eines seltenen Leidens („Orphan Drugs“), die nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 zugelassen sind, gilt gemäß § 35a SGB V Abs. 1 Satz 10 Halbsatz 1 der medizinische Zusatznutzen durch die Zulassung als belegt. Es müssen keine Nachweise zum medizinischen Nutzen und medizinischen Zusatznutzen vorgelegt werden. Das Ausmaß des Zusatznutzens ist allerdings für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, nachzuweisen (4). Das Ausmaß des Zusatznutzens von Translarna wurde aus der für die Zulassung relevanten, kontrollierten Studie nach Maßgabe der in Kapitel 5, § 5 Absatz 7 Satz 1 Nr. 1 bis 4 der VerFO (2014) festgelegten Kriterien hinsichtlich der therapeutischen Relevanz (qualitativ) bewertet (4). Der Vorschlag des IQWiGs zur quantitativen, Zielgrößen-spezifischen Operationalisierung findet dabei keine Anwendung (5).

Die Studie PTC124-GD-007-DMD ist die erste Zulassungsstudie für nmDMD. Vor Beginn der Studie waren weder validierte Endpunkte noch der natürliche Verlauf der Erkrankung umfassend beschrieben. Die klinische Studie ist eine randomisierte, doppelblinde Placebo-kontrollierte Studie über 48 Wochen bei gehfähigen Patienten mit nmDMD Muskeldystrophie im Alter von mindestens fünf Jahren, die im 6-Minuten Gehstest mindestens 75 m bewältigen

konnten. Zur kausalen Behandlung der nmDMD steht bislang keine zugelassene Therapie zur Verfügung. Daher war die Wahl von Placebo plus best supportive care als Referenztherapie in der zulassungsrelevanten Studie angemessen. Die Dauer der Studie betrug 48 Wochen. Die Bewertung von Translarna vs. Placebo basiert auf der patientenrelevanten Zielgröße der Morbidität mit den Nutzendimensionen der Verbesserung des Gesundheitszustands und der unerwünschten Ereignisse [UEs].

### **Datenquellen**

Die Quantifizierung des Zusatznutzens von Translarna basiert auf dem Studienbericht der Placebo - kontrollierten Phase 2b-Zulassungsstudie PTC124-GD-007-DMD in dem zugelassenen Anwendungsgebiet (6–8). Der Beurteilungsbericht von Ataluren durch das CHMP (2014) wird bei der Bewertung berücksichtigt (9). Für Arzneimittel zur Behandlung seltener Leiden sind nur die Abschnitte 4.4.2 und 4.4.3 auszufüllen (10); daher sind eine systematische Literaturrecherche sowie Recherchen in Patientenregistern nicht erforderlich.

### **Ein-/Ausschlusskriterien für Studien**

Bei Arzneimitteln für seltene Leiden sind die zulassungsrelevanten Studien in der Nutzenbewertung zu berücksichtigen und es sind nur die Abschnitte 4.4.2 und 4.4.3 zu bearbeiten (10). Somit entfällt ebenfalls die Verpflichtung zur Definition von Ein- und Ausschlusskriterien im Falle von Arzneimitteln zur Therapie seltener Leiden. Da randomisierte, kontrollierte Studien den höchsten Grad der evidenzbasierten Medizin repräsentieren und damit ein direkter Kausalitätsbeweis potentieller Unterschiede zwischen der Intervention und der Vergleichstherapie ermöglicht wird, werden in der folgenden Nutzenbewertung nur randomisierte kontrollierte Studien berücksichtigt. Die noch laufenden offenen Studien mit Translarna in der zugelassenen Dosierung finden keine Berücksichtigung.

### **Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen**

Die zulassungsrelevante randomisierte, kontrollierte (SOC) klinische Studie PTC124-GD-007-DMD wurde gemäß den Anforderungen des CONSORT 2010 Statements (Items 2b bis 14) und des CONSORT-Flow-Charts beschrieben (siehe Anhang 4-E), obwohl diese Angaben für ein Arzneimittel für seltene Leiden nicht erforderlich gewesen wären. Die Bewertung der Aussagekraft der Nachweise erfolgte für die eingeschlossene, kontrollierte Studie entsprechend der Vorgaben der Verfahrensordnung des G-BA (4) zur Beurteilung des Verzerrungspotentials endpunktübergreifend und endpunktspezifisch (Abschnitt 4.3, Anhang 4-F). Das Verzerrungspotential der Endpunkte wurde nur für diejenigen Endpunkte bewertet, die tatsächlich in der Bewertung des Ausmaßes des Zusatznutzens von Translarna Berücksichtigung fanden. Diese Daten werden auf freiwilliger Basis zur strukturierten Darstellung sowie zur Bewertung der Ergebnissicherheit präsentiert.

Die prospektiv definierten primären und sekundären Endpunkte wurden patientenrelevanten Endpunkten wie Verbesserung des Gesundheitszustandes, Verkürzung der Krankheitsdauer,



Verlängerung des Überlebens, Verringerung von Nebenwirkungen oder Verbesserung der Lebensqualität zugeordnet (4). Der primäre Endpunkt der Studie PTC124-GD-007-DMD war die Änderung der Gehstrecke im 6-Minuten-Gehtest ("6-min-walk test", 6MWT) von Baseline (BL) bis Woche 48 unter oraler Therapie mit Ataluren (40 oder 80 mg/kg KG pro Tag) vs. Placebo. In der Nutzenbewertung werden i.A. nur die Ergebnisse der zugelassenen Dosierung von 40mg/kg Körpergewicht pro Tag bewertet. Der 6MWT ist von den Gesundheitsbehörden als etablierter und validierter primärer Endpunkt zur Messung der Ausdauer und funktionalen Mobilität in Studien bei DMD-Patienten anerkannt und wird als Maß für die Krankheitsprogression angesehen, bei dem die minimal klinisch relevanten Änderungen ("minimal clinically important difference", MCID) bekannt sind (11–13).

Als Sekundärendpunkt wird die Fähigkeit der Patienten, bestimmte funktionelle Tätigkeiten möglichst schnell ("timed function tests", TFTs) auszuüben wie z.B. vier Treppen hinauf oder hinab zu steigen, oder 10 m zu gehen. Diese Tests sind einfach anzuwenden und reproduzierbar, sie sind prädiktiv in Bezug auf klinische bedeutsame Meilensteine der Krankheitsprogression (14), und zeigen eine verlässliche Response auf Intervention mit Kortikosteroiden (15,16). Für die TFTs wurden ebenfalls MCIDs publiziert (11,13,17). TFTs werden als Maß angesehen, inwieweit ein DMD-Patient in der Lage ist, wichtige Alltagstätigkeiten zu verrichten (9).

In dem Beurteilungsbericht zu Ataluren bestätigt das CHMP (2014) die Patientenrelevanz der genannten Endpunkte und unterstützt auch die MCID von 30 m im 6MWT als klinisch bedeutsam für nmDMD Patienten (9). Die Änderung der Gehstrecke im 6MWT und die TFTs sind der Zielgröße Verbesserung des Gesundheitszustands zuzuordnen.

Daten zur Lebensqualität wurden in der Studie PTC124-GD-007-DMD ebenfalls erhoben. Die Patientenrelevanz dieser Ergebnisse ist derzeit nicht eindeutig zu beurteilen, da sich das verwendete Instrument einerseits als wenig sensitiv herausstellte, und andererseits für die verwendeten Instrumente keine *a priori* definierten Responsekriterien für nmDMD Patienten zur Verfügung stehen. Die patientenrelevante Zielgröße Mortalität konnte in der 48-Wochenstudie und angesichts der Seltenheit der Erkrankung nicht evaluiert werden, zumal hier nicht-gefähige Patienten primär betroffen wären.

In Tabelle 4-1 sind die in der Nutzenbewertung berücksichtigten patientenrelevanten Endpunkte zusammenfassend dargestellt.

Tabelle 4-1: Darstellung der patientenrelevanten Endpunkte

Patientenrelevante Zielgrößenkategorie	Endpunktparameter
Morbidität (Nutzendimension: Abschwächung schwerwiegender Krankheitssymptome bzw. Folgekomplikationen)	<ul style="list-style-type: none"> <li>- Gehstrecke im 6MWT bis Woche 48</li> <li>- Anteil der Patienten mit mindestens 10%iger Verschlechterung der Gehstrecke im 6MWT</li> <li>- Zeit bis zu einer mindestens 10%igen Verschlechterung der Gehstrecke im 6MWT</li> <li>- Anteil der Patienten mit mindestens 0.5 x SD Verschlechterung der Gehstrecke im 6MWT</li> <li>- TFTs: Zeit (Maximum 30 sec), die benötigt wird, um <ul style="list-style-type: none"> <li>• 10 m zu gehen oder zu rennen</li> <li>• vier Treppenstufen hinauf oder hinunter zu gehen</li> </ul> </li> </ul>
Nebenwirkungen (Nutzendimension: Vermeidung von [schwerwiegenden] Nebenwirkungen)	<ul style="list-style-type: none"> <li>- Nicht schwerwiegende (bzw. schwerwiegende) unerwünschte Arzneimittelwirkungen</li> </ul>

Zur Bewertung des Ausmaßes des Zusatznutzens von Ataluren wird die zulassungsrelevante Studie PTC124-GD-007-DMD herangezogen. Es handelte sich um eine randomisierte, Placebo-kontrollierte Doppelblindstudie der Phase 2b und der Evidenzstufe 1b. Nur ein Patient brach die Studie vorzeitig ab, sodass der Datensatz der ITT-Population und der Patienten mit vollständigem Datensatz nahezu identisch ist. Das ITT-Prinzip wurde adäquat umgesetzt.

Nach Entblindung des Datensatzes wurde allerdings erkannt, dass die BL-Werte der Gehstrecke im 6MWT von zwei Patienten, die kurz zuvor akute Verletzungen der unteren Extremitäten erlitten hatten und protokollgemäß durch den Prüfarzt als nicht valide zu klassifizieren gewesen wären, inkorrekt als valide klassifiziert worden waren. Diese im Vergleich zur Screening- und Folgevisite (Woche 6) unrealistisch niedrigen BL-Werte wurden in der ITT-Analyse berücksichtigt, was zu einer Verzerrung dieser Analysen geführt haben könnte. Daher wurden die Daten post-hoc in einer sogenannten korrigierten ITT-Population (kITT) ausgewertet. In der kITT-Analyse wurden die nicht validen BL-Werte durch die zuvor erhobenen Screening-Werte, die die tatsächlichen Fähigkeiten der Patienten widerspiegeln, ersetzt. Die kITT-Analyse wird als adäquat und klinisch sowie wissenschaftlich begründet eingestuft und dementsprechend in der vorliegenden Nutzenbewertung als *primär* angesehen. Auch das CHMP (2014) bewertete die kITT-Analyse als akzeptabel selbst unter Berücksichtigung der Tatsache, dass es sich um eine post-hoc Analyse handelt (9).

Bei der Bewertung der Studie PTC124-GD-007-DMD ist zu berücksichtigen, dass zum Zeitpunkt der Planung der Studie nur wenige Daten zum natürlichen Verlauf der Erkrankung bekannt waren. Beispielsweise war unsicher, welcher Endpunkt die Krankheitsprogression am besten widerspiegeln würde, wie hoch die Variabilität dieses Endpunkts wäre, oder welche Änderung klinisch relevant wäre. Im Ergebnis war die Variabilität des primären Endpunkts zu Studienende tatsächlich höher, als bei der Fallzahlschätzung angenommen worden war. Die

Ataluren-Studie selbst, sowie die Ergebnisse der Verlaufsstudien anderer Arbeitsgruppen haben zum besseren Verständnis der Pathogenese und des Verlaufs der Erkrankung sowie der geeigneten Erhebungsinstrumente beigetragen. Die Daten zeigen, dass bestimmte Patientengruppen, bzw. Patienten in einer bestimmten Phase des Krankheitsverlaufes (z.B.  $\geq 7$  Jahre, Gehstrecke  $< 350$  m im 6MWT) eine deutlich schlechtere Prognose zum Verlust der Gehfähigkeit haben als Patienten, die diesen Kriterien nicht entsprechen. Die Patientenpopulation, bei der die Krankheitsprogression besonders schnell fortschreitet (sogenannte "ambulatory decline phase"-Gruppe) erscheint daher besonders geeignet, innerhalb eines limitierten Zeitrahmens (z.B. ein Jahr in einer Placebo-kontrollierten Studie) einen Effekt eines Therapeutikums zu zeigen, welches primär die Muskelfunktion stabilisiert. Post-hoc Analysen der Studie PTC124-GD-007-DMD bestätigen diese Annahme. Die neu gewonnenen Erkenntnisse wurden beim Design der noch andauernden konfirmatorischen Phase 3 Studie mit Ataluren genutzt.

In der Studie PTC124-GD-007-DMD wurde besonderer Wert auf eine adäquate Operationalisierung der Endpunkte gelegt. Beispielsweise unterzogen sich die Studienzentren einem speziellen 6MWT-Trainingsprogramm und den Zentren wurde ein ausführliches Manual zur Verfügung gestellt, in dem die verschiedenen Testverfahren sowie die Interpretation der Ergebnisse detailliert beschrieben waren.

In der drei-armigen Studie PTC124-GD-007-DMD wurden die p-Werte auf Grund der in der Zulassung geforderten Hypothesentestung als adjustierte p-Werte angegeben (die p-Werte wurden auf Grund der Mehrfachtestung für die beiden unterschiedlich dosierten Ataluren-Arme adjustiert). Da sich in der klinischen Studie nur ein Dosierungsregime (Ataluren 10/10/20 mg/kg KG pro Tag) als wirksam erwies, und auch nur dieses Dosierungsregime zugelassen wurde, basiert die Quantifizierung des Zusatznutzens in diesem Dossier auf der niedrigeren Dosis (Ataluren 10/10/20 mg/kg KG pro Tag). Auch werden zur Beurteilung der Signifikanz der Studienergebnisse die nominalen p-Werte herangezogen, da das Ziel der Nutzenbewertung nicht die konfirmative Testung darstellt, sondern die Bestimmung der Therapie-Effektgröße auf Grundlage der evidenzbasierten Medizin.

Die Bewertung des Ausmaßes des Zusatznutzens von Ataluren erfolgt nach den Kriterien des §5 Absatz 7 des 5. Kapitels der VerfO (4).

### **Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen**

Die vorliegende Nutzenbewertung basiert auf einer randomisierten, Placebo-kontrollierten Doppelblindstudie bei Patienten mit genetisch gesicherter nmDMD im Alter von mindestens 5 Jahre, die mindestens 75 m gehen konnten (n=174 randomisierte Patienten, n=174 ITT und kITT). Die Patienten wurden randomisiert im Verhältnis 1:1:1 der oralen Behandlung mit Ataluren 10/10/20 oder 20/20/40 mg/kg KG pro Tag oder Placebo dreimal täglich über 48 Wochen zugeordnet. Die Dosis von Ataluren 10/10/20 mg/kg KG pro Tag entspricht der in Europa zugelassenen Dosierung.

Die Studie PTC124-GD-007-DMD hat nach der in der Verfahrensordnung des G-BA dargestellten Evidenzhierarchie eine Evidenzstufe 1b (4). Das Verzerrungspotential wird für die eingeschlossene Studie endpunktübergreifend als gering eingeschätzt. Das Verzerrungspotential wird ebenfalls für die Endpunkte ‚Gehstrecke im 6MWT‘ und ‚unerwünschte Ereignisse‘ als gering sowie für die Ergebnisse der ‚TFTs‘ als hoch eingestuft. Für die Bewertung des Ausmaßes des Zusatznutzens von Ataluren werden nur patientenrelevante Endpunkte berücksichtigt. Da der Erhalt der Gehfähigkeit für DMD-Patienten (sowie deren Angehörige) von höchster klinischer Bedeutung ist und einen „Meilenstein“ in der Krankheitsprogression darstellt, stützt sich die vorliegende Bewertung hauptsächlich auf die Ergebnisse des 6MWT. Die in der Studie erhobenen Surrogatparameter wie Expression von Dystrophin im Muskel flossen nicht in die Bewertung ein. Die Ergebnissicherheit der vorgelegten Daten wird unter Berücksichtigung der Seltenheit und des typischen Krankheitsverlaufs als insgesamt hoch eingeschätzt.

Primäres Ziel der Studie war es, den Effekt von Ataluren auf den Erhalt der Gehfähigkeit im 6MWT über 48 Wochen zu ermitteln. Die Gehfähigkeit im 6MWT ist ein zur Untersuchung von DMD-Patienten akzeptierter und validierter Endpunkt (siehe Kapitel 4.4.2.3). Das primäre Ziel der Studie (signifikanter Unterschied der Änderung der Gehstrecke bis Woche 48 zugunsten mindestens einer der Ataluren-Gruppen gegenüber Placebo) wurde in der protokollgemäßen ITT-Analyse allerdings nicht erreicht.

Wie bereits zuvor erläutert, werden die Daten der ITT Population dadurch verzerrt, dass von zwei Patienten inkorrekt Weise falsche BL-Werte in die ITT-Analyse eingeschlossen worden sind. Die korrigierte ITT-Analyse (kITT-Analyse) wurde von der CHMP als klinisch und wissenschaftlich begründet eingeschätzt (9). In der zulassungsrelevanten kITT-Analyse, in der die BL-Werte durch die zuvor erhobenen Screening-Werte ersetzt wurden, beträgt der Unterschied der Änderung der Model-basierten 6MWD ("6-min walking distance") zwischen Ataluren 10/10/20 und Placebo **31.7 m** (95% Konfidenzintervall: **5.1, 58.3**), und der Unterschied ist im prospektiv definierten MMRM-Modell signifikant (**p=0.0197 nominal**; p=0.0367 adjustiert). Ein Unterschied von 30 m im 6MWT, der im Studienprotokoll prospektiv als therapierelevanter Effekt definiert war, ist gleichzeitig bei DMD-Patienten von klinischer Relevanz, da eine Verminderung der Gehfähigkeit um diese Distanz mit dem Risiko des Verlusts der Gehfähigkeit in den nächsten Jahren assoziiert ist (9). Die Ergebnisse verschiedener Sensitivitätsanalysen sind in ihrer Ausrichtung vergleichbar und zeigen bei der zu erwartenden relativ hohen Variabilität einen konsistenten Effekt zugunsten von Ataluren 10/10/20 vs. Placebo. Die Robustheit der Primäranalyse wird durch einen post-hoc Permutationstest, der Abweichungen von den getroffenen Annahmen z.B. zur Normalverteilung adressiert, bestätigt.

Der Anteil der Patienten, deren 6MWD sich innerhalb von 48 Wochen um mindestens 10% verschlechterte, ist in der Placebo-Gruppe deutlich höher als in der Ataluren 10/10/20-Gruppe (kITT, HR 0.508; 95% KI 0.273, 0.946; p=0.0326 nominal, p=0.0652 adjustiert), was einer Risikoreduktion um 49% unter Ataluren entspricht. Auch dieses Ergebnis ist statistisch

signifikant und klinisch relevant. Es stützt nach Ansicht des CHMPs (2014) die Robustheit der primären Analyse der 6MWD (9).

Subgruppenanalysen werden mit dem primären Ziel durchgeführt, Patienten mit bereits weiter fortgeschrittener Erkrankung (z.B. 6MWD <350 m [*a priori* geplant] und "ambulatory decline phase"-Gruppe) zu erfassen, da sich in dieser Population die Gehfähigkeit rapide verschlechtert, und im 6MWT besser erfassen lässt. Zudem erreichten auf Grund des typischen Verlaufs der Erkrankung nicht alle Patienten während der 48-wöchigen Studiendauer die Phase einer rapiden Verschlechterung. Die Subgruppenanalysen der 6MWD zeigten einen deutlich größeren Effekt zugunsten von Ataluren 10/10/20 gegenüber Placebo, als die Ergebnisse für die gesamte Studienpopulation. Dieses Ergebnis stützt nachträglich das in der Studie gewählte Stratifizierungskriterium (<350 m vs. ≥350 m). Insgesamt erweist sich Ataluren 10/10/20 gegenüber der Vergleichstherapie als überlegen. Dennoch liefern die Subgruppenanalysen wichtige Informationen für das Design der laufenden konfirmatorischen Phase 3 Studie, in welcher Patienten in der "ambulatory decline phase" rekrutiert werden.

Hinsichtlich der „timed function tests“ (TFTs) ergibt sich ein signifikanter Vorteil zugunsten von Ataluren 10/10/20 vs. Placebo in der Fähigkeit, vier Treppenstufen zu passieren. Die Ergebnisse der anderen TFTs waren gleichgerichtet (exklusive Aufstehen aus dem Liegen), die Unterschiede waren aber verglichen mit Placebo nicht signifikant. In der Subgruppe der Patienten mit Baseline <350 m waren die Effektunterschiede in den TFTs noch deutlich größer.

Patientenberichtete, gesundheitsbezogene Lebensqualität wurde mittels des PedsQL Instrumentes erhoben. Die Auswertungen zeigen, dass Patienten, die mit Ataluren 10/10/20 behandelt wurden eine höhere mittlere Verbesserung in der physischen Dimension haben, als Patienten die mit Placebo behandelt wurden. Auch hier waren die Effektunterschied in der Subgruppe der Patienten mit Baseline <350 m deutlich größer.

Hinsichtlich anderer mit Sicherheit patientenrelevanter Endpunkte, die sich nachhaltig auf Folgekomplikationen auswirken können, wie z.B. die Zahl der versehentlichen Stürze oder die Häufigkeit der Benutzung eines Rollstuhls ergeben sich Hinweise auf einen Zusatznutzen von Ataluren 10/10/20 verglichen mit Placebo.

Die Verträglichkeit von Ataluren 10/10/20 und 20/20/40 ist trotz hoher Ereignisraten in allen Gruppen insgesamt sehr gut. Keiner der Patienten brach die Behandlung aufgrund von Nebenwirkungen ab und es wurde kein mit der Behandlung mit Ataluren in Zusammenhang stehendes schwerwiegendes UE berichtet.

### **Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen**

Der Zusatznutzen von Ataluren zur Behandlung des seltenen Leidens nmDMD gilt als durch die Zulassung belegt. Bei der nmDMD handelt es sich um eine sehr schwerwiegende, progredient verlaufende Erkrankung mit hohem Leidenspotential, die typischer Weise in der zweiten Lebensdekade zum vollständigen Verlust der Gehfähigkeit führt (1,2). Die nmDMD

ist mit erheblichen Einschränkungen der Ausdauer und Mobilität bis hin zum Verlust der Gehfähigkeit assoziiert, so dass der Erhalt der Gehfähigkeit für jeden Patienten ein primäres Therapieziel darstellt. (12). Der systematische Verlust der Gehfähigkeit stellt eine körperliche Behinderung dar, die in Pflegebedürftigkeit und frühzeitiger Letalität mündet. Die damit verbundenen Einschränkungen und Belastungen für die Patienten (und deren Angehörige) beeinflussen das gesamte Leben, so dass es sich bei dem Verlust der Gehfähigkeit um ein sehr schwerwiegendes Krankheitssymptom handelt.

Für die Gehstrecke im 6MWT wurden bei DMD-Patienten Schwellenwerte eines klinisch relevanten Unterschiedes publiziert. Ein Unterschied von 30 m im 6MWT wird bei DMD-Patienten als für die Patienten bedeutsam eingestuft:

- Zwei verteilungsbasierte Methoden zeigten MCIDs von 28.5 bzw. 31.7 m im 6MWT bei nmDMD-Patienten (11).
- Longitudinale Studien zum natürlichen Verlauf der DMD zeigten, dass jede Reduktion der BL 6MWD um 30 m das Risiko erhöht, innerhalb von zwei Jahren die Gehfähigkeit zu verlieren (Mercuri et al. zitiert nach (30)).

Aufgrund der vorgenannten Daten schätzt das CHMP (2014) einen Unterschied von 30 m im 6MWT als klinisch bedeutsam ein (9). Da die 6MWD mit dem Risiko assoziiert ist, zukünftig rollstuhlpflichtig zu werden (Mercuri et al. zitiert nach (30)), kann die 6MWD als Maß der Krankheitsprogression angesehen werden. Der in der Studie PTC124-GD-007-DMD beobachtete Therapieunterschied von nahezu 32 m im 6MWT zugunsten von Ataluren 10/10/20 vs. Vergleichstherapie wird daher als für die Patienten eindeutig relevant eingestuft.

Prospektiv definierter Endpunkt der Studie PTC124-GD-007-DMD war auch die Verschlechterung der Gehstrecke um mindestens 10%. Der natürliche Krankheitsverlauf der DMD Patienten zeigt, dass Patienten, deren Gehstrecke innerhalb von 12 Monaten um 10% oder mehr abnimmt, ein signifikant höheres Risiko haben, innerhalb von vier Jahren ihre Gehfähigkeit zu verlieren (12). Die Verschlechterung der Gehstrecke um 10% wurde in einer kürzlich veröffentlichten Studie als klinisch relevant eingestuft (13). Daher muss die in der Studie PTC124-GD-007-DMD demonstrierte Risikoreduktion um nahezu 50% (für eine 10%ige Verschlechterung der 6MWD innerhalb eines Jahres) unter Ataluren 10/10/20 verglichen mit Placebo als klinisch relevant eingestuft werden. Der therapeutische Zusatznutzen in Bezug auf die Verlangsamung der Krankheitsprogression wurde zusätzlich durch eine *post-hoc* durchgeführte Responderanalyse bestätigt. Die Responderanalyse wurde unter Verwendung eines Responsekriteriums (MCID für die Krankheitsprogression) gerechnet, welches mittels einer etablierten statistischen Verteilungsmethode abgeleitet wurde. Für den Endpunkt 6MWD ergibt sich somit in striktem Einklang mit den Kriterien des § 5 Absatz 7 des 5. Kapitels der Verfo (4) ein beträchtlicher Zusatznutzen für Ataluren in der zugelassenen Dosierung.

Hinsichtlich der Fähigkeit, Alltags-Tätigkeiten zügig auszuführen (timed function tests, TFTs), ergeben sich Hinweise auf einen patientenrelevanten Zusatznutzen von Ataluren

10/10/20, insbesondere in Bezug auf die Fähigkeit, vier Treppenstufen zu passieren. Der Unterschied zwischen Ataluren und Placebo liegt oberhalb der für diesen Endpunkt (4 Treppenstufen) bei DMD-Patienten publizierten MCIDs (11,17) und wird daher für die Patienten im Alltag klinisch relevant sein.

Hinsichtlich anderer patientenrelevanter Endpunkte wie z.B. der Zahl der versehentlichen Stürze kann das Ausmaß eines Zusatznutzens von Ataluren mit den vorliegenden Daten nicht gezeigt werden. Aufgrund methodischer oder anderer Unsicherheiten bezüglich der Interpretation der Daten werden diese Ergebnisse bei der Bewertung des Ausmaßes des Zusatznutzens von Ataluren 10/10/20 nicht berücksichtigt.

Die Analyse der UEs liefert keine Hinweise auf ein relevantes Schadenspotential von Ataluren 10/10/20, welches sich negativ auf die Bewertung des Ausmaßes des Zusatznutzens auswirkt. Auch das CHMP (2014) kommt in dem Beurteilungsbericht zu Ataluren zu dem Schluss, dass das Sicherheitsprofil von Ataluren bei einer sehr schwerwiegenden und lebensbedrohlichen Erkrankung wie der DMD akzeptabel ist und mögliche Risiken durch entsprechende Maßnahmen handhabbar sind (9). Zudem ist zu berücksichtigen, dass kein mit Ataluren in Zusammenhang stehendes schwerwiegendes UE und nur relativ wenige schwere UEs (häufigstes: Krankheitsprogression) mit vergleichbarer Inzidenz in allen Gruppen berichtet wurden. Dies muss zu der Abschwächung eines schwerwiegenden Krankheits-symptoms durch Ataluren 10/10/20 in Relation gesetzt werden.

Zusammenfassend wird das Ausmaß des Zusatznutzens von Translarna (Ataluren) in einer Dosierung von 10/10/20 mg/kg KG pro Tag in der zugelassenen Indikation als beträchtlich eingestuft, da eine klinisch relevante Verbesserung in Bezug auf die Ausdauer und Mobilität gezeigt wird, die mit einer Abschwächung schwerwiegender, krankheitsbedingter Symptome verbunden ist.

## 4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

### 4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

*Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.*

Ziel der vorliegenden Nutzenbewertung ist es, das Ausmaß des Zusatznutzens von Translarna Granulat zur oralen Suspension (125, 250 und 1000 mg Ataluren pro Sacht) in der zugelassenen Indikation zur Behandlung der genetisch gesicherten durch eine Nonsense-Mutation verursachte Muskeldystrophie Typ Duchenne (nmDMD) bei gehfähigen Patienten im Alter von fünf Jahren und älter zu bewerten. Bei der DMD handelt es sich um ein seltenes Leiden. Für Translarna wurde am 27. Mai 2005 die Orphan Designation zur Behandlung der DMD gemäß der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 ausgesprochen (3). Bei der DMD handelt es sich um eine schwerwiegende tödlich verlaufende X-chromosomal vererbte Erkrankung, bei der kein oder nicht funktionales Dystrophin gebildet wird, wodurch es u.a. zum fortschreitenden Muskelabbau kommt (1,2). Bei einem Teil der DMD-Patienten wird die Erkrankung durch ein zusätzliches Stop-Codon im Dystrophingen verursacht. Durch Translarna wird das vorzeitige Stop-Codon überlesen, sodass die Translation bis zu dem normalen Stop-Codon fortschreiten kann und funktionales Dystrophin entsteht. Translarna ist das erste kausal wirksame Medikament bei nmDMD.



Gemäß der gesetzlichen Vorgaben (§ 35a Absatz 1 Satz 10 SGB V) gilt für Medikamente zur Behandlung seltener Leiden der medizinische Zusatznutzen bereits durch die Zulassung als belegt; Nachweise zum medizinischen Nutzen und zum medizinischen Zusatznutzen im Verhältnis zur zweckmäßigen Vergleichstherapie müssen nicht vorgelegt werden. Lediglich das Ausmaß des Zusatznutzens ist für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, nachzuweisen (5. Kapitel § 12 Nr. 1 Satz 2 Verfahrensordnung [VerfO] des Gemeinsamen Bundesausschusses [G-BA]) (4). Der G-BA bestimmt bei Arzneimitteln für seltene Leiden, deren Umsatz die gesetzliche Grenze von 50 Millionen Euro nicht überschreitet (Anlage II - Format und Gliederung des Dossiers), das Ausmaß des Zusatznutzens auf der Grundlage der Zulassungsstudien.

Die Bewertung bezieht sich auf die Gesamtpopulation der Patienten in der zugelassenen Indikation, d.h. gehfähige Patienten mit genetisch gesicherter nmDMD im Alter von mindestens fünf Jahren.

Die Bewertung des Ausmaßes des Zusatznutzens von Translarna basiert auf den patientenrelevanten Endpunkten und der relevanten Zulassungsstudie (6–8). Die vorliegende Bewertung erfolgte für die Intervention Translarna im Vergleich zu Placebo anhand der patientenrelevanten Endpunkte Veränderung der Gehstrecke im 6-Minuten-Gehtest („6-min walk test“, 6MWT), 10%ige Verschlechterung der Gehstrecke im 6-Minuten-Gehtest, zeitlich begrenzter Funktionstests ("timed function tests", TFTs) und unerwünschter Ereignisse (UEs). Die Endpunkte sind den Zielgrößenkategorien Morbidität bzw. Nebenwirkungen zuzuordnen. Für Translarna liegen bislang keine Daten für die Endpunktkategorie gesundheitsbezogene Lebensqualität vor. Der patientenrelevante Endpunkt Mortalität konnte innerhalb der Dauer der eingeschlossenen Studie (48 Wochen) nicht evaluiert werden.

Da randomisierte, kontrollierte Studien den höchsten Grad der evidenzbasierten Medizin repräsentieren und damit ein direkter Kausalitätsbeweis potentieller Unterschiede zwischen der Intervention und der Vergleichstherapie ermöglicht wird, werden in der folgenden Untersuchung nur randomisierte kontrollierte Studien berücksichtigt. In die Bewertung wurde eine randomisierte, doppelblinde, Placebo-kontrollierte Parallelgruppenstudie mit Translarna der Phase 2b über ein Jahr einbezogen. Eine weitere confirmatorische, Placebo-kontrollierte Phase 3 Studie bei Patienten mit nmDMD ist noch nicht abgeschlossen. Bei der Planung dieser Studie wurden die Erkenntnisse der Translarna-Studie sowie die Daten anderer Arbeitsgruppen zum natürlichen Verlauf der DMD berücksichtigt. In der confirmatorischen Phase 3 Studie werden Patienten mit einem hohem Risiko einer Krankheitsprogression (sogenannte „ambulatory decline phase“) rekrutiert, die erwartungsgemäß innerhalb der Studiendauer von einem Jahr am meisten von einer Therapie profitieren, die die Muskelfunktion via Regeneration von Dystrophin stabilisiert oder verbessert. Aufgrund der niedrigeren Ergebnissicherheit werden keine offenen, nicht kontrollierten Studien für die Bewertung des Ausmaßes des Zusatznutzens herangezogen.

Da es sich bei Translarna um ein Arzneimittel zur Behandlung seltener Leiden handelt sind die Abschnitte 4.4.2 und 4.4.3 zu bearbeiten (10). Zur strukturierten Darstellung wurden dennoch einige weitere Abschnitte des Dossiers bearbeitet.

## 4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister erfolgen, während ein Kongressabstract allein in der Regel nicht für eine Studienbewertung ausreicht.

*Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar.*

Ausgehend von den gesetzlichen Vorgaben (§ 35a Absatz 1 Satz 10 SGB V) bestimmt der G-BA bei Arzneimitteln zur Behandlung seltener Leiden, die den 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 (10). Entsprechend basiert die vorliegende Nutzenbewertung auf der der Zulassung zugrunde liegenden randomisierten, Placebo-kontrollierten, Phase 2b-Studie PTC124-GD-007-DMD. Aus dieser Studie leiten sich die zugelassene Indikation sowie die Anwendungsbedingungen von Ataluren ab.

## 4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

### 4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

### 4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

*Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.*

Für die Bewertung des Ausmaßes des Zusatznutzens eines Arzneimittels zur Behandlung seltener Leiden sind die Zulassungsstudien relevant (10). Translarna (Ataluren) wurde am 31.07.2014 durch die europäische Zulassungsbehörde in der gegenständlichen Indikation zugelassen. Für das zu bewertende Arzneimittel wurde im Zeitraum von 12 Monaten vor dem Zeitpunkt der Dossiererstellung keine Zulassung in dem für das vorliegende Dossier relevanten Anwendungsgebiet erteilt. Für das vorliegende Arzneimittel ist eine bibliografische Literaturrecherche daher nicht erforderlich.

### 4.2.3.3 Suche in Studienregistern

Eine Suche in öffentlich zugänglichen Studienregistern ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten

vollständig identifiziert werden und in Studienregistern vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern [clinicaltrials.gov](http://www.clinicaltrials.gov) ([www.clinicaltrials.gov](http://www.clinicaltrials.gov)), EU Clinical Trials Register (EU-CTR, [www.clinicaltrialsregister.eu](http://www.clinicaltrialsregister.eu)), Klinische Prüfungen PharmNet.Bund (<http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm>) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: <http://apps.who.int/trialsearch/>) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmen ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister einzeln und mit einer für das jeweilige Studienregister adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

*Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.*

Die Suche in Studienregistern ist für Arzneimittel zur Behandlung seltener Leiden gemäß VerFO nicht erforderlich (4). Für Arzneimittel zur Behandlung seltener Leiden sind nur die Abschnitte 4.4.2 und 4.4.3 auszufüllen (10).

#### **4.2.3.4 Selektion relevanter Studien**

*Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2 und 4.2.3.3 beschriebenen Rechenschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.*

Das Ausmaß des Zusatznutzens von Arzneimitteln zur Behandlung seltener Leiden ist auf der Grundlage der Zulassung und der die Zulassung begründenden Studien zu bestimmen (10). Eine Selektion relevanter Studien entfällt.

#### 4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

##### A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie der behandelnden Personen
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

##### B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

*Beschreiben Sie die für die Bewertung der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

Der medizinische Zusatznutzen von Arzneimitteln zur Behandlung seltener Leiden gilt nach § 35a Absatz 1 Satz 3 Nummer 2 und 3 SGB V durch die Zulassung als belegt. Nachweise gemäß § 5 Absatz 1 bis 6 VerfO müssen nicht vorgelegt werden (4). Eine Bewertung der Aussagekraft der Nachweise ist für Arzneimittel zur Behandlung seltener Leiden daher nicht erforderlich (10). Dennoch wird zur Bewertung der dieser Nutzenbewertung zugrunde liegenden kontrollierten Studie das Verzerrungspotenzial sowohl auf Studienebene als auch auf Endpunktebene systematisch gemäß der oben beschriebenen Vorgaben und Methoden analysiert und im Bewertungsbogen in Anhang 4-F dargestellt, um die Aussagekraft der Nachweise und Ergebnissicherheit der der Zulassung zugrunde liegenden Studie abzuschätzen. Das Verzerrungspotential wurde lediglich für diejenigen Endpunkte analysiert, die tatsächlich bei der Bewertung des Ausmaßes des Zusatznutzens von Ataluren berücksichtigt werden.

Entsprechend des Vorgehens des G-BA wurde das Ausmaß des Zusatznutzens von Ataluren (qualitativ) nach Maßgabe der im 5. Kapitel § 5 Abs. 7 Satz 1 Nr. 1 bis 4 VerfO festgelegten Kriterien bewertet (4). Die seitens des IQWiG vorgeschlagenen inferenzstatistischen Schwellenwerte (Hypothesengrenzen) werden bei der Bewertung des Ausmaßes des Zusatznutzens von Ataluren nicht angewendet (5).

## **4.2.5 Informationssynthese und -analyse**

### **4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien**

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem

CONSORT-Flow-Chart)<sup>1</sup>. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-<sup>2</sup> bzw. STROBE-Statements<sup>3</sup> folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

*Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.*

Für Arzneimittel zur Behandlung seltener Leiden müssen die Abschnitte 4.3.1 und 4.3.2 nicht ausgefüllt werden (10). Das Design und die Methodik der eingeschlossenen Zulassungsstudie wurde in Kapitel 4.2.5.1, Kapitel 4.4.2, Anhang 4-E und Anhang 4-F entsprechend der Vorgaben für randomisierte Studien dargestellt. Insoweit erfüllt die Darstellung der zulassungsrelevanten, randomisierten kontrollierten Studie mindestens die Anforderungen des CONSORT-Statements.

#### **4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien**

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 4.3.2 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

*Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.*

---

<sup>1</sup> Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.

<sup>2</sup> Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Publ Health 2004; 94(3): 361-366.

<sup>3</sup> Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007; 147(8): 573-577.

Die Charakteristika der in die zulassungsrelevante Studie eingeschlossenen Patienten sind in Kapitel 4.4.2.4.1 beschrieben. Angaben zur Operationalisierung, Validität und Patientenrelevanz der dieser Bewertung zugrunde liegenden Endpunkte sind in Kapitel 4.3.1.2 und 4.4.2.3 dargestellt. Surrogatendpunkte werden bei der Bewertung des Ausmaßes des Zusatznutzens von Ataluren nicht berücksichtigt.

### 4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten<sup>4</sup> erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet<sup>5</sup> werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand des Maßes  $I^2$  und des statistischen Tests auf Vorliegen von Heterogenität<sup>6</sup> erfolgen. Ist die Heterogenität der Studienergebnisse nicht bedeutsam, soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden. Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe

<sup>4</sup> DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3): 177-188.

<sup>5</sup> Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

<sup>6</sup> Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.



Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

*Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

Da die vorliegende Bewertung auf einer Zulassungsstudie basiert, entfällt eine quantitative Zusammenfassung der Studien in einer Meta-Analyse.

#### **4.2.5.4 Sensitivitätsanalysen**

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

*Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

Folgende Sensitivitätsanalysen der 6MWD waren im statistischen Analyseplan (SAP) der PTC124-GD-007-DMD-Studie *a priori* spezifiziert (RANCOVA; siehe auch Kapitel 4.3.1.2.1) (7):

- Fehlende Patientencompliance mit dem Protokoll
- Einschluss von Zwillingspaaren
- Vorzeitige Entblindung
- Fehlende Daten
- Dynamische Randomisierung (Permutationstest)

#### **4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren**

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodi-

fikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. –stadium
- Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen.

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

*Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.*

Folgende Subgruppen- und supportiven Analysen der 6MWD waren im SAP der PTC124-GD-007-DMD-Studie *a priori* spezifiziert (RANCOVA; siehe auch Kapitel 4.3.1.2.1) (7):

- Alter (<9 vs. ≥9 Jahre)
- 6MWD (≥350 vs. <350 m)
- Kortikosteroidanwendung (ja vs. nein)
- Altersverteilung (e. g. 5-6, 7-8, 9-11, ≥12 Jahre)
- Typ des nm-Stop-Codons (UGA, UAG, UAA)
- Lokalisation des nm-Stop-Codons (<40 vs. ≥40)
- BL Größe und Gewicht
- Arzneimittel zur Prophylaxe/Behandlung des kongestiven Herzversagens

Da die Größe und das Gewicht der Patienten zu Beginn der Studie zwischen den Behandlungsgruppen ausgeglichen waren und nur wenige Patienten kardiale Medikationen einnahmen, wird auf diese Analysen verzichtet.

#### 4.2.5.6 Indirekte Vergleiche

Zurzeit sind internationale Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich<sup>7</sup>. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“<sup>8</sup>, „Multiple-Treatment-Meta-Analysen“<sup>9</sup> oder auch „Netzwerk-Meta-Analysen“<sup>10</sup>, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

Grundannahme für solche komplexen Analysen ist die Annahme der Konsistenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist<sup>11</sup>.

Da das Ergebnis eines indirekten Vergleichs maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen kann, ist die Wahl des Brückenkomparators bzw. der Brückenkomparatoren zu begründen. Dies gilt insbesondere dann, wenn eine Beschränkung auf ein oder mehrere Brückenkomparatoren vorgenommen wird, obwohl Daten zu anderen Therapieoptionen, die ebenfalls als Brückenkomparatoren in Frage kommen, vorliegen.

Insgesamt ist es notwendig, die zugrunde liegende Methodik genau und reproduzierbar zu beschreiben und die Annahme der Konsistenz zu untersuchen<sup>12</sup>.

---

<sup>7</sup> Glenny AM, Altman DG, Song F, Sakarovich C, Deeks JJ, D'Amico R et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005; 9(26): 1-148.

<sup>8</sup> Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

<sup>9</sup> Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; 331(7521): 897-900.

<sup>10</sup> Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17(3): 279-301.

<sup>11</sup> B. Schöttker, D. Lühmann, D. Boulkhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

<sup>12</sup> Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

*Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:*

- *Benennung des Brückenkomparators bzw. der Brückenkomparatoren und Begründung für die Auswahl.*
- *Genauere Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayesianischen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Nicht zutreffend.

### 4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2)
- Weitere Untersuchungen (Abschnitt 4.3.2.3)

#### 4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

##### 4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

###### 4.3.1.1.1 Studien des pharmazeutischen Unternehmers

*Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:*

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*
- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

*Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus*

*(abgeschlossen, abgebrochen, laufend), Studiendauer und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.*

Tabelle 4-2: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
PTC124-GD-007-DMD	ja	ja	abgeschlossen	48 Wochen 28. Februar 2008 bis 17. Dezember 2009	Placebo Ataluren 10/10/20 mg/kg KG (40 mg/kg KG/Tag) Ataluren 20/20/40 mg/kg KG (80 mg/kg KG/Tag)

*Geben Sie an, welchen Stand die Information in Tabelle 4-2 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.*

Da es sich bei Ataluren um ein Arzneimittel zur Behandlung seltener Leiden handelt und die vorliegende Bewertung auf einer Zulassungsstudie basiert, entfällt die Notwendigkeit der Darstellung des aktuellen Studienstatus.

*Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-2 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-3: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
Nicht zutreffend	

#### 4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

*Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese*

*zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.*

*[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]*

*Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.*

*Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.*

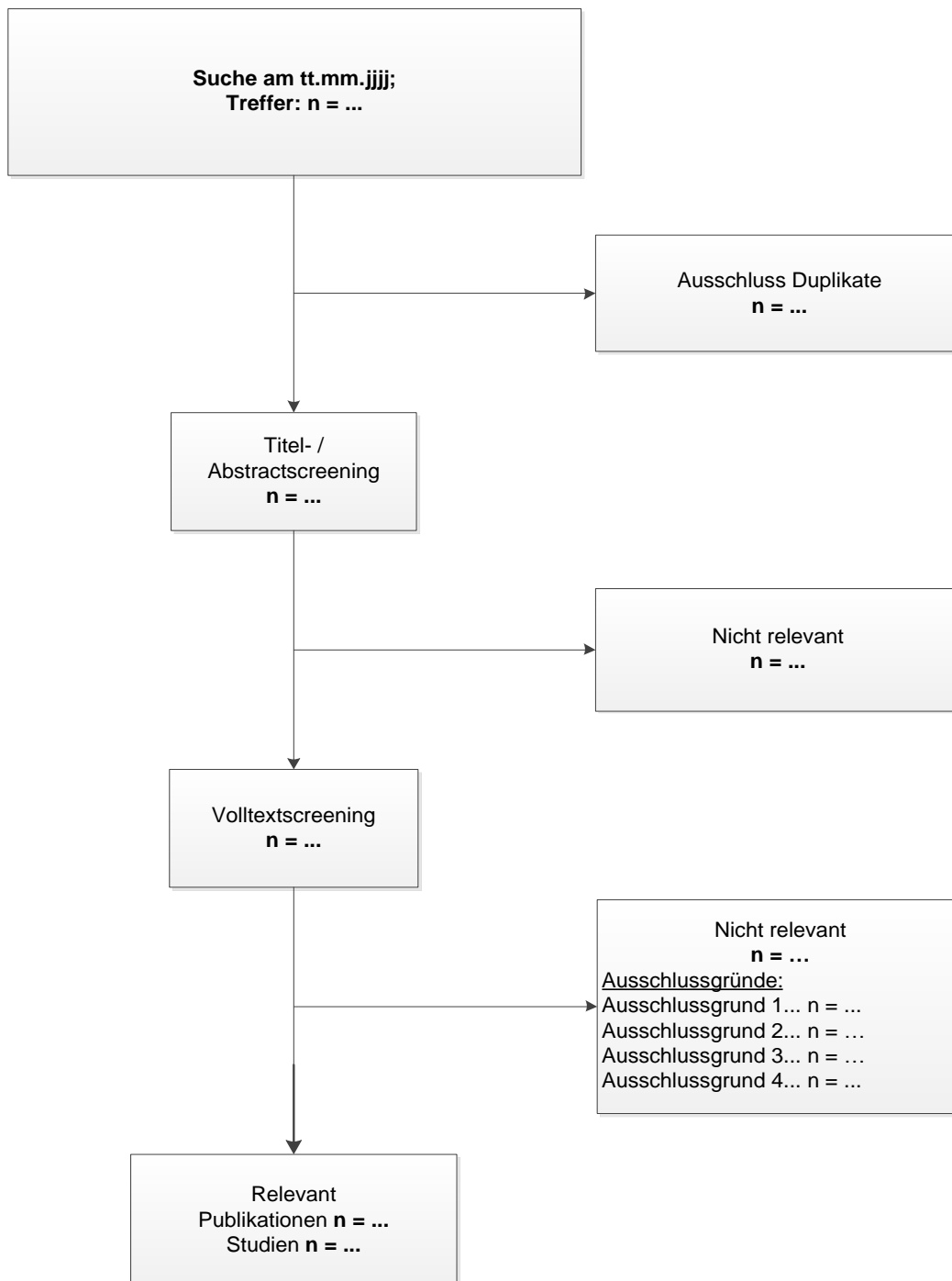


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Da es sich bei Translarna um ein Arzneimittel zur Behandlung seltener Leiden handelt, entfällt die Verpflichtung einer bibliografischen Literaturrecherche.



#### 4.3.1.1.3 Studien aus der Suche in Studienregistern

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern identifiziert wurden. Geben Sie dabei an, in welchem Studienregister die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-2) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-4: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters und Angabe der Zitate <sup>a</sup> )	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Status (abgeschlossen/ abgebrochen/ laufend)
PTC124- GD-007- DMD	clinicaltrials.gov (NCT00592553, EudraCT) (19,20)	ja	Nicht durchgeführt	abgeschlossen
PTC124- GD-020- DMD	clinicaltrials.gov (NCT01826487) (21)	ja	Nicht durchgeführt	laufend

a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.

Geben Sie an, welchen Stand die Information in Tabelle 4-4 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dokuments maßgeblichen Zeitpunkt liegen.

Da es sich bei Translarna um ein Arzneimittel zur Behandlung seltener Leiden handelt, entfällt die Verpflichtung zur Suche relevanter Studien in Studienregistern. Entsprechend wurde keine systematische Suche in Studienregistern durchgeführt.

#### 4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich

zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Quellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-5: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
<b>placebokontrolliert</b>						
PTC124 -GD- 007- DMD	ja	ja	nein	ja (8)	ja (19,20)	ja (22)
<b>aktivkontrolliert, zweckmäßige Vergleichstherapie(n)</b>						
	Nicht zutreffend					
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.7 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

Nicht zutreffend.

### 4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

#### 4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

*Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.*

Das Design der zulassungsrelevanten Studie PTC124-GD-007-DMD und die Charakteristika der Interventionen sind in Tabelle 4-6 bzw. Tabelle 4-7 beschrieben.

Tabelle 4-6: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
PTC124-GD-007-DMD (8)	Phase 2b, internationale, multi-zentrische, randomisierte, doppelblinde, Placebo-kontrollierte Dosisfindungsstudie	Gehfähige männliche Patienten mit nmDMD im Alter von $\geq 5$ Jahren; Fähigkeit, während des Screening-6MWTs $\geq 75$ m ohne Hilfen zu gehen	Placebo (n=57) Ataluren 10/10/20 mg/d (n=57) Ataluren 20/20/40 mg/d (n=60)	Screening-Phase mit 2 Visiten Behandlung: 48 Wochen Nachbeobachtung: Follow-up-Visite 6 Wochen ( $\pm 7$ d) nach der letzten Dosis oder Teilnahme an einer offenen Verlängerungsstudie; bei mit der Behandlung in Verbindung stehenden/schwerwiegenden UEs Nachbeobachtung bis UEs abgeklungen oder stabil Patienten, die $\geq 1$ Dosis Ataluren erhalten hatten: Follow-up über 5 Jahre geplant (Daten noch nicht verfügbar)	International (Australien, Belgien, Deutschland, Frankreich, Israel, Italien, Kanada, Spanien, Schweden, UK, USA) 28. Februar 2008 - 17. Dezember 2009	<u>Primärer Endpunkt:</u> 6MWD; <u>Patienten-relevante sekundäre Endpunkte:</u> Körperliche Funktion: TFTs (proximale Muskelfunktion), Muskelkraft (Myometrie), Aktivität im Alltag (SAM); Durch die Patienten/Betreuer berichtete Endpunkte: Gesundheitsbezogene Lebensqualität (PedsQL), Zufriedenheit mit der Behandlung (TSQM), akzidentielle Stürze pro Tag (Tagebuch) Kognitive Funktion: „Digit span task“ Kardiale Funktion: Herzrate

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

<b>Studie</b>	<b>Studiendesign</b> <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	<b>Population</b> <relevante Charakteristika, z. B. Schweregrad>	<b>Interventionen</b> (Zahl der randomisierten Patienten)	<b>Studiendauer</b> <ggf. Run-in, Behandlung, Nachbeobachtung>	<b>Ort und Zeitraum der Durchführung</b>	<b>Primärer Endpunkt; patientenrelevante sekundäre Endpunkte</b>
						Arzneimittelsicherheit (PD-Parameter siehe Tabelle 4-53)

Tabelle 4-7: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Placebo	Ataluren Niedrig- Dosis	Ataluren Hoch-Dosis	Komedikationen
PTC124-GD-007-DMD (8)	Dreimal täglich morgens, mittags und abends p.o. etwa 30 min nach den Mahlzeiten	10, 10, 20 mg/kg KG/d, analog Placebo	20, 20, 40 mg/kg KG/d, analog Placebo	Bei mit Kortikosteroiden (etwa 70% in allen Gruppen) oder kardialen Medikationen (15.8% Placebo, etwa 10% Ataluren) vorbehandelten Patienten sollte die Dosis während der Studie wenn möglich stabil bleiben. Nur minimale Änderungen dieser Medikationen wurden während der Studie dokumentiert.

Tabelle 4-8: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Alter (Jahre)	Geschlecht w/m (%)	<i>ggf. weitere Spalten mit Populationscharakteristika z. B. Dauer der Erkrankung, Schweregrad, weitere Basisdaten projektabhängig</i>
<Studie 1> <Gruppe 1> <Gruppe 2>		Siehe Tabelle 4-32		

*Beschreiben Sie die Studien zusammenfassend. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

*Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.*

### **Studie PTC124-GD-007-DMD (Studienbericht, -protokoll und SAP) (6–8)**

Die Wirksamkeit und Sicherheit von Ataluren bei Patienten mit nmDMD wurde in einer Phase 2b, multizentrischen, randomisierten, doppelblinden, Placebo-kontrollierten, GCP-konformen Studie über 48 Wochen überprüft. An dieser Stelle werden ausschließlich das Design und spezifische methodische sowie statistische Aspekte der Studie dargestellt. Eine zusammenfassende Darstellung insbesondere der Ergebnisse zur Wirksamkeit und Sicherheit findet sich in Kapitel 4.4.2.

**Primäres Ziel der Studie:** Effekt von Ataluren auf die Gehfähigkeit der Patienten

**Studienpopulation:** Männliche Patienten im Alter von  $\geq 5$  Jahren mit phänotypischer Evidenz einer DMD und Dokumentation einer Nonsense-Mutation im Dystrophin-Gen via Gensequenzierung durch ein zertifiziertes Labor wurden in die Studie eingeschlossen. Nur Patienten, die in der Lage waren während der Screening-Visite  $\geq 75$  m ohne Hilfen zu gehen, konnten an der Studie teilnehmen (siehe Anhang 4-E für die Einschluss- und Ausschlusskriterien).

**Intervention:** Die Patienten nahmen die Prüfpräparate (Placebo, Ataluren 10/10/20 mg/kg oder 20/20/40 mg/kg) dreimal täglich morgens, mittags und abends vorzugsweise 30 min nach den Mahlzeiten oral ein (6-, 6- bzw. 12-Stunden Dosisintervalle [ $\pm 1$  h]). Die Pulverformulierungen wurden in 30-60 ml stillem Wasser oder Milch aufgelöst, gerührt und anschließend eingenommen. Eventuelle Rückstände der Prüfpräparate wurden durch erneute Zugabe von Flüssigkeit suspendiert. Die geplante Behandlungsdauer betrug 48 Wochen. Visiten waren in Woche 6, 12, 18, 24, 30, 36, 42 und 48 vorgesehen. Eine Behandlungsdauer von mehr als einem Jahr wurde in einer Placebo-kontrollierten Studie unter Berücksichtigung der Schwere der Erkrankung als ethisch nicht vertretbar angesehen.

**Vergleichstherapie:** Zum Zeitpunkt der Studienplanung stand keine medikamentöse Standardtherapie zur Behandlung der DMD zur Verfügung, sodass die Wahl von Placebo plus SOC als Vergleichstherapie angemessen war. Das CHMP stimmte der Wahl von Placebo als Referenztherapie zu, da Ataluren das erste Medikament zur Behandlung der DMD ist („first-in-class“) und keine zugelassene Standardtherapie existiert (9).

***Primäre und sekundäre Endpunkte:*****Primärer Endpunkt:**

Änderung der 6MWD von BL bis Woche 48.

Zeit bis zur dauerhaften 10%igen Verschlechterung der 6MWD (supportiv für den primären Endpunkt)

**Sekundäre Endpunkte:**

Körperliche Funktion:

- Änderung der proximalen Muskelfunktion in TFTs (Aufstehen aus Rückenlage, 10 m rennen oder gehen und vier Treppenstufen hinauf- oder hinuntersteigen)
- Änderung der Kraft während einer Kniebeugung- oder Streckung, Streckung des Ellbogens oder Schulterabduktion ermittelt durch Myometrie.
- Änderung der Aktivität (Treppensteigen) im Alltag mittels eines Pedometer-ähnlichen Instruments (SAM, „Stepwatch<sup>®</sup> Activity Monitor“)

PROs („patient [parent/caregiver]-reported outcomes“):

- Änderung der Gesundheits-bezogenen Lebensqualität erfasst im PedsQL (körperliche, emotionale, soziale und Schul-bezogene Domäne sowie der Gesamt-Fatigue-Score)
- Änderung der Zufriedenheit mit der Behandlung (TSQM)
- Änderung der Rate der akzidentellen Stürze pro Tag erfasst in Tagebüchern

Kognitive Funktion: Änderung des verbalen Gedächtnisses und der Aufmerksamkeit im „digit span test“.

Kardiale Funktion: Änderung der Herzrate vor, während und nach dem 6MWT (oder in Ruhe bei Patienten, die nicht an dem Test teilnehmen konnten) gemessen mittels Polar<sup>®</sup> RS400.

Pharmakodynamik:

- Änderung der Serumkonzentration der Kreatinkinase.
- Änderung der Dystrophin-Expression im Bizeps mittels Immunfluoreszenz

Plasma Ataluren-Konzentrationen

Die Sicherheitsendpunkte schlossen UEs, Laborabnormalitäten, Vitalzeichen, EKG-Änderungen, renale Ultraschall-Untersuchungen und körperliche Untersuchungen ein. Die Compliance der Patienten wurde erfasst.

***Fallzahlschätzung:*** Basierend auf Studien bei Patienten mit anderen chronischen



Erkrankungen und Verlaufsdaten bei DMD-Patienten wurde angenommen, dass die mittlere Änderung der 6MWD von BL bis Woche 48 in einer der Ataluren-Gruppen mindestens 30 m größer als in der Placebo-Gruppe ausfallen würde. Unter der Annahme einer üblichen SD von etwa 50 m in allen Gruppen und einer 1:1:1 Randomisierung waren 150 Patienten notwendig, um eine Differenz von etwa 30 m im 6MWT mit einer >85%igen Power in wenigstens einer der Ataluren-Gruppen bei einem Signifikanzniveau von 0.042 (2-seitiger Dunnetts-Test) zu detektieren. Um einer antizipierten vorzeitigen Abbruchrate von etwa 10% Rechnung zu tragen, sollten etwa 160 Patienten (etwa 55 pro Arm) eingeschlossen werden.

**Randomisierung:** Die Patienten wurden randomisiert im Verhältnis 1:1:1 und stratifiziert nach Alter (<9 vs. ≥9 Jahre), Anwendung von Kortikosteroiden zu BL (ja / nein) und BL 6MWD (≥350 m vs. <350 m; initial: <270 vs. ≥270 m, geändert in Amendment 3.0 des Studienprotokolls) den Behandlungsgruppen zugeteilt. Die Zuteilung erfolgte zentral mittels IVR/IWR-System. Um das Risiko von Ungleichheiten zwischen den Gruppen zu minimieren, wurde bei der Randomisierung eine adaptive Methode eingesetzt (Anhang 4-E).

**Interim-Analysen:** Ein unabhängiges Experten-Komitee („data monitoring committee“, DMC) war für die Überwachung der Sicherheits-, Wirksamkeits- und anderer relevanter Daten verantwortlich und sollte basierend auf den Daten Empfehlungen hinsichtlich der Fortführung der Studie und gegebenenfalls notwendiger Design-Modifikationen aussprechen. Zwei Interim-Analysen waren prospektiv geplant, nachdem ≥45 Patienten ≥12 Wochen Behandlung abgeschlossen hatten und nachdem ≥90 Patienten (initial 100 Patienten) ≥24 Wochen Behandlung abgeschlossen hatten.

### **Studienpopulationen**

- ITT-Population: Die Population schloss alle randomisierten Patienten ein, die entsprechend des Randomisierungsplans zugeteilt worden waren, unabhängig davon, ob der Patient irgendeines der Prüfpräparate oder ein von der Randomisierung abweichendes Prüfpräparat erhalten hatte. Es musste mindestens ein valider BL 6MWD- und ein valider 6MWD-Wert nach BL vorliegen. Alle Zwillingspaare wurden in die ITT-Population eingeschlossen. Diese Population wurde in der Analyse der 6MWD, der zusätzliche Analysen der 6MWD und der Sekundärendpunkte ausgewertet. Zur Definition einer post-hoc korrigierten ITT-Population (kITT) siehe unten.
- Wie behandelt ("as-treated"): Die Population schloss alle randomisierten Patienten ein, die tatsächlich die randomisiert zugeteilten Prüfpräparate erhalten hatten. Diese Population wurde in der Analyse der Sicherheitsparameter und der Verabreichung der Behandlung ausgewertet.
- PP ("per-protocol")-Population: Die Population schloss alle randomisierten Patienten ein, die tatsächlich die randomisiert zugeteilten Prüfpräparate erhalten hatten, die Studie abgeschlossen hatten und einen validen 6MWD-Wert in Woche 48 hatten. Diese Population wurde in zusätzlichen Analysen der 6MWD ausgewertet.
- Auswertbare ("evaluable") Population: Die Population schloss alle Patienten der "as

treated"-Population ein (d.h. alle Patienten, die  $\geq 1$  Dosis der Prüfpräparate erhalten hatten) mit hinreichend vielen Werten zu BL und während der Studie, um interpretierbare Ergebnisse in den entsprechenden Analysen zu liefern (z.B. Ataluren-Plasmakonzentrationen oder Dystrophin-Expression).

### ***Statistische Analyse inklusive unterstützende und Sensitivitätsanalysen:***

Der primäre Endpunkt (Änderung der 6MWD von BL bis Woche 48) sollte mittels MMRM („mixed-model repeated-measure“-Analyse der Kovarianz, RANCOVA) ausgewertet werden. In die Primäranalyse wurde die ITT-Population eingeschlossen, d.h. alle randomisierten Patienten mit einem validen BL 6MWD-Wert und mindestens einem validen Wert nach BL. Das Modell erlaubte den Einschluss aller Patienten inklusive derjenigen mit fehlenden Werten aufgrund vorzeitigen Ausscheidens oder zeitweise fehlenden Werten ohne Imputation der fehlenden Werte. Die Varianzstruktur der Analysen wurde durch Modellierung der Daten unter Annahme einer zusammengesetzten Symmetrie („compound symmetry“), autoregressiven oder unstrukturierten Kovarianzstruktur bestimmt. Das Akaike-Informationskriterium wurde eingesetzt, um zu bestimmen, welche der hypothetischen Strukturen unter Berücksichtigung der Zahl der Parameter am besten die beobachtete Struktur abbildete.

Die Faktoren Behandlung, BL 6MWD, Alter ( $< 9$  oder  $\geq 9$  Jahre), Kortikosteroid-Anwendung (ja / nein), Visite und die Behandlung-Visite-Interaktion wurden in das Modell eingeschlossen. Da die BL 6MWD in dem Modell als Kovariate enthalten war, wurde die Stratifizierung ( $\geq 350$  oder  $< 350$  m) in dem Modell nicht berücksichtigt. Die LSMs und die Varianz wurden kalkuliert und die Änderungen in den Ataluren-Gruppen wurden mit denjenigen in der Placebo-Gruppe verglichen. Mittels Shapiro-Wilks W-Test wurde die Normalverteilung der Änderung der 6MWD bis Woche 48 bei einem Signifikanzniveau von 0.05 überprüft. Sollten die Werte signifikant von der Normalverteilung abweichen, sollten die log- oder falls notwendig rank-transformierten Daten in der Analyse eingesetzt werden. Die Änderung der 6MWD (Mittelwert und Varianz) von BL bis Woche 48 wurde zwischen den beiden Ataluren-Gruppen und Placebo verglichen (RANCOVA). Zusätzlich wurde der Dunnetts-Test eingesetzt, um der Multiplizität der Vergleiche Rechnung zu tragen. Im Falle einer Analyse der rank-transformierten Daten sollte die Holms-Methode der sequentiellen p-Wert-Adjustierung eingesetzt werden.

In einem MMRM-Modell mit BL 6MWD als Kovariate wurde der Einfluss der folgenden unabhängigen Variablen als Kovariaten auf die 6MWD überprüft: Stratifizierungsfaktoren (Alter, Anwendung von Kortikosteroiden), zusätzliche Altersklassen (z. B. 5-6, 7-8, 9-11 und  $\geq 12$  Jahre), Typ des Stop-Codons (UGA, UAG, UAA) und Lokalisation der Nonsense-Exonmutation. Auf die im Protokoll spezifizierten Analysen nach Größe und Körpergewicht zu BL und Anwendung von Medikationen zur Prophylaxe/Therapie der kongestiven Herzinsuffizienz wurde verzichtet. In einem MMRM-Modell, das die BL 6MWD als Kovariate einschloss, wurde der Effekt von Interaktionen zwischen der Behandlung und den Stratifikationsfaktoren (Behandlung nach Altersgruppen [ $< 9$  Jahre vs.  $\geq 9$  Jahre] und Behandlung nach 6MWD [ $\geq 350$  m vs.  $< 350$  m]) und einer Kortikosteroid-Anwendung auf den

Behandlungseffekt im 6MWT überprüft. Für jede Kovariate wurde das Interaktionsmodell eingesetzt, um den Behandlungseffekt für jede Kovariate abschätzen zu können.

Folgende Sensitivitätsanalysen waren geplant: Fehlende Compliance mit dem Protokoll, Einschluss von Zwillingspaaren, fehlende Daten (Analyse nach multipler Imputation) und dynamische Randomisierung in einem MMRM-Modell (Permutationstest). In dem Permutationstest wurden die Ergebnisse des primären Endpunkts mit den Ergebnissen aus 10'000 möglichen Kombinationen, in denen die Patienten randomisiert den Behandlungen hätten zugeordnet werden können, verglichen. Für jeden Datensatz wurden die Patienten dynamisch in derselben Reihenfolge wie sie initial in die Studie eingeschlossen worden waren re-randomisiert, und zufälligen „Seeds“ (Start-Seed 14576) zugeordnet. Um den möglichen Einfluss der Zuordnung von Zwillingspaaren zu derselben Behandlungsgruppe zu analysieren, wurde die primäre Analyse unter Berücksichtigung verschiedener Wege der Zuordnung der Zwillingspaare wiederholt. Der Mittelwert der resultierenden Modell-Parameter wurde zur Hypothesentestung eingesetzt. In einer zusätzlichen Sensitivitätsanalyse zur Überprüfung der Robustheit der Primäranalyse wurden fehlende Daten durch LOCF-Imputation (letzte verfügbare 6MWD nach BL) ersetzt und ein ANCOVA-Modell mit denselben Kovariaten eingesetzt. In eine zusätzliche Analyse wurde die PP-Population eingeschlossen. Wie in der Primäranalyse wurde das RANCOVA-Modell gefolgt von einem Dunnetts-Test in der PP-Analyse eingesetzt.

Als weitere primäre Analyse wurde der Anteil der Patienten, die auf die Therapie ansprachen bzw. deren Erkrankung fortschritt ( $\geq 10\%$  Verbesserung bzw. Verschlechterung der 6MWD in Woche 48), in der ITT- und PP-Population analysiert. Der Cochran-Mantel-Haenszel Chi-Quadrat-Test wurde eingesetzt, um den Anteil der Patienten, die auf die Therapie ansprachen bzw. deren Erkrankung im 6MWT fortschritt, pro Behandlungsarm zu vergleichen. Die ORs mit den zugehörigen 95% KIs wurden ebenfalls kalkuliert. Der mögliche Einfluss der Stratifizierungsfaktoren ( $< 9$  Jahre vs.  $\geq 9$  Jahre, Kortikosteroid-Anwendung und  $\geq 350$  m vs.  $< 350$  m 6MWD zu BL) und anderer BL-Charakteristika wurde mittels multipler logistischer Regression überprüft.

Wie im SAP festgelegt wurde die Zeit bis zu einer andauernden mindestens 10%igen Verschlechterung oder Verbesserung der 6MWD (letzter Zeitpunkt, zu dem die 6MWD nicht 10% schlechter oder besser als der Wert zu BL war) als supportive primäre Analyse ausgewertet. Bei Patienten ohne mindestens 10%ige Verschlechterung oder Verbesserung im 6MWT wurde der Wert zum Zeitpunkt des letzten 6MWT verwendet. Unterschiede zwischen den Ataluren-Gruppen und der Placebo-Gruppe wurden mittels Kaplan-Meier-Methode und einem stratifizierten log-rank-Test analysiert. Der mögliche Einfluss von Stratifizierungsfaktoren und anderer Patientencharakteristika zu Beginn der Studie (BL) auf diese Endpunkte wurde mittels eines Cox-Regressionsmodells ausgewertet.

In allen Analysen wurde bei Patienten, die aufgrund einer Krankheitsprogression nicht mehr in der Lage waren, an dem 6MWT teilzunehmen, ab diesem Zeitpunkt ein Wert von 0 m eingegeben. In den Analysen der log-transformierten Daten wurde für Null-Werte in dem

Original-Datensatz ein Wert von 1 und der niedrigste beobachtete 6MWD-Wert mit optimaler Normalverteilung eingegeben.

Die Sekundärparameter wurden in der ITT- und PP-Population ausgewertet. Wenn nicht anders angegeben, wurden die Sekundärparameter in dem MMRM-Modell ausgewertet (RANCOVA gefolgt von einem Vergleich der Änderung von BL bis Woche 48 im Dunnetts-Test). Der BL-Wert der jeweiligen Variable diente dabei als Kovariate und die unabhängigen Variablen schlossen die Behandlung, Visite, Behandlung-Visite und die Stratifikationsfaktoren (Alter, Kortikosteroid-Anwendung und BL 6MWD) ein. Die Daten wurden zu allen Untersuchungszeitpunkten analysiert. Die in der Primäranalyse eingesetzte Varianzstruktur sollte verwendet werden. Die LSMs, SEs und 95% KIs wurden für jeden Behandlungsarm, jede Visite und die Stratifizierungsfaktoren berechnet (Ataluren vs. Placebo in PROC MIXED). Ein Shapiro-Wilks-Test wurde durchgeführt und, falls notwendig, sollten die transformierten Daten (log- oder rank-basiert) analysiert werden. Falls die Original- oder log-transformierten Daten ausgewertet wurden, sollte der Dunnetts-Test eingesetzt werden, um für zwei Vergleiche vs. Placebo zu adjustieren. Der Holms-Test wurde wie oben beschrieben eingesetzt. Die TFTs wurden mittels eines Random-Effekt-Modells für Ordinalzahlen in SAS ausgewertet (PROC NLMIXED).

#### *Post-hoc-Analysen und Änderungen des SAP*

Die MMRM-Methode diente als primäre Analyse der 6MWD und die ANCOVA mit LOCF-Imputation als unterstützende Analyse. 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 des Biostatistikers Prof. Gary Koch (Universität Nord-Carolina) wurde vermutet, dass ein BL-Visite Interaktionsterm in dem MMRM-Modell fehlt, da die Beziehung zwischen BL 6MWD und post-BL 6MWD mit der Zeit variieren kann. Die BL-Visite Interaktion war hoch signifikant ( $p<0.001$ ), was andeutet, dass der Effekt der BL 6MWD über die Zeit nicht konstant ist. Die Diskrepanz zwischen der MMRM- und ANCOVA-Analyse verschwand, wenn der zusätzliche Interaktionsterm eingeschlossen wurde.

Nach Aussage von Prof. Dr. Koch liefert der Permutationstest der nicht transformierten Daten (10'000 Re-Randomisierungen) akkuratere Daten zum Signifikanzniveau der Behandlungsvergleiche als p-Werte der rank-transformierten Daten. Der Permutationstest wurde zusätzlich in der prä-spezifizierten ANCOVA-LOCF-Analyse der nicht transformierten Daten eingesetzt. Zur Adjustierung der Multiplizität wurde der Anteil der 10000 Permutationen, in welchen die maximale Effektgröße der beiden Vergleiche (Ataluren 10/10/20 und 20/20/40 vs. Placebo) die beobachtete maximale Effektgröße überstieg, bestimmt. Zusätzlich wurde ein Permutationstest der Zahl der unbeabsichtigten Stürze eingeführt, um das protokollgemäße GEE-Modell zu validieren.

Zu Studienbeginn lagen keine normativen Daten zur 6MWD bei DMD-Patienten vor. Da derartige Daten zum Zeitpunkt der Studiauswertung verfügbar waren, wurden post-hoc Auswertungen zur erwarteten 6MWD in Prozent über 48 Wochen verglichen mit gesunden Jungen ausgeführt. Da sich in der Studie eine inverse Dosis-Effekt-Beziehung zeigte, wurde zusätzlich der Zusammenhang zwischen den Ataluren-Plasma-Konzentrationen und den Wirksamkeitsparametern analysiert.

Auf die geplante Auswertung eines Summenscores der Muskeltests (Myometrie) der oberen bzw. unteren Körperpartien wurde verzichtet, da insgesamt nur geringe Änderungen beobachtet wurden

Vor Studienbeginn mussten sich die für die 6MWTs verantwortlichen Personen in den Studienzentren einem Training unterziehen, um möglichst standardisierte Testbedingungen zu gewährleisten. Nach Entblindung der 6MWD-Daten zeigte sich, dass zwei Patienten der Placebo- bzw. der Ataluren 20/20/40-Gruppe in einem Studienzentrum innerhalb weniger Tage vor dem BL 6MWT akute Verletzungen an den unteren Extremitäten erlitten hatten, die sich verglichen mit dem Screening- und dem nachfolgenden 6-Wochen-Wert deutlich negativ auf die BL-Gehstrecke ausgewirkt hatten (Patient Nr. 501-014: 395, 309 [BL] bzw. 418 m; Patient Nr. 501-012: 303, 125 [BL] bzw. 309 m; siehe Abbildung 9). Vor der Studie waren Regeln festgelegt worden, wie die Validität eines 6MWT-Versuchs zu bewerten war und unter welchen Bedingungen der Test zu wiederholen war (siehe Tabelle 4-11). Protokollgemäß hätte der klinische Beurteiler die Werte der beiden vorgenannten Patienten als nicht valide klassifizieren müssen. Die Werte wurden allerdings inkorrekt als valide eingestuft. Wie im SAP festgelegt durften nur Patienten mit einer validen BL 6MWD und mindestens einer validen post-BL 6MWD in die ITT-Population eingeschlossen werden. Daher wurden post-hoc-Analysen durchgeführt, in denen die invaliden BL 6MWD-Werte der beiden Patienten durch die Screening-Werte ersetzt wurden, die die tatsächlichen Fähigkeiten dieser Patienten besser widerspiegeln als die Tests unmittelbar nach den Beinverletzungen (kITT). Die Ergebnisse der TFTs wurden entsprechend post-hoc für die kITT Population analysiert. Zusätzliche post-hoc Analysen, in denen die Werte dieser beiden Patienten ausgeschlossen wurden, oder durch andere Werte ersetzt wurden, dienten als Sensitivitätsanalysen.

Da die kITT-Analyse wissenschaftlich und klinisch begründet ist und die tatsächlichen Fähigkeiten der Patienten widerspiegelt, bezieht sich die vorliegende Nutzenbewertung schwerpunktmäßig auf diese Analysen. Die Ergebnisse in der ITT-Population (mit den beiden invaliden 6MWD-Werten) wurden aus informativen Gründen dargestellt wenn dies als sinnvoll erachtet wurde. Das CHMP (2014) kommentierte die kITT-Analyse dahingehend, dass die Analyse methodisch gesehen als akzeptabel eingestuft werden könne, es sich aber *de facto* um post-hoc-Ergebnisse handle. Die Ergebnisse der kITT-Analyse flossen in die Bewertung von Ataluren durch das CHMP (2014) ein (9).

Die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext wird in Kapitel 4.4.2 diskutiert.

### 4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-9: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
PTC124-GD-007-DMD	ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential wird auf Studienebene unter Berücksichtigung der in Tabelle 4-9 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Bei der Studie PTC124-GD-007-DMD handelt es sich um eine multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte internationale Parallelgruppenstudien der Phase 2b. Eine adäquate Erzeugung der Randomisierungssequenz liegt vor. Eine adaptive Randomisierung unter Einsatz akzeptierter Minimierungstechniken und Permutationsalgorithmen wurde angewandt (Anhang 4-E), um mögliche Ungleichheiten zwischen den Gruppen hinsichtlich der prä-definierten Stratifikierungskriterien und der Zahl der Patienten pro Arm zu vermeiden bzw. deren Risiko zu minimieren. Die Gruppenzuteilung erfolgte verdeckt. Sowohl die Patienten als auch die Prüfarzte und das medizinische Personal waren bis zur Endauswertung der Studie verblindet. Lediglich ein Patient (Ataluren 20/20/40) brach die Studie etwa in Woche 6 vorzeitig ab. Die BL- und Woche 6 Daten dieses Patienten wurden in allen MMRM-Analysen berücksichtigt. Das ITT-Prinzip wurde adäquat umgesetzt. In die Analyse der kITT-Population wurden wie im Protokoll vorgesehen nur Patienten mit validen 6MWD-Werten eingeschlossen, so dass die verletzungsbedingt unrealistisch niedrigen BL-Werte von zwei Patienten durch deren Screening-Werte ersetzt wurden. Diese Analyse wird trotz ihres post-hoc-Charakters als klinisch und wissenschaftlich begründet eingestuft (siehe 4.3.1.2.1), was sich mit der Einschätzung des CHMPs in dem Beurteilungsbericht von Ataluren deckt (9). Wie bei einer sehr seltenen Erkrankung zu erwarten, ist die Fallzahl relativ niedrig. Unter Berücksichtigung der Seltenheit der Erkrankung ist die Zahl der eingeschlossenen Patienten allerdings als

akzeptabel einzuschätzen. Es liegen keine Hinweise auf eine ergebnisgesteuerte Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, vor.

#### 4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-10: Matrix der Endpunkte in der eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität	Gesundheitsbezogene Lebensqualität	Morbidität
PTC124-GD-007-DMD	nein	ja (PedsQL)	ja (6MWD inklusive Progressor-Analyse, TFTs, UEs)

In der Studie wurden weitere u.U. patientenrelevante Endpunkte erhoben wie z.B. die Zahl der unbeabsichtigten Stürze oder die Häufigkeit der Benutzung eines Rollstuhls. Da diese Daten (inklusive Daten zur Lebensqualität, PedsQL) allerdings keine Rückschlüsse zum Ausmaß des Zusatznutzens zulassen oder hinsichtlich des Patientennutzens nicht eindeutig interpretierbar sind, wird auf eine detaillierte Darstellung dieser Endpunkte verzichtet (siehe Kapitel 4.3.1.2.1 und 4.4.2.4.4 für eine kurze Übersicht). Im Folgenden werden die Operationalisierung und das Verzerrungspotential nur für diejenigen Endpunkte dargestellt, die tatsächlich bei der Bewertung des Ausmaßes des Zusatznutzens von Ataluren berücksichtigt wurden. Die Ergebnisse der als patientenrelevant erachteten und für die Einschätzung des Ausmaßes des Zusatznutzens von Ataluren relevanten Endpunkte werden in Kapitel 4.4.2 zusammenfassend dargestellt.

##### 4.3.1.3.1 Gehstrecke im 6MWT – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind

- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler
- Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden.

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

*Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.*



Tabelle 4-11: Operationalisierung von Endpunkt 6MWD

Studie	Operationalisierung
PTC124-GD-007-DMD	<p>Das Gehvermögen der Patienten wurde im 6MWT nach einer modifizierten Version der ATS-Richtlinie (23) mittels eines an der Universität Kalifornien-Davis entwickelten, standardisierten Verfahrens erfasst (24). Den Prüfbüros wurde ein Manual zur Verfügung gestellt, in dem der Test detailliert beschrieben war. Modifikationen gegenüber der ATS-Version schlossen ein kurzes Video zur Orientierung vor dem Test, eine kontinuierliche Motivation durch das Klinikpersonal und eine Person zur Sicherung, die dem Patienten während des Tests folgte, ein. Die Tests sollten in einem ausreichend breiten Flur mit harter Oberfläche, der mit einem 25 m langen Massband markiert war, stattfinden. Der Start- bzw. Wendepunkt sollte jeweils mit Kegeln markiert werden. Die einzelnen Testschritte inklusive der Überwachung des Pulses vor, während und nach dem Test waren in dem Manual detailliert beschrieben und die Phrasen zur Motivation der Patienten waren vorgegeben, um möglichst standardisierte Testbedingungen zu schaffen. Die benötigte Zeit oder gegebenenfalls Zeit bis zu einem Sturz wurde mittels digitaler Stoppuhr erfasst. War der Patient nach einem Sturz während des Tests unverletzt und war der Patient dazu in der Lage, konnte der Test fortgeführt werden. War dies nicht der Fall wurde die Gehstrecke zum Zeitpunkt des Sturzes als 6MWD erfasst. Benötigte der Patient eine Ruhepause konnte er sich an der Wand anlehnen und wurde angehalten, den Test so bald wie möglich fortzusetzen.</p> <p>Vor Studienbeginn nahmen die Zentren an einem Endpunkttraining und einer Standardisierungssitzung teil, um die Testprotokolle und Logistik zu harmonisieren. Nach etwa einem Jahr fand ein zentrales erneutes Training statt.</p> <p>Wie im SAP (7) festgelegt wurde der 6MWT als valide eingestuft, wenn der Patient den Test wie geplant abschloss oder den Test ausschließlich aufgrund von Erschöpfung vorzeitig abbrach. Unter folgenden Bedingungen galt der 6MWT als nicht valide: Der Patient</p> <ul style="list-style-type: none"> <li>- folgte nicht den Anweisungen,</li> <li>- behielt den Kurs während der Dauer des Tests nicht ein,</li> <li>- ging zu irgendeinem Zeitpunkt während des Tests in eine entgegengesetzte Richtung und/oder</li> <li>- brach den Test aufgrund fehlender Compliance oder anderer Gründe vorzeitig ab (&lt;6 min), die nicht mit Erschöpfung in Zusammenhang standen (z.B. Verletzung durch Sturz, Hinsetzen).</li> </ul> <p>Die Zentren wurden angehalten im Falle eines nicht validen initialen Tests oder für den Fall, dass der Patient zu erschöpft war, um den ersten Test zu beenden zu versuchen, den Test zu wiederholen (nach einer mindestens 1-stündigen Ruhepause an Tag 1 oder an Tag 2). Bei gehfähigen Patienten ohne validen Test, sollte die 6MWD als fehlend eingegeben werden.</p> <p>Die 6MWD wurde zur Screening-Untersuchung, Tag 1 (Visite 3 vor Beginn der Bahndlung), Woche 6 (Visite 4), Woche 12 (Visite 5), Woche 18 (Visite 6), Woche 24 (Visite 7), Woche 30 (Visite 8), Woche 36 (Visite 9), Woche 42 (Visite 10) und Woche 48 (Visite 11 oder Ende der Behandlung) mittels 6MWT erfasst. An Tag 1 sowie in Woche 24 und 48 sollten die 6MWTs mittels Videokamera dokumentiert werden sofern eine Einverständniserklärung der Patienten/Betreuer zur Aufzeichnung des Tests vorlag, die keinen Einfluss auf die weitere Teilnahme an der Studie hatte.</p> <p>Wie in Kapitel 4.3.1.2.1 beschrieben war die geplante primäre Analyse die Änderung der 6MWD von BL bis Woche 48 für den Vergleich der Ataluren-Gruppen vs. Placebo in der ITT-Population (mindestens eine der Ataluren-Gruppen ist Placebo signifikant überlegen). Post-hoc wurde eine so genannte kITT-Population definiert und ausgewertet, die in Übereinstimmung mit dem CHMP (2014) als methodisch begründet eingeschätzt wird (9).</p> <p>Zusätzlich waren Responder- und Progressor-Analysen und Analysen der Zeit bis zum Therapieeffekt bzw. einer Krankheitsprogression (10%ige Verbesserung oder Verschlechterung der 6MWD) im SAP spezifiziert (siehe Kapitel 4.3.1.2.1 und Anhang 4-E).</p> <p>Der 6MWT ist ein mittlerweile etablierter und validierter primärer Endpunkt in Studien bei DMD-Patienten, der als Maß der Krankheitsprogression angesehen wird und für den minimal klinisch relevante Änderungen ("minimal clinically important difference", MCID) bekannt sind</p>

(11–13), die auch seitens des CHMPs (2014) als klinisch bedeutsam eingeschätzt werden (9). Die statistische Auswertung ist in Kapitel 4.3.1.2.1 und Anhang 4-E beschrieben und Daten zur Validität und Patientenrelevanz des Endpunkts und der Analysen sind in Kapitel 4.4.2.3 dargestellt.

*Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.*

*Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.*

Tabelle 4-12: Bewertung des Verzerrungspotenzials für den Endpunkt 6MWD (inklusive Progressor-Analyse in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PTC124-GD-007-DMD	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt 6MWD wird unter Berücksichtigung der in Tabelle 4-12 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Das Verzerrungspotential auf Studienebene wurde ebenfalls als gering eingestuft. Da es sich um eine randomisierte, kontrollierte Doppelblindstudie handelte, lag eine Verblindung der Endpunkterheber vor. Die Prüfarzte, das medizinische Personal und die Patienten waren verblindet. Es fand eine adäquate Operationalisierung des Endpunkts statt. Der 6MWT wurde unter standardisierten Bedingungen nach einem Trainingsprogramm in den Studienzentren und unter Einsatz eines Manuals, in dem die Testbedingungen inklusive Angaben zur Validität eines Tests detailliert beschrieben waren, durchgeführt. Wie in Kapitel 4.4.2.3 beschrieben ist der 6MWT ein bei DMD-Patienten validierter und mittlerweile akzeptierter primärer Endpunkt in DMD-Studien (z.B. McDonald 2013a). Auch die Wahl des Stratifizierungskriteriums (<350 m vs. ≥350 m) und des Kriteriums zur Definition einer Krankheitsprogression (mindestens 10%ige anhaltende Verschlechterung der 6MWD) sind klinisch begründet (siehe Kapitel 4.4.2.3). Eine adäquate Umsetzung des ITT Prinzips fand statt. Alle Patienten, die mindestens eine Dosis der Prüfpräparate erhalten hatten, wurden in die ITT-Analyse eingeschlossen. Die Zahl der randomisierten Patienten entsprach der ITT-Population. Wie oben angesprochen wurden in der kITT-Analyse die invaliden BL-Daten von zwei Patienten, die kurz zuvor akute

Verletzungen der unteren Extremitäten erlitten hatten und inkorrekt Weise in die ITT-Population eingeschlossen worden waren, durch deren Screening-Werte ersetzt. Die unrealistisch niedrigen BL-Werte spiegeln nicht die tatsächlichen Fähigkeiten der beiden Patienten wider, so dass eine Verzerrung der protokollgemäßen Primäranalyse der 6MWD nicht angenommen werden muss. Dementsprechend wird die kITT-Analyse in der vorliegenden Bewertung als die klinisch und wissenschaftlich begründete (primäre) Analyse eingestuft und die geplante ITT-Primäranalyse als unterstützende Auswertung. Diese Vorgehensweise wird durch das CHMP (2014) bestätigt (9). Die Robustheit der Primäranalyse wurde durch verschiedene Sensitivitätsanalysen zusätzlich bestätigt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder das Vorhandensein von sonstigen Faktoren, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

*Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

Tabelle 4-13: Ergebnisse für <Endpunkt xxx> aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)
<Studie 1>	Nicht zutreffend

Da es sich bei dem gegenständlichen Arzneimittel um ein Arzneimittel zur Behandlung eines seltenen Leidens handelt, werden die Ergebnisse des Endpunkts 6MWD in Übereinstimmung mit den Vorgaben des G-BA in Kapitel 4.4.2 dargestellt.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

<Abbildung Meta-Analyse>

Abbildung 2: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Nicht zutreffend

### 4.3.1.3.2 Zeitlich gemessene Funktionstests – RCT

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

*Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-14: Operationalisierung von Endpunkt TFT („timed function tests“)

Studie	Operationalisierung
PTC124-GD-007-DMD	<p>Während der TFTs wurde die Zeit, die benötigt wurde, um aus der Rückenlage aufzustehen, 10 m zu gehen oder zu rennen und vier Treppenstufen von Standardgröße hinauf- oder hinabzusteigen, erfasst. Die Patienten hatten für die Tests maximal 30 sec Zeit und bei Patienten, die aufgrund von Krankheitsprogression nicht in der Lage waren, an den Tests teilzunehmen oder die Tests innerhalb von 30 sec abzuschließen, wurden ab diesem Zeitpunkt 30 sec dokumentiert. Bei Patienten, bei denen die Tests aus anderen Gründen fehlten (z.B. Knochenbrüche), sollten die Werte als fehlend eingegeben werden.</p> <p>Die Tests waren in einem Manual, das den Prüffzentren zur Verfügung gestellt wurde, detailliert beschrieben. Die Tests sollten pro Visite vorzugsweise nur einmal ausgeführt werden, ausgenommen im Falle eines Testfehlers oder der erste Versuch war nicht auswertbar. Die Tests sollten in folgender Reihenfolge stattfinden: 1) 10 m Gehen oder Rennen, 2) Aufstehen aus liegender Position, 3) 4 Treppenstufen hinauf und hinabsteigen. Ein gerader Flur mit einer Länge von mindestens 12 m, der mit einer Start- und Ziellinie (exakt 10 m) markiert war, musste zur Verfügung stehen. Der Zeitpunkt, zu dem der zweite Fuß die Ziellinie kreuzte sollte mittels Stoppuhr gemessen werden. Bei Patienten, deren Zustand sich während der Studie derart verschlechterte, dass sie lange Beinschienen oder die Unterstützung durch eine andere Person während des 10-m-Gehtests benötigten, sollte die Zeit nicht dokumentiert werden. Die Fähigkeiten in den TFTs wurden zusätzlich mittels Skalen bewertet.</p> <p>Die TFTs wurden in einem Random-Effekt-Modell für Ordinalzahlen in SAS (PROC NL-MIXED) analysiert. Die statistische Analyse ist in Kapitel 4.3.1.2.1 näher beschrieben.</p>

*Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.*

*Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.*

Tabelle 4-15: Bewertung des Verzerrungspotenzials für den Endpunkt „TFTs“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PTC124-GD-007-DMD	niedrig	ja	ja	ja	nein	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt TFTs wird unter Berücksichtigung der in Tabelle 4-15 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Wie in Kapitel 4.3.1.3.1 und 4.3.1.2.2 dargelegt wurde das Verzerrungspotential auf Studienebene als niedrig eingestuft, da die Endpunkte verblindet erhoben wurden, das ITT-Prinzip adäquat umgesetzt wurde und keine Hinweise auf eine ergebnisgesteuerte Berichterstattung vorliegen. Auch für die TFTs wird die kITT-Analyse als die primär relevante und klinisch sowie wissenschaftlich begründete Analyse eingestuft (siehe Kapitel 4.3.1.3.1). Die Endpunkte wurden adäquat operationalisiert, die Endpunkterheber wurden trainiert und ihnen wurde ein ausführliches Manual, in dem die Tests beschrieben waren, zur Verfügung gestellt. TFTs sind ein in DMD-Studien häufig eingesetztes und etabliertes Instrument.

*Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

Tabelle 4-16: Ergebnisse für &lt;Endpunkt xxx&gt; aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)
<Studie 1>	Nicht zutreffend

Da es sich bei dem gegenständlichen Arzneimittel um ein Arzneimittel zur Behandlung eines seltenen Leidens handelt, werden die Ergebnisse des Endpunkts TFTs in Übereinstimmung mit den Vorgaben des G-BA in Kapitel 4.4.2 dargestellt.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum*

*eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

<Abbildung Meta-Analyse>

Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Nicht zutreffend

#### **4.3.1.3.3 Unerwünschte Ereignisse – RCT**

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

*Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-17: Operationalisierung des Endpunkts „Unerwünschte Ereignisse“ (UEs)

Studie	Operationalisierung
PTC124-GD-007-DMD	<p>Ein UE konnte jedes nachteilige oder nicht beabsichtigte Zeichen (einschließlich abnormale Laborbefunde), Symptome oder Erkrankungen sein, die temporär mit der Anwendung eines medizinischen (investigativen) Produkts assoziiert waren, unabhängig davon, ob diese mit dem medizinischen (investigativen) Produkt in Zusammenhang standen oder nicht. Es wurden Kriterien spezifiziert, wie der Zusammenhang mit den Prüfpräparaten zu bewerten war (wahrscheinlich, möglich, unwahrscheinlich, kein Zusammenhang). Für regulatorische Zwecke wurde ein UE als mit der Behandlung in Zusammenhang stehend gewertet, wenn der Prüfarzt oder der Sponsor dieses Ereignis als möglicherweise oder wahrscheinlich behandlungsbedingt einstufte.</p> <p>UEs wurden entsprechend MedDRA klassifiziert. Die Schwere der UEs wurde nach CTAE, Version 3.0 sofern möglich, eingeschätzt. Für den Fall, dass ein Patient dasselbe UE mehrmals entwickelte wurde das UE nur einmal mit dem höchsten CTAE-Schweregrad dokumentiert. Ein behandlingsspezifisches UE war definiert als ein UE, welches im Zeitraum zwischen der ersten Dosis der Prüfmedikationen bis zu sechs Wochen nach der letzten Dosis der Prüfmedikationen auftrat (vorausgesetzt der Patient nahm nicht an der offenen Verlängerungsstudie teil). Alle UEs sollten bis sie abgeklungen waren oder die Situation chronisch oder stabil war beobachtet werden. Spezielle Aufmerksamkeit wurde CTCAE-UEs Grad 3 oder höher, mit den Prüfpräparaten in Zusammenhang stehende UEs, adrenale, hepatische und renale Ereignisse, die zu spezifischen diagnostischen Maßnahmen führten, sowie schwerwiegenden UEs gewidmet.</p> <p>Ein schwerwiegendes UE (SUE) war unabhängig vom Zusammenhang mit den Prüfpräparaten als ein Ereignis definiert, welches</p> <ul style="list-style-type: none"> <li>– den Tod zur Folge hatte (einschließlich sechs Wochen nach der letzten Dosis oder länger wenn das Ereignis bereits während der Berichtsperiode auftrat) inklusive Tod durch DMD - Krankheitsprogression.</li> <li>– lebensbedrohlich war. Die Definition schloss keine Ereignisse ein, die zum Tod hätten geführt haben können, wenn sie in einer schwereren Form aufgetreten wären.</li> <li>– eine Hospitalisierung erforderte oder diese verlängerte (exklusive Hospitalisierungen im Zusammenhang mit der Verabreichung der Prüfpräparate, der Studienprozeduren oder diagnostischen Maßnahmen [DMD]; andere geplante Hospitalisierungen; Hospitalisierungen allein aufgrund eines Fortschreitens der Erkrankung). Behandlungen in der Notfallstation wie z.B. eine Hydratation, die keine Klinikeinweisung erforderlich machten, oder ein Aufenthalt von weniger als 24 h in der Notfallstation wurden nicht als schwerwiegend bewertet.</li> <li>– zu dauerhafter oder signifikanter Behinderung/Invalidität führte und nicht mit der DMD in Zusammenhang stand.</li> <li>– sowie jedes andere medizinisch signifikante Ereignis, welches aus Sicht des Prüfarztes oder Sponsors oder der lokal zuständigen Behörde als schwerwiegend einzustufen war.</li> </ul> <p>Ein Ereignis, das ausschließlich eine Folge eines Rückfalls, einer erwarteten Änderung oder Progression der DMD war, sollte als UE und nicht als SUE berichtet werden.</p>

*Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.*

*Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.*

Tabelle 4-18: Bewertung des Verzerrungspotenzials für den Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PTC124-GD-007-DMD	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt UEs wird unter Berücksichtigung der in Tabelle 4-18 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft, die Studie war doppelblind und das ITT-Prinzip wurde adäquat umgesetzt. Der Endpunkt wurde angemessen operationalisiert. Es liegen keine sonstigen Hinweise auf eine Verzerrung vor, die die Ergebnisse relevant beeinflusst haben könnten.

*Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.*

Tabelle 4-19: Ergebnisse für den Endpunkt „Unerwünschte Ereignisse“ aus RCT mit dem zu bewertenden Arzneimittel

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung oben)
<Studie 1>	Nicht zutreffend

Da es sich bei dem gegenständlichen Arzneimittel um ein Arzneimittel zur Behandlung eines seltenen Leidens handelt, werden die Ergebnisse des Endpunkts Unerwünschte Ereignisse in Übereinstimmung mit den Vorgaben des G-BA in Kapitel 4.4.2 dargestellt.

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden.*



*Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

<Abbildung Meta-Analyse>

Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Nicht zutreffend

#### **4.3.1.3.4 Subgruppenanalysen – RCT**

Für die tabellarische Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die tabellarische Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.

*Beschreiben Sie die Ergebnisse von Subgruppenanalysen (einschließlich der Interaktionsterme). Stellen Sie dabei die Ergebnisse in den Subgruppen zunächst für jede einzelne Studie in tabellarischer Form dar. Diese Anforderung gilt sowohl für Subgruppenanalysen auf Basis individueller Patientendaten als auch für solche auf Basis aggregierter Daten. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Kennzeichnen Sie in einzelnen Studien a priori geplante Subgruppenanalysen.*

*Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.*

*Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.*

Die Ergebnisse der im Studienprotokoll spezifizierten Subgruppenanalysen sind in Übereinstimmung mit den Vorgaben des G-BA für Arzneimittel zur Behandlung seltener Leiden in Kapitel 4.4.2 dargestellt.

#### **4.3.1.3.5 Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien**

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen aus randomisierten kontrollierten Studien geben. Die Zusammenfassung soll Aussagen zu allen in Abschnitt 4.3.1.3 präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

*Fassen Sie die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen aus randomisierten kontrollierten Studien zusammen.*

Da es sich bei dem gegenständlichen Arzneimittel um ein Arzneimittel zur Behandlung eines seltenen Leidens handelt, werden die Ergebnisse aus RCTs in Übereinstimmung mit den Vorgaben des G-BA in Kapitel 4.4.2 dargestellt.

## 4.3.2 Weitere Unterlagen

### 4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

#### 4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie*

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht zutreffend

#### 4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

*Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche herangezogen wurden, und bewerten Sie deren Verzerrungspotenzial. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.*

Nicht zutreffend

### 4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-20: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
<Studie 1>	nein	ja	ja	ja	nein

#### 4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in 3 Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 4.3.1.3.1.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielhaft wäre folgende Darstellung denkbar:

Tabelle 4-21: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Intervention	<Vergleichs- therapie 1>	<Vergleichs- therapie 2>	<Vergleichs- therapie 3>
1	<Studie 1>	●		●	●
2	<Studie 2> <Studie 3>	● ●		● ●	
1	<Studie 4>		●	●	●
etc.	etc.	etc.	etc.		

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-22: Operationalisierung von &lt;Endpunkt xxx&gt;

Studie	Operationalisierung
<Studie 1>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-23: Bewertung des Verzerrungspotenzials für &lt;Endpunkt xxx&gt; in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
<Studie 1>	<hoch / niedrig>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>	<hoch / niedrig>

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-24: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1)
<Studie 1>	

Nicht zutreffend

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere Widersprüche zwischen direkter und indirekter Evidenz.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt, für den ein indirekter Vergleich vorgenommen wird, fortlaufend in einem eigenen Abschnitt dar.

#### **4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT**

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.4.

Nicht zutreffend

### 4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

#### 4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie*

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Die Studie PTC124-GD-007-DMD war für die Zulassung und die zugelassenen gebiete von Ataluren ausschlaggebend und lässt aufgrund ihres Designs Aussagen zum Ausmaß des Zusatznutzens von Ataluren zu. Auf die Darstellung der noch laufenden offenen Studien mit Ataluren 10/10/20 wird daher verzichtet.

#### 4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

*Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.*

*Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.*

*Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.*

Tabelle 4-25: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend

#### 4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

##### 4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-26: Operationalisierung von &lt;Endpunkt xxx&gt;

Studie	Operationalisierung
<Studie 1>	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-27: Verzerrungsaspekte für &lt;Endpunkt xxx&gt; – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
<Studie 1>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein / unklar>	<ja / nein>

*Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.*

Nicht zutreffend

*Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

Nicht zutreffend

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

#### **4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien**

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.4.***

Nicht zutreffend

#### **4.3.2.3 Weitere Untersuchungen**

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.



#### 4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie*

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht zutreffend

#### 4.3.2.3.2 Charakteristika der weiteren Untersuchungen

*Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.*

*Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.*

*Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.*

Nicht zutreffend

#### 4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

##### 4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

*Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-28: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
<Studie 1>	

*Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind,*

*gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.*

Nicht zutreffend

*Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.*

Nicht zutreffend

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

#### **4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen**

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.4.***

Nicht zutreffend

#### **4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen**

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 4.3.2.2 und 4.3.2.3) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

*Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.*

Nicht zutreffend

### **4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens**

#### **4.4.1 Beurteilung der Aussagekraft der Nachweise**

*Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.*

Bei der für die Nutzenbewertung herangezogenen Studie handelt es sich um eine randomisierte, Placebo-kontrollierte Doppelblindstudie der Evidenzstufe 1b. Da es sich bei DMD um ein sehr seltenes Leiden handelt, ist die Anzahl der in die Studie eingeschlossenen Patienten als akzeptabel einzustufen. Die *a priori* festgelegten Endpunkte wurden adäquat operationalisiert und das Verzerrungspotential wurde auf Studienebene und für die

bewertungsrelevanten Endpunkte (6MWD-Analysen und UEs) als niedrig eingestuft, so dass insgesamt von einer hohen Aussagekraft der Nachweise ausgegangen werden kann.

#### **4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß**

*Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie, worin der Zusatznutzen besteht.*

*Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.*

*Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):*

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

*Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.*

##### **4.4.2.1 Hintergrund und Fragestellung**

Translarna ist zur Behandlung der DMD aufgrund einer genetisch nachgewiesenen Nonsense-Mutation bei gehfähigen Patienten im Alter von fünf Jahren und älter zugelassen. Die DMD ist eine schwerwiegende X-chromosomal vererbte Erkrankung vorwiegend bei Jungen, die zu fortschreitendem Muskelabbau, assoziiert mit Muskelschwäche, Muskelkontrakturen und Skoliose, bis hin zur Rollstuhlabhängigkeit, Pflegebürtigkeit und Tod im zweiten bzw. dritten Lebensjahrzehnt führt (1,2,25). Mit Fortschreiten der Erkrankung werden die Muskelfasern zunehmend durch fibrotisches und Fettgewebe ersetzt (26). Nach longitudinalen Beobachtungsstudien scheint das Alter von 7 Jahren bei DMD-Patienten einen Wendepunkt darzustellen. Bis zum Alter von 7 Jahren nimmt das Gehvermögen von DMD-Patienten vermutlich aufgrund von Wachstum im Mittel zu oder bleibt stabil, während es bei älteren

Kindern ( $\geq 7$  Jahre) abnimmt (27). Mit ca. 15-16 Jahren sind die Kinder und Jugendlichen i.A. nicht mehr gehfähig. Da die Erkrankung auch die respiratorischen Muskeln betrifft, entwickeln die Patienten üblicherweise ab einem Alter über 15 Jahren eine respiratorische Insuffizienz (25). Ab einem Alter von 18 Jahren zeigt sich bei DMD-Patienten häufig eine Kardiomyopathie (26).

DMD wird durch verschiedene Mutationen in der DNA des Dystrophin-Gens verursacht, wodurch es zur Bildung nicht funktionalen Dystrophins kommt (2). Diese schließen Gendeletionen (meist Exon 45-53; 60% bis 80%), Duplikationen (7% bis 11%) und kleine Mutationen (10% bis 30%) ein. Bei einem Anteil der DMD-Patienten enthält die mRNA des Dystrophin-Gens ein zusätzliches Stop-Codon (etwa 11,5%), wodurch die Translation vorzeitig abgebrochen und ein verkürztes, nicht funktionales Protein gebildet wird (2,11). Durch Ataluren wird das vorzeitige Stop-Codon ignoriert, sodass die Translation bis zu dem normalen Stop-Codon fortschreiten und funktionales Dystrophin synthetisiert werden kann (2). Zur kausalen Behandlung der DMD steht bislang keine Therapie zur Verfügung. Die Behandlung der DMD umfasst heute symptomatische Maßnahmen zur Unterstützung der kardialen und respiratorischen Funktion sowie Kortikosteroide zur Verbesserung der Muskelstärke und -funktion. Orthopädische Korrekturen und physiotherapeutische Interventionen sind ebenfalls Teil der Therapie (2,25,26). Kortikosteroide sind in dieser Indikation in Europa nicht zugelassen. Sie haben z.T. signifikante Nebenwirkungen und hinsichtlich der/des optimalen Dosisregimes besteht kein Konsens (26). Translarna ist somit das erste zugelassene kausal wirksame Arzneimittel bei nmDMD.

Mit der Zulassung von Translarna (Ataluren) als Arzneimittel zur Behandlung eines seltenen Leidens („Orphan Drug“) gilt der Zusatznutzen nach § 35a Abs. 1 Satz 10 SGB V mit der Zulassung als belegt. Das Ausmaß des Zusatznutzens von Ataluren wird anhand der im 5. Kapitel § 5 Abs. 7 Nr. 1 bis 4 der VerfO festgelegten Kriterien bewertet (4). Demnach liegt ein beträchtlicher Zusatznutzen vor, wenn eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens im Sinne von § 3 Absatz 1 erreicht wird, insbesondere eine Abschwächung schwerwiegender Symptome, eine moderate Verlängerung der Lebensdauer, eine für die Patientinnen und Patienten spürbare Linderung der Erkrankung, eine relevante Vermeidung schwerwiegender Nebenwirkungen oder eine bedeutsame Vermeidung anderer Nebenwirkungen (4). Der Verlust der Gehfähigkeit ist als schwerwiegendes Symptom der DMD für die in Frage stehende Population anzusehen, bzw. führt deterministisch zu schwerwiegenden Folgekomplikationen.

Die Verfahrensordnung des G-BA nach § 35a SGB V definiert patientenrelevante Effekte in Paragraph 3 Absatz 1 als Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen oder Verbesserung der Lebensqualität (4). Für die Bewertung des Ausmaßes des Zusatznutzens von Ataluren wurden die patientenrelevante Zielgröße der Morbidität (Nutzendimensionen: Verbesserung des Gesundheitszustands [Mobilität] sowie Nebenwirkungen) herangezogen. Die Zielgröße Mortalität konnte in der vorliegenden 48-Wochen-Studie nicht evaluiert werden. Die

Lebensqualität wurde in der Studie erfasst (PedsQL); die Daten erlauben allerdings keine Aussagen hinsichtlich des Ausmaßes des Zusatznutzens von Ataluren, da eine *a priori* Validierung von geeigneten Messinstrumenten und Responsekriterien, insbesondere in dieser paediatrischen Population, an methodische Grenzen stößt. Auch ist eine deutliche Verschlechterung der Lebensqualität erst mit dem kompletten Verlust der Gehfähigkeit zu erwarten (siehe Modul 3.2.1), diese Phase wurde in der vorliegenden Studie nicht erreicht. Tabelle 4-29 gibt einen Überblick der in dieser Nutzenbewertung berücksichtigten patientenrelevanten Endpunkte. Der 6MWT ist ein Maß für die Ausdauer und funktionale Mobilität, wobei beide Aspekte unmittelbar mit der Gehfähigkeit der betroffenen Patienten zusammenhängen. Eine signifikante Veränderung dieses Outcome-Parameters stellt somit eine Abschwächung der mit der Erkrankung einhergehenden schwerwiegenden Symptome dar.

Tabelle 4-29: Darstellung der berücksichtigten patientenrelevanten Endpunkte

<b>Patientenrelevante Zielgrößenkategorie</b>	<b>Endpunktparameter</b>
Morbidity (Nutzendimension: Abschwächung schwerwiegender Krankheitssymptome bzw. Folgekomplikationen)	<ul style="list-style-type: none"> <li>• Gehstrecke im 6MWT bis Woche 48</li> <li>• Anteil der Patienten mit mindestens 10%iger Verschlechterung der Gehstrecke im 6MWT</li> <li>• Zeit bis zu einer mindestens 10%igen Verschlechterung der Gehstrecke im 6MWT</li> <li>• Anteil der Patienten mit mindestens 0.5 x SD Verschlechterung der Gehstrecke im 6MWT</li> <li>• TFTs: <ul style="list-style-type: none"> <li>- vier Treppenstufen hinauf zu gehen</li> </ul> </li> </ul>
Nebenwirkungen (Nutzendimension: Vermeidung von [schwerwiegenden] Nebenwirkungen)	- Nicht schwerwiegende (bzw. schwerwiegende) unerwünschte Arzneimittelwirkungen

Die nachfolgend und in Kapitel 4.4.3 dargestellte Beschreibung und Begründung des Ausmaßes des Zusatznutzens basiert auf den Ergebnissen der Zulassungsstudie PTC124-GD-007-DMD, die für die zugelassene Indikation und die Anwendungsbedingungen von Translarna (Ataluren) entscheidend war. Bei der Studie handelt es sich um eine Placebo-kontrollierte Doppelblindstudie der Evidenzstufe 1b, die eine Quantifizierung des Zusatznutzens nach heutigen wissenschaftlichen Standards erlaubt. Die Bewertung des Ausmaßes des Zusatznutzens von Ataluren erfolgte nach den Kriterien des §5 Absatz 7 des 5. Kapitels der VerfO (4). Bei der Quantifizierung des Ausmaßes des Zusatznutzens von Ataluren ist zu berücksichtigen, dass es sich um eine sehr schwerwiegende, progredient verlaufende Erkrankung handelt, für die bislang keine Referenztherapie zugelassen ist.

#### 4.4.2.2 Design der Studie PTC124-GD-007-DMD

Das Design und die statistische Auswertung der Zulassungsstudie PTC124-GD-007-DMD ist in Kapitel 4.3.1.2.1, das Verzerrungspotential auf Studien- bzw. Endpunktebene in Kapitel 4.3.1.3.1 bzw. 4.3.1.2.2 und den zugehörigen Anhängen (Anhang 4-E und Anhang 4-F) beschrieben. Die Operationalisierung der bewertungsrelevanten Endpunkte ist in Kapitel 4.3.1.3 dargestellt.

Es handelt sich um eine randomisierte, Placebo-kontrollierte, dreiarmlige, doppelblinde Studie im Parallelgruppendesign über 48 Wochen. Eingeschlossen wurden männliche gehfähige Patienten mit genetisch gesicherter DMD im Alter von 5 Jahren und älter (siehe Anhang 4-F für eine Beschreibung der Einschluss- bzw. Ausschlusskriterien). Die Patienten mussten in der Lage sein, zur Screening-Untersuchung mindestens 75 m ohne Hilfsmittel zu gehen. Die Patienten wurden randomisiert der oralen Behandlung mit Placebo oder Ataluren (10/10/20 oder 20/20/40 mg/kg KG pro Tag) dreimal täglich zugeordnet. Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft (siehe Kapitel 4.3.1.3.1 und Anhang 4-F).

#### 4.4.2.3 Validität und Patientenrelevanz der Endpunkte

*Gehstrecke im 6MWT inklusive Anteil der Patienten mit mindestens 10%iger Verschlechterung*

Bei Patienten mit DMD ist die funktionale Mobilität der bedeutendste Endpunkt und der Verlust der Gehfähigkeit stellt für die Patienten ohne Zweifel einen relevanten Wendepunkt im Verlauf der Erkrankung dar (13,26). Der 6-MWT ist ein bei verschiedenen Erkrankungen gut etabliertes und robustes Instrument, welches die für DMD-Patienten relevanten Aspekte der Muskelfunktion und Ausdauer erfasst (11,13). Bei gehfähigen DMD-Patienten wird die Gehfähigkeit als der bedeutendste Endpunkt angesehen und der 6MWT wird weltweit in klinischen Studien als primärer Endpunkt sowie in longitudinalen Studien zum natürlichen Verlauf bei gehfähigen DMD-Patienten eingesetzt (12,13,25,27).

Da der molekulare Defekt der DMD direkt die Muskulatur betrifft (Skelettmuskulatur, Atemmuskulatur, Herzmuskel, etc.), stellt der 6-MWT auch ein direktes Maß für die Schwere der Erkrankung bzw. für die deterministische Krankheitsprogression dar. Die verminderte Gehfähigkeit der DMD Patienten ist eines der frühzeitigst auftretenden Symptome und als sensitivstes Maß der Krankheitsprogression zu werten. Die Krankheitsprogression in der DMD führt zu einer schwerwiegenden Erkrankung (§ 2 der Richtlinie des Gemeinsamen Bundesausschusses zur Umsetzung der Regelungen in § 62 für schwerwiegend chronisch Erkrankte, 20. August 2008). Der Grad der Schädigungsfolgen (GdS) der DMD führt bis hin zu schweren Auswirkungen (Grad 90-100) von Muskelkrankheiten (18.6, Teil B der Versorgungsmedizin-Verordnung vom 10. Dezember 2008 (BGBl. I S. 2412), die zuletzt durch Artikel 1 der Verordnung vom 11. Oktober 2012 (BGBl. I S. 2122) geändert worden ist).

Eine auf 50-55% des prognostizierten Wertes verkürzte Gehstrecke, sowie eine um 30 Meter oder 10% reduzierte Gehstrecke stellen für DMD Patienten Marker für den weiteren raschen

deterministischen, und irreversiblen Verlauf der Erkrankung dar (12,27). Eine therapeutische Intervention, die den krankheitsbedingten Defekt der betroffenen Muskulatur adressiert, und zur Verhinderung/Verlangsamung der Abnahme der Gehfähigkeit führt, stellt somit eine Verhinderung/Verlangsamung der durch die Krankheitsprogression entstehenden schwerwiegenden Symptome (bzw. Folgekomplikationen) dar. Aus diesem Grund ist eine therapeutisch relevante Verbesserung in der 6-MWD für diese Patientenpopulation als beträchtlicher Zusatznutzen einzustufen.

Das "National Institute of Health" stuft den 6MWT als globales Instrument der Gehfähigkeit und Ausdauer ein (28). Der 6MWT wurde für den Einsatz bei ambulanten Jungen mit DMD modifiziert und hat sich als gut durchführbar, sicher und zuverlässig erwiesen (11,13). In einer Online-Umfrage wurde untersucht, welche Eigenschaften einer Therapie für Betreuer von DMD-Patienten besonders wichtig sind. Unter den vorgeschlagenen Attributen war der Erhalt der Muskelfunktion oder ein Aufhalten der fortschreitenden Muskelschwäche für die Befragten von größter Bedeutung und wurde höher bewertet als eine Verlängerung der Lebenserwartung. Die Befragten waren bereit, bei einer (nicht kurativen) Therapie, die das Fortschreiten der Muskelschwäche aufhält oder verlangsamt, selbst ein schwerwiegendes Risiko zu akzeptieren (29).

McDonald et al. (2013) analysierten die Screening- und BL Daten im Abstand von 42 d (Median, Bereich 0-91 d) der Studie PTC124-GD-007-DMD, um die Validität und minimal klinisch relevanten Änderungen („minimal clinically important difference“, MCID) des 6MWT und anderer klinischer Endpunkte zu ermitteln. Die 6MWD zeigte eine hohe Reliabilität nach wiederholter Testung (Screening vs. BL) mit einem Pearson r und ICC von 0.92. Die MCID wurde mittels verteilungsbasierter Methoden auf Basis der SEM- ( $SD\sqrt{(1-r)}$ ) und 1/3 SD-Methode bestimmt. Die geschätzten MCIDs betragen respektive 28.5 m bzw. 31.7 m im 6MWT bei DMD-Patienten und entsprachen 8.0% bzw. 8.9% der mittleren BL 6MWD (11). Eine Gehstreckenveränderung von 30 m wird für diese Patientenpopulation auch seitens des CHMPs (2014) als klinisch bedeutsam eingeschätzt (9).

Auch longitudinale Daten zum natürlichen Verlauf bei DMD, die darauf hinweisen, dass jede Reduktion der BL 6MWD um 30 m das Risiko, die Gehfähigkeit innerhalb von zwei Jahren zu verlieren, erhöht, stützen dieses Kriterium (siehe Abbildung 5) (30).

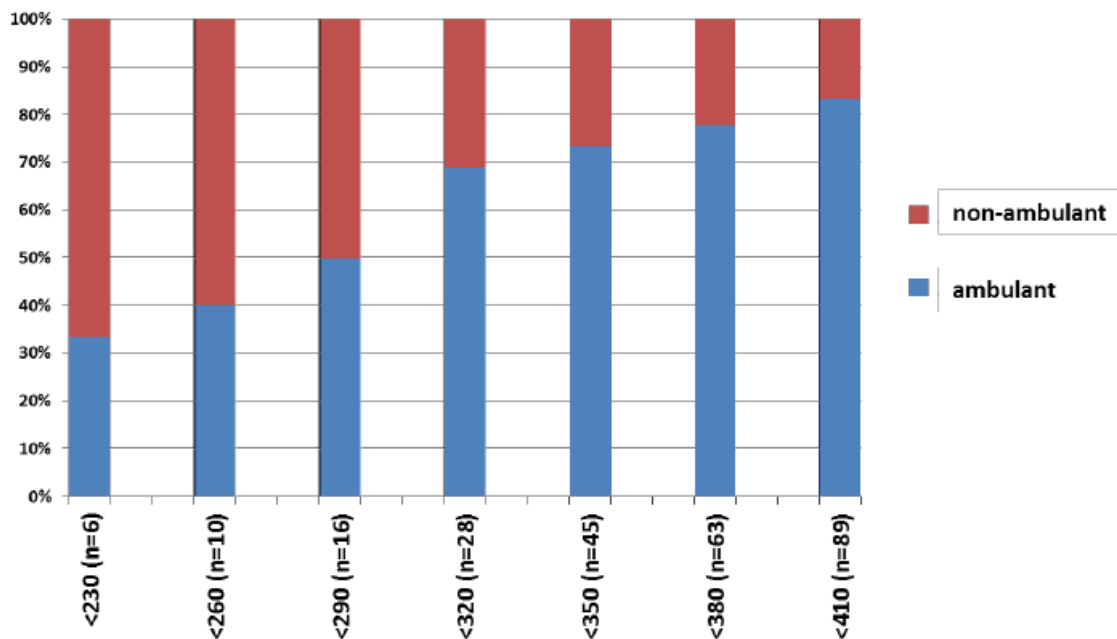


Abbildung 5: Anteil der Jungen mit DMD stratifiziert nach BL 6MWD (30-m-Intervalle), die über zwei Jahre ihre Gehfähigkeit verloren (Mercuri et al. zitiert nach (30))

Basierend auf Daten zum natürlichen Verlauf der DMD haben Patienten, deren Gehstrecke innerhalb von 12 Monaten um 10% oder mehr abnimmt, ein signifikant höheres Risiko, innerhalb von vier Jahren ihre Gehfähigkeit zu verlieren (12). Allerdings scheint die MCID bei DMD-Patienten bei Betrachtung auf individueller Patientenebene in Abhängigkeit vom Alter und der Ausgangs-Gehfähigkeit zu variieren (25). Henricson et al. (2013) schlossen 20 gehfähige Jungen im Alter von 4-12 Jahren mit DMD in eine Studie ein, um die MCID von 30 m im 6MWT in Abhängigkeit von der initialen Gehfähigkeit zu analysieren (13). Als Bezugspunkt (PRO, „person-reported outcome“) wurde die Domäne „Transfer und Basismobilität“ im PODCI („POSNA Pediatric Outcomes Data Collection Instrument“) gewählt, welche von den Eltern/Stellvertretern berichtet wurde. Auf Basis der BL-Werte des PODCI-Scores wurde eine MCID von 4.5 Punkten kalkuliert ( $1/3$  SD) und ermittelt, welche Änderungen im 6MWD in Abhängigkeit vom PODCI-Score erforderlich sind, um innerhalb eines Jahres eine klinisch bedeutsame Änderung des PODCI-Scores zu erreichen. Die Änderung der therapeutisch relevanten 6MWD war abhängig von der Ausgangsmobilität für konstante klinisch relevante Änderungen im PODCI (Tabelle 4-30). Beispielsweise ging eine klinisch bedeutsame Verbesserung des PODCI-Scores (4.5 m) von einem niedrigen PODCI Bereich von 30 auf 34.5 mit einer Änderung von nur 5.6 m im 6MWD einher. Ausgehend von einem PODCI-Score von 90 Punkten war eine Änderung um 4.5 Punkte demgegenüber mit einer Zunahme der relevanten 6MWD um fast 46 m assoziiert. Wie aus Tabelle 4-30 ersichtlich betrug die MCID für die 6MWD unabhängig von der Ausgangsgehstrecke (153-454 m stratifiziert nach PODCI-Bereichen) jeweils etwa 10% oder weniger des 6MWD-Ausgangswertes ( $\leq 10.1\%$ ). Dies rechtfertigt die klinische Relevanz der Analyse des Anteils der Patienten mit 10%iger Verschlechterung im 6MWT bis Woche 48 in der Ataluren-Studie.



Tabelle 4-30: Änderung des 6MWT stratifiziert nach PODCI-Ausgangswert, die bei Jungen mit DMD mit einer bedeutsamen Änderung des PODCI um 4.5 Punkte (MCID, Transfer und basale Mobilität) innerhalb eines Jahres einhergehen (13)

<b>PODCI Bereich (Punkte)</b>	<b>6MWT Bereich (m)</b>	<b>6MWT Mittelwert (m)</b>	<b>Relevante Änderung der 6MWD (m)</b>	<b>Relevante Änderung der 6MWD (%)</b>
30-34.5	150.3-155.9	153,10	5.6	3,66%
40-44.5	165.0-174.7	169,85	9.7	5,71%
50-54.5	189.5-202.2	195,85	12.7	6,48%
60-64.5	225.9-246.8	236,35	20.9	8,84%
70-74.5	276.7-304.9	290,80	28.2	9,70%
80-84.5	344.3-380.8	362,55	36.5	10,07%
90-94.5	431.1-477.0	454,05	45.9	10,11%

Unabhängig von der von Henricson (2013) durchgeführten Ableitung eines MCID für die Krankheitsverschlechterung wurde *post-hoc* eine weitere etablierte Methode eingesetzt, um die therapeutische Relevanz der Therapieunterschiede zwischen Ataluren 10/10/20 und Vergleichstherapie zu adressieren. In diesem Ansatz wird ein empirisch gut bestätigter minimal relevanter Therapieunterschied von 0.5 x SD (Baseline) als MCID auf der individuellen Patientenebene eingesetzt (18).

#### *Unterschiedliche Phasen im Krankheitsverlauf*

Eine Analyse der Placebo-Gruppe der Ataluren-Studie lieferte Daten zum natürlichen Verlauf der DMD, die zeigen, dass das Alter (Abbildung 7), sowie die Ausgangsfunktion der Gehfähigkeit Prädiktoren für den aktuellen Stand der Krankheitsprogression darstellen. Bis zu einem Alter von 7 Jahren blieb die Gehstrecke im 6MWT über 48 Wochen tendenziell stabil oder stieg sogar an, allerdings in der überwiegenden Mehrzahl nur bei Patienten mit besserer Ausgangsfunktion (>350 m im 6MWT oder >60% des vorhergesagten Wertes) (12). Bei den Placebo-Patienten der Ataluren-Studie blieb die 6MWD bei Patienten, die initial 350 m oder mehr gehen konnten weitgehend stabil, während die Gehstrecke bei Patienten, die initial weniger als 350 m im 6MWT gehen konnten, innerhalb eines Jahres deutlich abfiel (Abbildung 8) (31). In einer Longitudinal-Studie bei DMD-Patienten erwies sich eine initiale Gehstrecke im 6MWT von 330 m als der beste Prädiktor für den Erhalt der Gehfähigkeit und nur Patienten älter als 7 Jahre verloren innerhalb von zwei Jahren ihre Gehfähigkeit (27). Weiters zeigt die Longitudinalstudie von Pane (2014), dass sich die Kinetik der Gehstreckenverluste in Abhängigkeit vom Baseline Wert (<350 m versus > 350 m) in den ersten beiden Jahren deutlich und signifikant unterscheidet. Auf Grund des deterministischen Krankheitsverlaufes konvergieren die Raten erst nach etwa 3 Jahren (siehe Abbildung 6) (32).

Der Therapieeffekt von Ataluren (gemessen mittels des 6MWT) sollte sich daher für Patienten mit BL-Werten von <350 m stärker zeigen, als bei Patienten mit BL Werten >350 m.

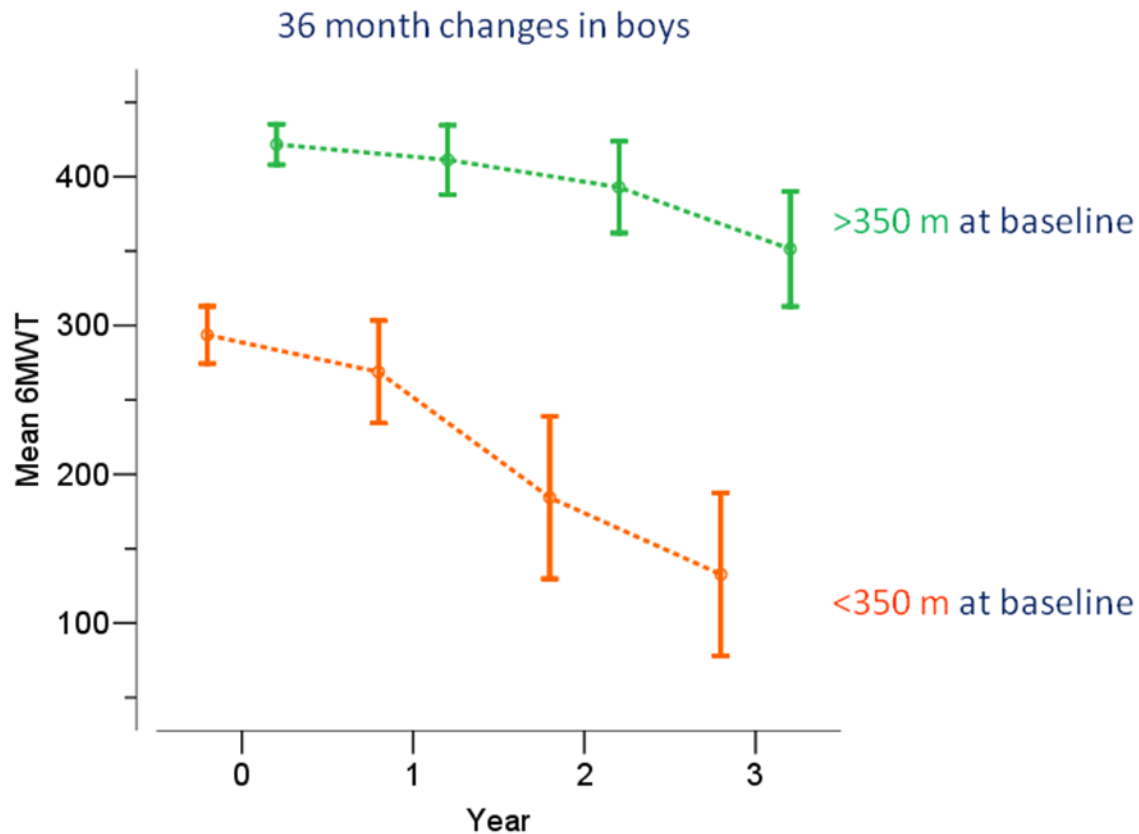


Abbildung 6: Verlauf der 6MWD in Abhängigkeit von BL-Werten

6MWD von DMD Patienten zu Baseline und nach 1-3 Jahren für BL-Werte über 350 Meter (grün) und unter 350 Meter (orange)

Diese Erkenntnisse zeigen deutlich, dass es schwierig ist, in einer klinischen Studie mit einer heterogenen Patientenpopulation (was das Stadium der Krankheitsprogression betrifft), den wahren therapeutischen Effekt von Ataluren mittels des 6MWT *effizient* für alle Patienten zu bestimmen. Da sich in der Phase 2b Studie nicht alle Patienten gleichzeitig in der „Decline-Phase“ befanden, in welcher der therapeutische Effekt auf die Gehfähigkeit am effizientesten gemessen werden kann, wird in dieser Studie der therapeutische Effekt einer früh begonnenen Langzeitbehandlung wahrscheinlich deutlich unterschätzt (siehe Abbildung 7 und Abbildung 8). Aus diesem Grund ist es gerechtfertigt, den Therapieeffekt von Ataluren in der Patientenpopulation mit BL-Werten von <350 m zu untersuchen (siehe Kapitel 4.4.2.4.2.4).

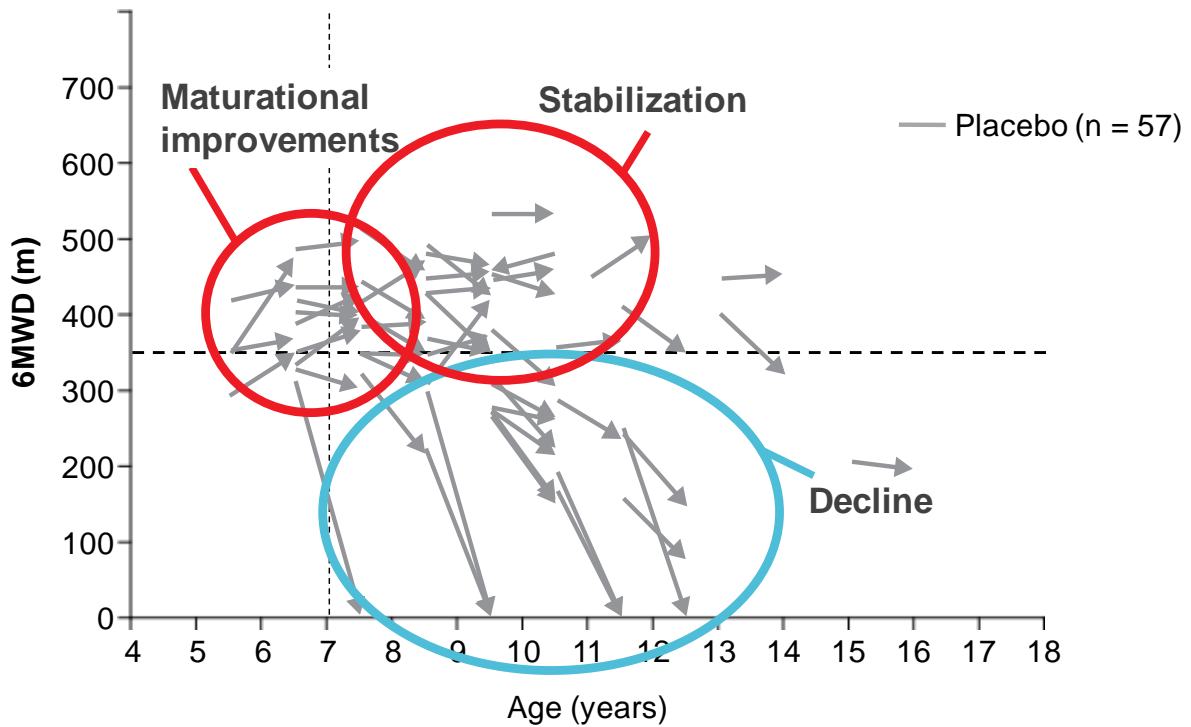


Abbildung 7: Änderung der 6MWD über 48 Wochen bei 57 DMD-Patienten in der Placebo-Gruppe der Phase 2b Ataluren-Studie (12)

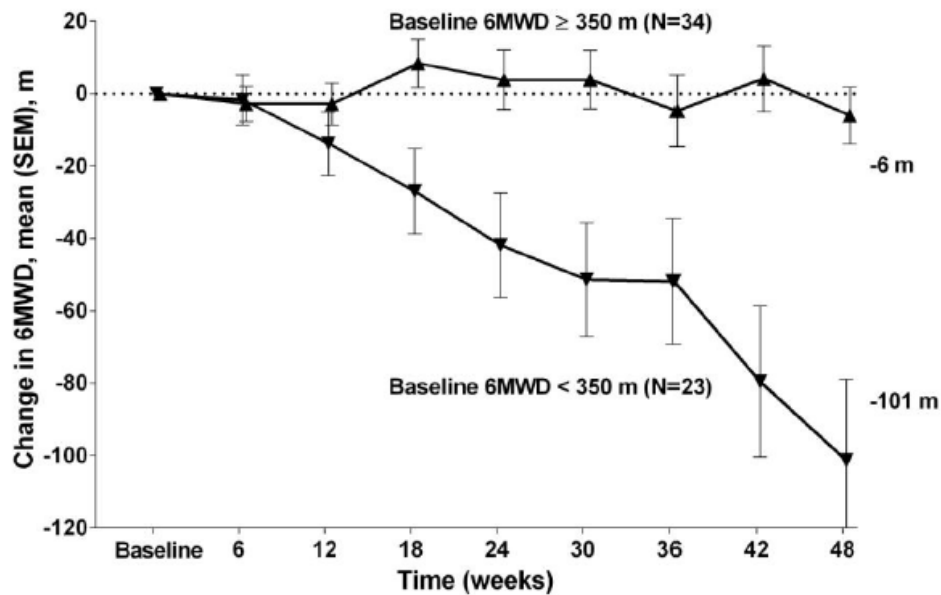


Abbildung 8: Mittlere Änderung der 6MWD bis Woche 48 stratifiziert nach BL 6MWD (<350 vs. ≥350 m) in der Placebo-Gruppe der Phase 2b Ataluren-Studie (31)

*TFTs ("timed function tests")*

Neben dem Verlust der Gehfähigkeit, stuft eine internationale DMD-Expertengruppe Fähigkeiten wie Treppenstufen zu ersteigen, unabhängig zu gehen oder selbst essen zu können als Patienten-orientierte „Meilensteine“ in der Krankheitsprogression ein, die für Patienten und deren Familien von unmittelbarer klinischer Bedeutung sind (33). TFTs (Zeit, die benötigt wird, um 10 m zu gehen oder zu rennen oder 4 Standardtreppen hinauf- oder hinabzusteigen) erfassen funktionale Fähigkeiten bei gehfähigen Patienten, waren bei DMD-Patienten reproduzierbar und sprachen auf die Behandlung mit Steroiden an (11). Der 10-m-Gehtest ist ein akzeptiertes und häufig zur Überprüfung der Krankheitsprogression bei DMD eingesetztes Instrument, mit welchem kurzfristig die physische Spitzenaktivität nicht aber die Ausdauer gemessen werden kann (11). Bei Steroid-naiven DMD-Patienten erwies sich die Zeit, die benötigt wurde, um 10 m zu gehen oder zu rennen, als prädiktiv für das Fortschreiten der Behinderung. Patienten, die mehr als 12 sec benötigen, um 10 m zu gehen oder zu rennen, hatten ein hohes Risiko, ihre Gehfähigkeit innerhalb von 12 Monaten zu verlieren, während es bei Patienten, die 10 m innerhalb von weniger als 6 sec bewältigten, wahrscheinlicher war, dass sie gehfähig blieben ( $p < 0.001$ ) (34). In einer longitudinalen Beobachtungsstudie in Italien hatten DMD-Patienten im Alter von 4.1-17 Jahren, die mehr als 7 sec benötigten, um 10 m zu gehen/rennen, ein signifikant erhöhtes Risiko, innerhalb von zwei Jahren rollstuhlpflichtig zu werden (OR 7.9, 95% KI 2.2-28.3) (26).

In einer Beobachtungsstudie bei 52 gehfähigen DMD-Patienten (4-15 Jahre) zeigten sich Assoziationen zwischen den Ergebnissen der TFTs und der physischen Funktion in zwei Instrumenten zur gesundheitsbezogenen Lebensqualität (PODCI und weniger ausgeprägt PedsQL berichtet durch die Eltern/Stellvertreter), was die klinische Bedeutung der TFTs als Endpunkte unterstützt (34). In einer Analyse der Placebo-Gruppe der Ataluren-Studie verloren 40% der Patienten ( $n=6/15$ ), die zu BL nicht in der Lage waren, aus dem Liegen aufzustehen, innerhalb von 48 Monaten ihre Gehfähigkeit verglichen mit 2.3% der Patienten ( $n=1/42$ ), die dies konnten (12). Auch die Fähigkeit, vier Treppenstufen zu gehen ( $< 8$  sec,  $\geq 8$  sec oder Unfähigkeit) hat sich bei DMD-Patienten als prädiktiv für den Verlust der Gehfähigkeit erwiesen. Etwa 40% derjenigen Patienten, die dies nicht konnten, verloren innerhalb von 12 Monaten ihre Gehfähigkeit (34).

In einer randomisierten Doppelblindstudie bei Jungen mit DMD im Alter von 4-10 Jahren wurden Äquivalenzgrenzen von  $\pm 0.4$  log sec (entsprechend 1.5 sec) für die Änderung der TFTs (aufstehen, 10 m gehen, 4 Stufen hinauf) innerhalb von 12 Monaten ohne nähere Begründung festgelegt (14). Henricson et al. (2013) kalkulierten bei 20 gehfähigen Jungen mit DMD eine MCID von 0.19 m/sec (1/3 SD) im Test, 10 m zu gehen oder zu rennen (13).

*Andere sekundäre Endpunkte*

Das CHMP (2014) stuft die Häufigkeit der durch die Patienten/Betreuer berichteten unbeabsichtigten Stürze als relevanten Endpunkt ein, da diese mit Frakturen der Gliedmaßen assoziiert sind, welche die Aktivität weiter vermindern und den Verlust der Gehfähigkeit beschleunigen können. Zwischen der Zahl der Stürze und der 6MWD ergab sich keine hohe

Korrelation, so dass die Stürze wahrscheinlich eine andere Aktivität erfassen als der 6MWT (9).

Das während der zentralen Zulassung von Translarna konsultierte Expertengremium aus Neurologen kam zu dem Schluss, dass die meisten der erhobenen Sekundärparameter primär eine Änderung der Funktion der Muskelstärke erfassten. Basierend auf aktuellen Daten führen minimale Zunahmen der Dystrophin-Produktion zu funktionalen Verbesserungen, um die Muskelstärke zu verbessern sind allerdings nahezu normale Dystrophin-Spiegel in den Muskelfasern erforderlich (9). Hinsichtlich der Muskelstärke (Streckung des Knies) wurden MCIDs von 2.1 bzw. 2.4 Pfund berichtet (SE und 1/3 SD Methode) (12). In einer Studie bei Jungen mit DMD im Alter zwischen vier und 10 Jahren wurde eine Äquivalenzgrenze von  $\pm 2$  Pfund für die Muskelfunktionstests gewählt (etwa 1 SD der Ausgangsverteilung) (14). In der Studie PTC124-GD-007-DMD wurde die Muskelstärke mittels Myometrie erfasst. Bei den Placebo-Patienten der Ataluren-Studie nahm die Muskelstärke über 48 Wochen nur minimal ab und laut McDonald ist die Muskelstärke kein angemessener Endpunkt für Therapeutika, die funktionale Verluste stabilisieren ohne die Stärke zu ändern (z.B. Restaurierung von Dystrophin) (34).

#### *Unerwünschte Ereignisse*

Nebenwirkungen stellen einen Aspekt der therapiebedingten Morbidität dar und sind daher unmittelbar patientenrelevant.

Folgende UEs sind in der Fachinformation (SmPC) von Translarna (Ataluren) gelistet (35):

<b>SOK</b>	<b>Sehr häufig (<math>\geq 1/10</math>)</b>	<b>Häufig (<math>\geq 1/100</math> to <math>&lt; 1/10</math>)</b>	<b>Häufigkeit unbekannt</b>
Stoffwechsel und Ernährungsstörungen		Verminderter Appetit	Änderung des Lipidprofils (erhöhte Triglyceride und Cholesterin)
Erkrankungen des Nervensystems	Kopfschmerzen	Schwindel	
Vaskuläre Erkrankungen		Hypertension	
Respiratorische, thorakale und mediastinale Erkrankungen		Husten, Epistaxis	
Gastrointestinale Erkrankungen	Übelkeit, Erbrechen	Oberer Abdominalschmerz, Flatulenz, Diarrhoe, Magenbeschwerden, Abdominalschmerz, Konstipation, Regurgitation	
Erkrankungen der Haut und des Unterhautzellgewebes		Erythem	
Skellett Muskulatur- und Bindegeweberkrankungen		Schmerz in den Extremitäten	

SOK	Sehr häufig ( $\geq 1/10$ )	Häufig ( $\geq 1/100$ to $< 1/10$ )	Häufigkeit unbekannt
Erkrankungen der Niere und der Harnwege		Bettnässen, Nierenzyste, Pollakisurie, Farbe des Urin abnormal	Änderung von Nierenfunktionstests (erhöhtes Kreatinin, Blut-Harnstoff-Stickstoff, Cystatin C)
Allgemeine Erkrankungen und Erkrankungen am Applikationsort		Fieber, Fatigue, Gewicht reduziert	

Folgende Nebenwirkungen sind in der SmPC von Translarna spezifiziert (35):

#### *Serumlipide*

Während der kontrollierten Studie waren Gesamtcholesterin und Triglyzeride zu BL im Mittel im Normbereich und stiegen auf grenzwertig hohe oder hohe Werte an. Es ergab sich während der Studie ein Trend hin zu einer frühen Stabilisierung der Werte und sie stiegen unter fortgesetzter Behandlung nicht weiter an.

#### *Nierenfunktionstests*

Während der kontrollierten Studie bei nmDMD wurde kleine Anstiege des mittleren Serumkreatinins, von Blut-Harnstoff-Stickstoff und Cystatin C beobachtet. Es ergab sich ein Trend hin zu einer frühen Stabilisierung der Werte und sie stiegen unter fortgesetzter Behandlung nicht weiter an.

### **4.4.2.4 Ergebnisse der Studie PTC124-GD-007-DMD – RCT**

Wenn nicht anders angegeben beziehen sich die nachfolgend dargestellten Daten auf den Studienbericht (8).

#### **4.4.2.4.1 Disposition der Patienten und Patientencharakteristika**

Insgesamt 185 Patienten wurden auf Eignung überprüft, von denen 11 Patienten nicht den Einschlusskriterien entsprachen. Insgesamt 174 Patienten wurden randomisiert der Behandlung mit Placebo (n=57), Ataluren 10/10/20 mg/kg/d (n=57) und 20/20/40 mg/kg/d (n=60) zugeteilt. Ein Patient der Ataluren 20/20/40-Gruppe brach die Studie aufgrund fehlender Compliance in Woche 6 vorzeitig ab. Die verbleibenden 173 Patienten beendeten die Studie. Alle randomisierten Patienten (n=174) wurden in die Wirksamkeitsanalyse eingeschlossen. Die PP-Population (alle Patienten, die die Studie abschlossen und einen validen 6MWD-Wert in Woche 48 hatten) bestand aus 169 von 174 Patienten (97%). Die Patienten wurden in 37 Zentren in 11 Ländern rekrutiert (siehe Tabelle 4-6). Die überwiegende Mehrzahl der Patienten wurden in den USA (n=87/174) und in europäischen Ländern rekrutiert, deren Versorgungskontext demjenigen in Deutschland ähnlich sein dürfte. Insgesamt 90.2% der eingeschlossenen Patienten waren kaukasischer Herkunft, was der Situation in Deutschland etwa entspricht.

Hinsichtlich der vordefinierten Stratifizierungsfaktoren (<9 Jahre vs.  $\geq 9$  Jahre, Kortikosteroid-Anwendung ja vs. nein, 6MWD  $\geq 350$  m vs. <350 m) war die Gruppenverteilung

ausgeglichen. Die Mehrzahl der Patienten war jünger als 9 Jahre (etwa 56%), wendete Kortikosteroide an (etwa 70%) und war in der Lage, im 6MWT mindestens 350 m ohne Hilfen zu gehen (im Mittel 55% bis 60%). Das Alter der Patienten schwankte zwischen fünf und 20 Jahren und die Gruppen waren hinsichtlich des Alters, der Größe, des Körpergewichtes und des BMIs ausgeglichen (siehe Tabelle 4-32). Insgesamt sechs Zwillingspaare nahmen an der Studie teil. Die Gruppen waren hinsichtlich der Krankheitscharakteristika (Alter bei Diagnosestellung und Phänotyp) vergleichbar und die mittlere 6MWD zu BL unterschied sich um <10 m im Gruppenvergleich. Hinsichtlich des Typs des Stop-Codons im Dystrophingen dominierte der Typ UGA und die Nonsense-Mutationen waren über die 79 Exons des Dystrophingens gleichmäßig verteilt (siehe Tabelle 4-32).

Hinsichtlich der Kortikosteroid-Anwendung (Art, Häufigkeit und Dauer der Vorbehandlung), die Einfluss auf die Wirksamkeitsanalyse haben könnte, ergaben sich im Gruppenvergleich keine Unterschiede (Tabelle 4-31). Bei nur einem Patienten (Ataluren 20/20/40) wurden Kortikosteroide während der Studie eingeleitet und keiner der vorbehandelten Patienten brach die Kortikosteroid-Therapie ab. Nur etwa 12% der Patienten nahmen zu Studienbeginn Herz-Kreislauf-Medikamente ein und nur vier Patienten begannen während der Studie eine Therapie mit kardialen Medikationen. Im Gruppenvergleich ergaben sich keine wesentlichen Unterschiede hinsichtlich der Anwendung nicht-arzneilicher Therapie-Maßnahmen wie Physio-, Hydro-, Psycho- oder Sprachtherapie, Einsatz von Orthesen oder anderer Hilfsmittel oder eines Trainingsprogramms.

Tabelle 4-31: Kortikosteroidanwendung zum Zeitpunkt der Randomisierung („as treated“-Population)

<b>Kortikosteroidtherapie, n (%)</b>	<b>Placebo (n=57)</b>	<b>Ataluren 10/10/20 (n=57)</b>	<b>Ataluren 20/20/40 (n=60)</b>
Kortikosteroidanwendung <sup>1</sup>	40 (70.2)	41 (71.9)	43 (71.7)
Deflazacort	17 (29.8)	17 (29.8)	20 (33.3)
Täglich	14 (24.6)	16 (28.1)	16 (26.7)
Jeden zweiten Tag	0 (0.0)	1 (1.8)	2 (3.3)
Andere	3 (5.3)	0 (0.0)	2 (3.3)
Prednisolon	11 (19.3)	14 (24.6)	9 (15.0)
Täglich	10 (17.5)	11 (19.3)	7 (11.7)
Jeden zweiten Tag	0 (0.0)	1 (1.8)	1 (1.7)
Andere	1 (1.8)	2 (3.5)	1 (1.7)
Prednison	12 (21.1)	10 (17.5)	14 (23.3)
Täglich	6 (10.5)	6 (10.5)	6 (10.0)
Jeden zweiten Tag	1 (1.8)	1 (1.8)	0 (0.0)
Andere	5 (8.8)	3 (5.3)	8 (13.3)

Quelle (Studienreport PTC124-GD-007-DMD): S. 84

<sup>1</sup> Bei den Patienten, die die Kortikosteroide täglich einnahmen, lag die Dosis zwischen 7.5 und 33 mg für Deflazacort, 10-30 mg für Prednisolon und 10-35 mg für Prednison.

In nachfolgender Tabelle 4-32 sind die Charakteristika der Studienpopulation dargestellt.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-32: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Alter (Jahre) M (±SD) [Bereich]	Geschlecht w/m (%)	Kauka- sier n (%)	Körper- größe (cm) M (±SD) [Bereich]	Körper- gewicht (kg) M (±SD) [Bereich]	BMI (kg/m <sup>2</sup> ) M (±SD) [Bereich]	Alter bei Diagnose- stellung M (±SD) [Bereich]	Phänotyp <sup>1</sup> n (%)					
									A	B	C	D	E	
PTC124-GD-007-DMD														
Placebo	57	8.3 (2.33) 5-15	0/57 (0%/100%)	54 (94.7%)	123.4 (11.8) [104-163]	28.6 (9.1) [16-55]	18 (3.7) [13-29]	3.9 (2.3) [0-10]	50 (88)	44 (77)	47 (83)	56 (98)	12 (21)	
ATL 10/10/20	57	8.8 (2.91) 5-20	0/57 (0%/100%)	53 (93.0%)	124.5 (15.3) [99-173]	31.2 (12.1) [16-76]	19 (3.5) [14-31]	3.3 (1.8) [0-9]	49 (86)	43 (75)	50 (88)	48 (84)	19 (33)	
ATL 20/20/40	60	8.4 (2.53) 5-16	0/60 (0%/100%)	50 (83.3%)	126.2 (13.8) [99-173]	31.9 (12.8) [17-84]	19 (4.8) [14-41]	3.8 (2.0) [0-8]	52 (87)	49 (82)	51 (85)	55 (92)	13 (22)	

Quelle (Studienreport PTC124-GD-007-DMD): S. 81, 82

<sup>1</sup> A: Proximale Muskelschwäche, B: schwankender Gang, C: Gowers Manöver, D: Wadenhypertrophie, E: Andere wie Zehengang, Muskelkrämpfe, Lordose oder Entwicklungsverzögerung



Tabelle 4-32 (fortgesetzt)

Studie Gruppe	N	BL 6MWD <sup>1</sup> (m) M (±SD) [Range]	Stop Codon-Typ n (%)			Localisation des Exons n (%)	
			UGA	UAG	UAA	1-39	40-79
PTC124-GD-007- DMD							
Placebo	57	361.12 (87.524) [158.5, 533.3]	31 (54.4%)	12 (21.1%)	14 (24.6%)	33 (57.9%)	24 (42.1)
ATL 10/10/20	57	350.00 (97.553) [75.0, 525.0]	29 (50.9%)	17 (29.8%)	11 (19.3%)	25 (43.9%)	32 (56.1)
ATL 20/20/40	60	361.19 (99.654) [90.2, 553.6]	23 (38.3%)	19 (31.7%)	18 (30.0%)	36 (60.0%)	24 (40.0)

Quelle (Studienreport PTC124-GD-007-DMD): S. 82, 83, 357

<sup>1</sup> Zwei Patienten (jeweils n=1 Placebo und ATL 20/20/40) hatten Verletzungen der unteren Extremitäten, die Einfluss auf die BL 6MWD hatten d.h. diese Werte waren protokollgemäß nicht valide. In der post-hoc-Analyse wurden die BL 6MWD-Werte durch die entsprechenden Screening-Werte ersetzt; die Werte der post-hoc-Analyse sind in der Tabelle dargestellt. Die BL 6MWD (Mittelwert ± SD) betrug 359.6 (87.7) m, 350.0 (97.6) m bzw. 358.2 (104.0) m in der Placebo-, ATL 10/10/10- bzw. 20/20/40-Gruppe wenn die BL-Werte der beiden Patienten mit Verletzungen der unteren Extremitäten verwendet wurden.

#### 4.4.2.4.2 Änderung der 6MWD und Progressor-Analyse - RCT

Wie bereits dargelegt wird die kITT-Analyse (verletzungsbedingt invalide 6MWD BL-Werte bei zwei Patienten durch die Screening-Werte ersetzt) als primär angesehen, da sie die tatsächlichen Fähigkeiten der Studienpopulation realistischer widerspiegelt als die ITT-Analyse und wissenschaftlich sowie klinisch begründet ist (siehe Kapitel 4.3.1.2.1). Die ursprünglich geplante ITT-Analyse wird lediglich zur unterstützenden Information dargestellt. Auch das CHMP (2014) stufte in dem Beurteilungsbericht von Ataluren die kITT-Analyse als akzeptabel ein (9).

Bei der überwiegenden Mehrzahl der Patienten lagen zu BL (96.6%) und in Woche 48 (97.1%) valide 6MWD-Ergebnisse vor. Insgesamt 15 Patienten (6 Placebo, 4 bzw. 5 Ataluren 10/10/20 bzw. 20/20/40) verloren während der Studie aufgrund eines Fortschreitens der Erkrankung ihre Gehfähigkeit und waren nicht in der Lage, an den nachfolgenden 6MWTs teilzunehmen. Protokollgemäß wurde für diese Tests ein Wert von 0 m dokumentiert.

Zu Anfang der Studie blieb die Variabilität der 6MWD (SD) innerhalb des antizipierten und der Fallzahlschätzung zugrunde gelegten Bereichs von etwa 50 m und entsprechend sah das DMC keine Notwendigkeit, die Fallzahl während der geplanten Interim-Analyse zu adjustieren. Die SD der 6MWD lag in Woche 48 mit im Mittel 72.0-90.0 m allerdings deutlich oberhalb der erwarteten Variabilität, was die Möglichkeiten, einen statistisch signifikanten Behandlungseffekt zu zeigen, einschränkte. In diesem Zusammenhang ist zu berücksichtigen, dass zum Zeitpunkt der Planung der Studie nicht auf Erfahrungen aus früheren Studien zurückgegriffen werden konnte und keine longitudinalen 6MWT-Daten bei DMD-Patienten vorlagen. Der Shapiro-Wilks W-Test zeigte eine signifikante Abweichung ( $p=0.0001$ ) von der Normalverteilung der nicht transformierten sowie log-transformierten 6MWD-Ergebnisse (gemischte Modelle mit gemischter symmetrischer, autoregressiver symmetrischer und unstrukturierter Kovarianzstruktur), sodass die Analyse der rank-transformierten Daten protokollgemäß als primär anzusehen war.

##### 4.4.2.4.2.1 Primäranalyse und Sensitivitätsanalysen der 6MWD (ITT)

Bei den Patienten der Placebo- bzw. Ataluren 20/20/40-Gruppe nahm die 6MWD im Mittel um 42.6 m bzw. 41.8 m bis Woche 48 ab verglichen mit einer Abnahme von nur 12.9 m in der Ataluren 10/10/20-Gruppe. In der MMRM-Analyse (ITT) betrug die mittlere Änderung der nicht transformierten 6MWD von BL bis Woche 48 26.4 m (nominal  $p=0.0905$ ) in der Ataluren 10/10/20-Gruppe und -0.1 m (nominal  $p=0.9956$ ) in der Ataluren 20/20/40-Gruppe verglichen mit Placebo. Basierend auf den rank-transformierten 6MWD-Daten, die aufgrund des Abweichens von der Normalverteilung als primär anzusehen gewesen wären, betrogen die nominalen p-Werte der mittleren Änderung bis Woche 48  $p=0.1490$  bzw.  $p=0.4756$  für Ataluren 10/10/20 bzw. 20/20/40 vs. Placebo. Auch der Dunnetts-Test zeigte keine signifikanten Unterschiede zwischen einer der Ataluren-Gruppen und Placebo (Tabelle 4-33). Dementsprechend wurde das primäre Ziel der Studie (mindestens einer der Atalurenarme ist dem Placeboarm hinsichtlich des primären Endpunkts signifikant überlegen) formell nicht erreicht. Der Unterschied zwischen Placebo und Ataluren 10/10/20 von BL bis Woche 48

entsprach allerdings größenordnungsmäßig dem hypothetischen Behandlungseffekt (30 m), der der Fallzahlschätzung zugrunde gelegt worden war, und der vorgeschlagenen MCID.

Tabelle 4-33: Änderung der 6MWD von BL bis Woche 48 in der prä-spezifizierten MMRM-Analyse unter Anwendung einer nicht strukturierten Kovarianzstruktur (ITT)

Gruppe	n	M (±SD) BL (m)	M (±SD) Woche 48 (m)	Δ vs. Placebo, M ± SE (95% KI) (m) <sup>1</sup>	p-Wert nominal vs. Placebo <sup>1</sup>	p-Wert adjustiert <sup>1,3</sup>	p-Wert nominal vs. Placebo <sup>2</sup>	p-Wert adjustiert <sup>2,3</sup>
Placebo	57	359.6 (87.7)	317.4 (152.3)	-	-	-		
Ataluren 10/10/20	57	350.0 (97.6)	342.7 (142.0)	26.44± 15.53 (-4.21, 57.09)	0.0905	0.1592	0.1490	0.2539
Ataluren 20/20/40	60	358.2 (104.0)	316.7 (161.6)	-0.09± 15.34 (-30.37, 30.20)	0.9956	1.0000	0.4756	0.6959

Quelle (Studienbericht): S. 92, 94, S. 456-8

In der Analyse wurde ein gemischtes Modell für wiederholte Messungen der 6MWD als abhängige Variable eingesetzt und die unabhängigen Variablen schlossen die BL 6MWD, Behandlung, Visite, Alter und Kortikosteroidanwendung und die Interaktion zwischen Behandlung und Visite ein.

<sup>1</sup> Nicht transformierte Werte

<sup>2</sup> Rang-transformierte Werte

<sup>3</sup> Der Dunnetts-Test wurde zur Adjustierung des Vergleichs von zwei Dosen vs. Placebo eingesetzt.

Da nur wenige 6MWD-Werte fehlten ergab die ITT-Analyse mit LOCF-Imputation vergleichbare Ergebnisse. Die Änderung (± SD) der 6MWD von BL bis Woche 48 betrug -40.42 m (89.556) in der Placebo-Gruppe, -14.57 m (71.396) (p=0.1766) bzw. -41.03 m (88.678) (p=0.9989) in der Ataluren 10/10/20- bzw. 20/20/40-Gruppe (ITT-LOCF). Wie in Tabelle 4-33 dargestellt entsprachen die Sensitivitäts- und supportiven Analysen der nicht transformierten 6MWT-Ergebnisse in ihrer Ausrichtung der Primäranalyse und lieferten keine Hinweise auf einen systematischen Bias. Hinsichtlich der Gehstrecke im 6MWT bis Woche 48 zeigte sich übereinstimmend ein Vorteil von Ataluren 10/10/20 nicht aber Ataluren 20/20/40 gegenüber Placebo. Ein signifikanter Unterschied zwischen Ataluren 10/10/20 und Placebo wurde lediglich in der ANCOVA-LOCF-Analyse beobachtet (nominal p=0.0445). Wie bereits oben angesprochen (Kapitel 4.3.1.2.1) war der Unterschied der p-Werte zwischen der ANCOVA-LOCF- und der MMRM-Analyse angesichts nur weniger fehlender Werte überraschend und führte zu einer post-hoc-Modifikation des MMRM-Modells (zusätzlicher Interaktionsterm BL 6MWD-Visite; siehe Kapitel 4.2.5.1 und Anhang 4-E). Der zusätzliche Interaktionsterm erlaubte eine flexiblere BL-Adjustierung zu jeder Visite, da der Zusammenhang zwischen der BL 6MWD und der post-BL 6MWD über die Zeit variieren kann.

Im Permutationstest (10000 Re-Randomisierungen; ITT) betrug die LSM-Differenz 0.322 m bzw. 29.0072 m für den Vergleich von Ataluren 20/20/40 (p=0.9798 nominal, p=1.0000

adjustiert) bzw. Ataluren 10/10/20 ( $p=0.0584$  nominal,  $p=0.1106$  adjustiert) verglichen mit Placebo.

Tabelle 4-34: Prä-spezifizierte Sensitivitätsanalysen der nicht transformierten 6MWD-Ergebnisse

Analyse	$\Delta$ LSM ( $\pm$ SE), BL bis Woche 48 vs. Placebo (m)	95% KI (m)	p-Wert nominal vs. Placebo	p-Wert adjustiert vs. Placebo <sup>1</sup>
<i>ANCOVA (ITT-LOCF)<sup>2</sup></i>				
ATL 10/10/20	28.3570 (14.0056)	0.71, 56.01	0.0445	0.0808
ATL 20/20/40	-0.7284 (13.8163)	-28.00, 26.55	0.9580	0.9980
<i>MMRM (ITT-LOCF)<sup>3</sup></i>				
ATL 10/10/20	25.78 (15.31)	-4.44, 56.00	0.0940	0.1649
ATL 20/20/40	-0.83 (15.11)	-30.66, 29.01	0.9564	0.9978
<i>MRRM (ITT, multiple Imputation)<sup>3</sup></i>				
ATL 10/10/20	25.66 (15.558)	-4.836, 56.153	0.0991	n.a.
ATL 20/20/40	-0.27 (15.366)	-30.383, 29.852	0.9862	n.a.
<i>MMRM (PP)<sup>3</sup></i>				
ATL 10/10/20	29.67 (15.69)	-1.31, 60.64	0.0604	0.1083
ATL 20/20/40	0.77 (15.42)	-29.68, 31.21	0.9605	0.9982
<i>Effekt des Einschlusses von Zwillingspaaren (MMRM, ITT)<sup>4</sup></i>				
ATL 10/10/20	26.03 (15.48)	-4.31, 56.36	0.0949	n.a.
ATL 20/20/40	-1.68 (15.30)	-31.66, 28.30	0.9485	n.a.

Quelle (Studienreport): S. 95, S. 462, S. 466, S. 482, S. 997, S. 1058

<sup>1</sup> Dunnetts-Test

<sup>2</sup> Faktoren: Behandlung, Alter und Kortikosteroid-Anwendung, BL 6MWD als Kovariate

<sup>3</sup> In der Analyse wurde ein gemischtes Modell für wiederholte Messungen der 6MWD als abhängige Variable eingesetzt und die unabhängigen Variablen schlossen die BL 6MWD, Behandlung, Visite, Alter und Kortikosteroid-Anwendung und die Interaktion zwischen Behandlung und Visite ein.

<sup>4</sup> Alle Zwillingspaare wurden derselben Behandlungsgruppe zugeteilt. Alle Ergebnisse stellen gepoolte Werte auf Basis des finalen gemischten Modells für 64 verschiedene Zwillingpopulationen dar. Der gepoolte Schätzer entspricht dem Durchschnitt des korrespondierenden Schätzers der 64 Zwillingpopulationen. Der gepoolte Schätzer des SE und der gepoolte p-Wert wurden nach der „Outputations“-Methode nach Follmann kalkuliert.

#### 4.4.2.4.2.2 kITT Analyse und Post-hoc-Analysen des 6MWT

Wie in Kapitel 4.3.1.2.1 dargelegt, wurden die BL 6MWD-Werte bei zwei Patienten, die kurz zuvor Verletzungen der unteren Extremitäten erlitten hatten, inkorrekt Weise durch den klinischen Beurteiler in einem Studienzentrum als valide eingestuft und dementsprechend in die ITT-Analyse eingeschlossen. Die Verletzungen der unteren Extremitäten waren bei den beiden Patienten als UE dokumentiert worden. Lediglich Patienten mit validen BL 6MWD-Werten und mindestens einem validen post-BL-Wert durften per Protokoll in die ITT-Population eingeschlossen werden. Der Vergleich der Gehstrecke zur Screening-Visite und der post-BL-Werte zeigte deutlich, dass der BL-Wert bei beiden Patienten nicht das tatsächliche Gehvermögen widerspiegelte und die niedrigen BL-Werte auf die Verletzungen der unteren Extremitäten zurückzuführen waren (Abbildung 9). Die Analysen wurden nach Ersatz der invaliden BL- durch die Screening-Werte dieser beiden Patienten in der kITT-Population wiederholt. Wie oben angesprochen wurde in der primären kITT-Analyse der

6MWD ein zusätzlicher Interaktionsterm (BL-Visite) in dem MMRM-Modell berücksichtigt. Die MMRM-Analyse basierte primär auf den nicht transformierten Daten und nicht den Rang-transformierten Daten wie in der Primäranalyse.

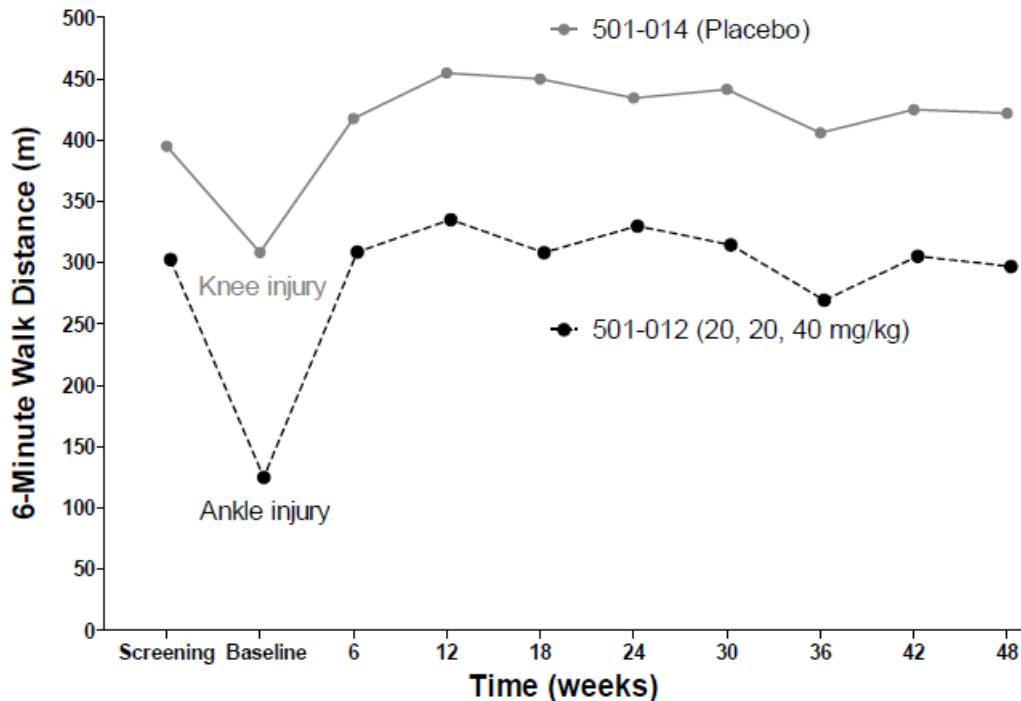


Abbildung 9: Ergebnisse der 6MWD bei zwei Patienten mit Verletzungen der unteren Extremitäten, die inkorrekt in die ITT-Population eingeschlossen worden waren

Der Shapiro-Wilks W-Test zeigte in der kITT-Population eine signifikante ( $p=0.0001$ ) Abweichung von der Annahme einer Normalverteilung der nicht transformierten und transformierten 6MWD-Werte in allen drei gemischten Modellen (mit gemischter symmetrischer, autoregressiver symmetrischer und unstrukturierter Kovarianzstruktur). Ein Permutationstest wurde post-hoc eingesetzt, um Abweichungen von der Annahme z.B. der Normalverteilung zu evaluieren.

Die Ergebnisse der 6MWD-Analysen in der kITT-Population und verschiedener Sensitivitätsanalysen sind in Tabelle 4-35 und Tabelle 4-36 dargestellt.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-35: Änderung der mittleren 6MWD über den Studienverlauf

Behandlungsgruppe	M (±SD) BL (m)	M (±SD) Woche 12 (m)	Δ Wo. 12 vs. BL, M (±SD) (m)	M (±SD) Woche 24 (m)	Δ Wo. 24 vs. BL, M (±SD) (m)	M (±SD) Woche 36 (m)	Δ Wo. 36 vs. BL, M (±SD) (m)	M (±SD) Woche 48 (m)	Δ Wo. 48 vs. BL, M (±SD) (m)
<i>kITT (beobachtete Werte)</i>									
Placebo	361.12 (87.524) (n=57)	354.30 (99.985) (n=57)	-7.15 (36.803)	346.48 (122.948) (n=57)	-14.64 (61.185)	342.38 (130.413) (n=53)	-22.49 (68.907)	317.38 (152.301) (n=55)	-44.14 (87.999)
ATL 10/10/ 20	350.00 (97.553) (n=57)	355.99 (118.170) (n=55)	5.81 (40.053) p=0.1259 <sup>1</sup>	350.51 (125.420) (n=56)	-2.89 (47.518) p=0.4289 <sup>1</sup>	344.91 (135.341) (n=54)	-8.81 (65.411) p=0.4851 <sup>1</sup>	342.73 (142.008) (n=55)	-12.86 (72.007) p=0.0851 <sup>1</sup>
ATL 20/20/ 40	361.19 (99.654) (n=60)	354.52 (111.372) (n=58)	-10.29 (35.666) p=0.8654 <sup>1</sup>	340.16 (139.995) (n=59)	-21.35 (57.824) p=0.7424 <sup>1</sup>	331.70 (153.004) (n=57)	-27.14 (72.148) p=0.9125 <sup>1</sup>	316.69 (161.572) (n=59)	-44.82 (84.776) p=0.9985 <sup>1</sup>
<i>PP (beobachtete Werte)</i>									
Placebo	359.94 (89.251) (n=55)	Daten nicht gezeigt	Daten nicht gezeigt	346.13 (125.185) (n=55)	-13.81 (64.587)	Daten nicht gezeigt	Daten nicht gezeigt	317.38 (152.301) (n=55)	-42.56 (90.046)
ATL 10/10/ 20	355.59 (94.395) (n=55)			353.84 (124.053) (n=55)	-1.75 (47.184) p=0.4621 <sup>1</sup>			342.73 (142.008) (n=55)	-12.86 (72.007) p=0.1186 <sup>1</sup>
ATL 20/20/ 40	358.49 (104.842) (n=59)			340.16 (139.995) (n=59)	-18.34 (64.630) p=0.8869 <sup>1</sup>			316.69 (161.572) (n=59)	-41.81 (89.234) p=0.9983 <sup>1</sup>
<i>kITT-LOCF</i>									
Placebo (n=57)	361.12 (87.524)	Daten nicht gezeigt	Daten nicht gezeigt	346.48 (122.948)	-14.64 (61.185)	Daten nicht gezeigt	Daten nicht gezeigt	319.18 (150.026)	-41.94 (87.610)
ATL 10/10/ 20 (n=57)	350.00 (97.553)			346.46 (128.012)	-3.54 (47.350) p=0.4613 <sup>1</sup>			335.43 (144.730)	-14.57 (71.396) p=0.1324 <sup>1</sup>
ATL 20/20/ 40 (n=60)	361.19 (99.654)			340.28 (138.807)	-20.92 (57.430) p=0.7662 <sup>1</sup>			317.20 (160.246)	-43.99 (84.297) p=0.9864 <sup>1</sup>

Quelle (Studienbericht): S. 337-341, S. 347-351, S. 357-361; <sup>1</sup> Dunnetts-Test vs. Placebo

Tabelle 4-36: Änderung der mittleren 6MWD über den Studienverlauf (augmentiertes gemischtes Modell mit unstrukturierter Kovarianzstruktur, kITT)

Behandlungs- gruppe	Woche 12 LSM (±SE) (m)	Woche 12 Δ (±SE) vs. Placebo [95% KI] (m)	Woche 12 p-Wert vs. Placebo nominal <sup>1</sup> , adjustiert <sup>2</sup>	Woche 24 LSM (±SE) (m)	Woche 24 Δ (±SE) vs. Placebo [95% KI] (m)	Woche 24 p-Wert vs. Placebo nominal <sup>1</sup> , adjustiert <sup>2</sup>	Woche 36 LSM (±SE) (m)	Woche 36 Δ (±SE) vs. Placebo [95% KI] (m)	Woche 36 p-Wert vs. Placebo nominal <sup>1</sup> , adjustiert <sup>2</sup>	Woche 48 LSM (±SE) (m)	Woche 48 Δ (±SE) vs. Placebo [95% KI] (m)	Woche 48 p-Wert vs. Placebo nominal <sup>1</sup> , adjustiert <sup>2</sup>
Placebo	349.67 (4.92)	-	-	341.21 (6.63)	-	-	325.94 (8.40)	-	-	312.08 (9.56)		
ATL 10/10/ 20	362.88 (4.98)	13.21 (6.81) [-0.24, 26.65]	0.0542 0.0977	354.53 (6.68)	13.33 (9.27) [-4.98, 31.63]	0.1525 0.2598	350.35 (8.45)	24.41 (11.80) [1.10, 47.72]	0.0402 0.0733	343.78 (9.62)	<b>31.71</b> <b>(13.46)</b> <b>[5.13,</b> <b>58.28]</b>	<b>0.0197</b> <b>0.0367</b>
ATL 20/20/ 40	344.35 (4.85)	-5.32 (6.73) [-18.61, 7.96]	0.4302 0.6443	334.30 (6.53)	-6.91 (9.17) [-25.01, 11.20]	0.4524 0.6700	328.51 (8.24)	2.57 (11.66) [-20.45, 25.60]	0.8257 0.9649	310.46 (9.36)	-1.62 (13.28) [-27.84, 24.61]	0.9031 0.9891

Quelle (Studienbericht): S. 555-557

Die Analysen basieren auf einem gemischten Modell für wiederholte Messungen der 6MWD als abhängige Variable und als unabhängige Variablen schloss das Modell die BL 6MWD, Behandlung, Visite, Alter und Kortikosteroid Stratifizierungsfaktoren sowie die Interaktion zwischen BL 6MWD und Visite ein.

<sup>1</sup> p-Wert Ataluren vs. Placebo<sup>2</sup> Dunnetts-Test

In der kITT-Analyse betrug die *mittlere Änderung* der 6MWD von BL bis Woche 48 -44.1 m in der Placebo-Gruppe (-42.6 m ITT), -12.9 m in der Ataluren 10/10/20-Gruppe (identisch in der ITT-Analyse) und -44.8 m in der Ataluren 20/20/40-Gruppe (41.8 m ITT). Der Unterschied zwischen Ataluren 10/10/20 und Placebo betrug **31.3 m** (Abbildung 10) und war somit der im Studienprotokoll spezifizierten Primäranalyse (ITT) gleichgerichtet und innerhalb des bei Studienplanung antizipierten Bereichs (30 m).

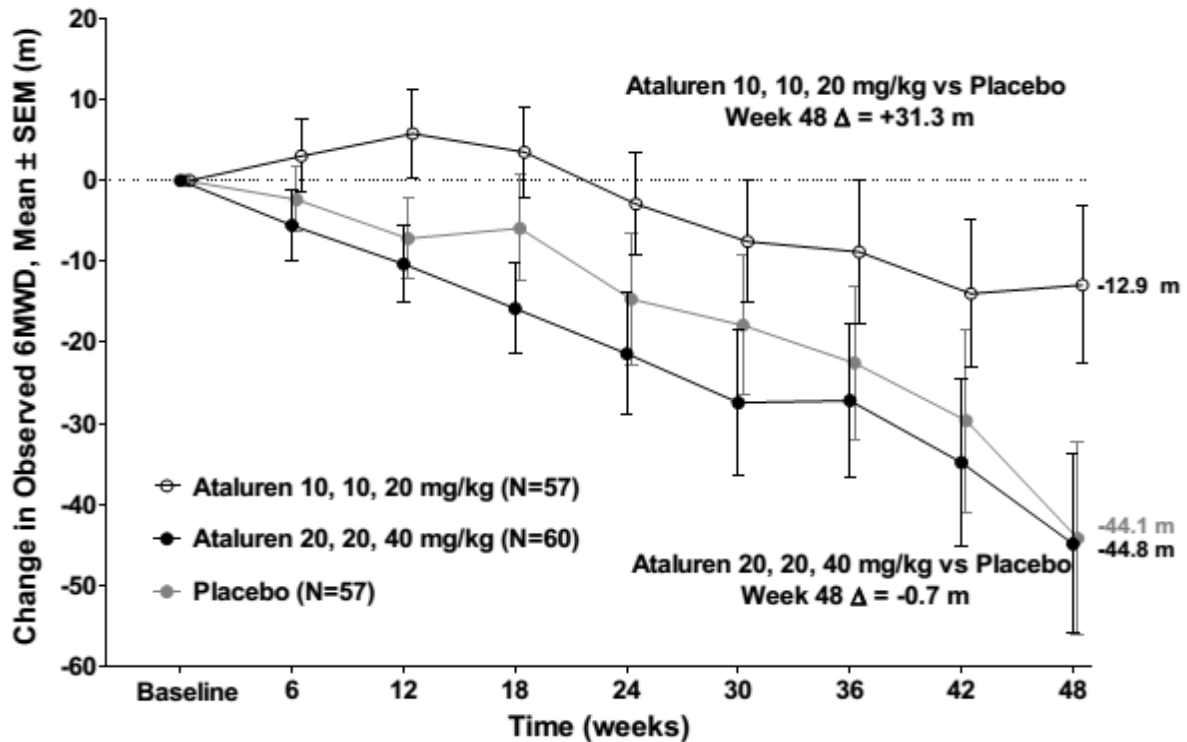


Abbildung 10: Mittlere Änderung der 6MWD bis Woche 48 (kITT)

Quelle (Studienbericht): S. 99

In der *primären kITT-Analyse* der 6MWD bis Woche 48 betrug der nominale *p-Wert* (adjustiert, Dunnett's Test) für den Vergleich von Ataluren 10/10/20 vs. Placebo **0.0197** (0.0367) und 0.9031 (0.9891) für den Vergleich von Ataluren 20/20/40 vs. Placebo. Der Unterschied zwischen Ataluren 10/10/20 und Placebo betrug **31.7 m (95% KI: 5.13, 58.28)**. Auch der Permutationstest zur Überprüfung der Robustheit der MMRM-kITT-Analyse zeigte einen signifikanten Unterschied zugunsten von Ataluren 10/10/20 vs. Placebo ( $p=0.0281$  nominal; Tabelle 4-37). Der Unterschied zwischen Ataluren 10/10/20 und Placebo zeigte sich bereits früh im Studienverlauf (etwa Woche 24-36) und blieb bis Woche 48 erhalten.



Tabelle 4-37: MMRM-Analyse der Änderung der untransformierten 6MWD und Permutationstest (kITT)

Analyse	Ataluren 10/10/20 vs. Placebo			Ataluren 20/20/40 vs. Placebo		
	$\Delta$ Mittelwert (95% KI)	p-Wert nominal	p-Wert adjustiert	$\Delta$ Mittelwert (95% KI)	p-Wert nominal	p-Wert adjustiert
MMRM <sup>1</sup>	31.7 (5.1, 58.3)	0.0197	0.0367 <sup>2</sup>	-1.62 (-27.8, 24.6)	0.9031	0.9891 <sup>2</sup>
Permutationstest <sup>3</sup>	-	0.0281 <sup>5</sup>	0.0561 <sup>4,6</sup>	-	0.9118 <sup>5</sup>	0.9910 <sup>4,6</sup>

Quelle (Studienreport): S. 101, S. 557, S. 1025

<sup>1</sup> MMRM-Modell: BL 6MWD (Kovariate) plus Behandlungsarm plus Visite plus Visite-Behandlungsarm plus BL 6MWD-Visite plus Altersgruppe (<9 vs.  $\geq$ 9 Jahre) plus Kortikosteroide (ja / nein); nicht strukturierte Varianz-/Kovarianzmatrix

<sup>2</sup> Der Dunnetts-Test wurde zur Adjustierung für den Vergleich von zwei Dosen vs. Placebo eingesetzt.

<sup>3</sup> Permutationstest von 10000 Re-Randomisierungen. Für jede Re-Randomisierung wurden die Patienten dynamisch in derselben Reihenfolge, in der sie ursprünglich in die Studie eingeschlossen worden waren, re-randomisiert (Start-Seed 14576).

<sup>4</sup> Basierend auf dem Anteil der 10000 Permutationen, bei denen die maximale Effektgröße der beiden Vergleiche (Ataluren 10/10/20 und 20/20/40 vs. Placebo) die beobachtete maximale Effektgröße überschritt.

<sup>5</sup> Der p-Wert entspricht der geschätzten Wahrscheinlichkeit, dass basierend auf dem Permutationstest die LSM-Differenz größer oder kleiner ist als der absolute Wert der aktuellen LSM-Differenz.

<sup>6</sup> Der p-Wert entspricht der geschätzten Wahrscheinlichkeit, dass basierend auf den Permutationen die maximale, absolute LSM-Differenz zwischen Ataluren und Placebo größer ist als die absolute, aktuelle LSM-Differenz zwischen Ataluren und Placebo.

Um die Robustheit der Analyse zu überprüfen wurden post-hoc Monte Carlo-Simulationen ausgeführt in denen randomisiert jeweils 10% der Patienten der Behandlungsarme entfernt wurden. Die Analyse wurde 1000fach in einer jeweils randomisiert unterschiedlichen Population wiederholt, um festzustellen, wie robust die Ergebnisse sind falls die Analysen konsistent im Vergleich zu Placebo waren. Alle Monte Carlo-Simulationen der 6MWD-Daten zeigten einen Trend zu einer besseren Wirksamkeit verglichen mit Placebo in der Gesamtpopulation mit einem deutlicheren Effekt in der “ambulatory decline”-Gruppe (siehe Legende der Abbildung 11). Diese Analysen bestätigten die interne Konsistenz der Analysen und Robustheit der Primäranalyse. Daher erscheint es unwahrscheinlich, dass die Ergebnisse zufälliger Natur oder auf Ausreißer zurückzuführen sind (31).

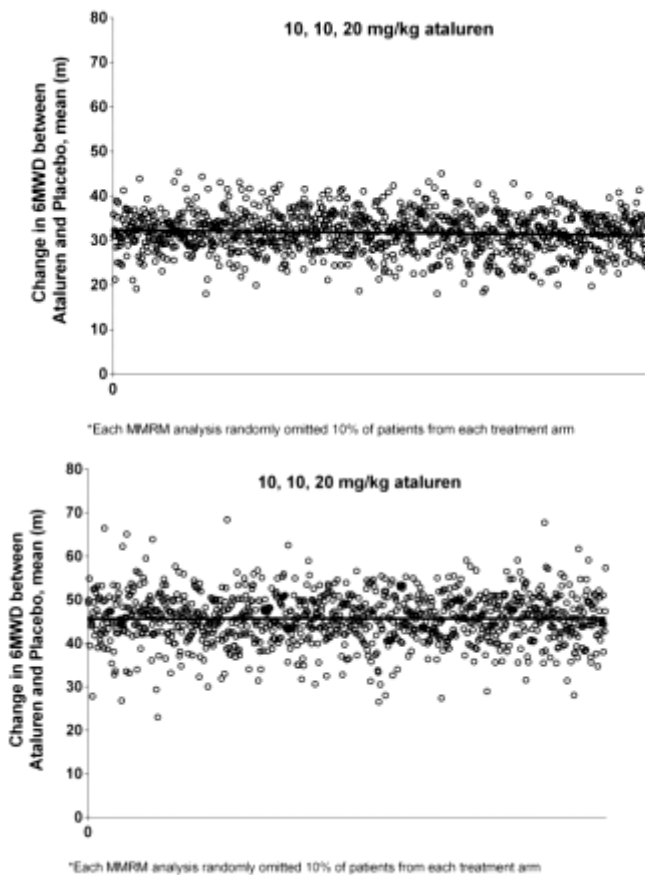


Abbildung 11: MMRM Monte Carlo-Analysen der 6MWD in der Gesamtpopulation (oben) und der „ambulatory decline“-Gruppe (unten) (31)

Jeder Kreis entspricht einer von insgesamt 1000 Monte Carlo-Analysen.

„Ambulatory decline“-Gruppe: Patienten im Alter von  $\geq 7$  bis  $\leq 16$  Jahren mit einer initialen 6MWD von  $\geq 150$  m und einer nach Alter und Größe erwarteten 6MWD von  $\leq 80\%$  unter stabiler Kortikosteroid-Therapie.

In zusätzlichen ITT-Analysen, in denen die Werte der beiden Patienten mit nicht validen BL-Werten ausgeschlossen wurden oder die BL-Werte durch andere Werte ersetzt wurden, war der Unterschied der 6MWD von BL bis Woche 48 zwischen der Ataluren 10/10/20 und der Placebo-Gruppe ebenfalls signifikant und die Effekte waren somit der primären Analyse gleichgerichtet (Tabelle 4-38).

Tabelle 4-38: Sensitivitätsanalysen Änderung der mittleren 6MWD (augmentiertes gemischtes Modell mit unstrukturierter Kovarianzstruktur)

Gruppe / Analyse	$\Delta$ ( $\pm$ SE) vs. Placebo (m)	95% KI (m)	p-Wert nominal	p-Wert adjustiert (Dunnetts)
<i>Ausschluss der beiden Patienten mit invalider BL-MWD (Studienreport S. 560)</i>				
ATL 10/10/ 20	32.70 (13.55)	5.95, 59.46	0.0169	0.0316
ATL 20/20/ 40	-1.75 (13.43)	-28.26, 24.77	0.8966	0.9876
<i>Mittelwert des Screening- und Woche 1-6MWD zum Ersatz des BL-Wertes (Studienreport S. 560)</i>				
ATL 10/10/ 20	36.19 (13.78)	8.98, 63.40	0.0094	0.0179
ATL 20/20/ 40	5.07 (13.60)	-21.77, 31.91	0.7097	0.9036
<i>Mittelwert des Screening- und Woche 1-6MWD zum Ersatz des BL-Wertes nach Ausschluss der beiden Patienten mit invaliden BL-Werten (Studienreport S. 568)</i>				
ATL 10/10/ 20	37.25 (13.88)	9.85, 64.64	0.0080	0.015
ATL 20/20/ 40	5.03 (13.75)	-22.11, 32.17	0.7148	0.9068
<i>Maximum des Screening- und Woche 1-6MWD zum Ersatz des BL-Wertes (Studienreport S. 573)</i>				
ATL 10/10/ 20	34.13 (14.30)	5.89, 62.38	0.0181	0.0339
ATL 20/20/ 40	1.99 (14.12)	-25.88, 29.86	0.8882	0.9855
<i>Maximum des Screening- und Woche 1-6MWD zum Ersatz des BL-Wertes nach Ausschluss der beiden Patienten mit invaliden BL-Werten (Studienreport S. 576)</i>				
ATL 10/10/ 20	35.57 (14.37)	7.20, 63.94	0.0143	0.0269
ATL 20/20/ 40	2.01 (14.24)	-26.10, 30.13	0.8877	0.9854
<i>PP (Studienreport S. 580)</i>				
ATL 10/10/ 20	31.11 (14.48)	2.51, 59.71	0.0332	0.0609
ATL 20/20/ 40	1.25 (14.23)	-26.85, 29.35	0.9303	0.9943

#### 4.4.2.4.2.3 10%ige Verschlechterung der 6MWD bis Woche 48 (Progressoranalyse)

Innerhalb von 48 Wochen verschlechterte sich die Gehfähigkeit bei 44% der Patienten der Placebo-Gruppe dauerhaft um mindestens 10% verglichen mit 26% in der Ataluren 10/10/20-Gruppe (kITT, nominal  $p=0.0386$ , adjustiert 0.0652; Abbildung 12). Der Anteil der Patienten, bei denen die Erkrankung bis Woche 48 fortschritt, unterschied sich nicht zwischen der ITT- und kITT-Analyse. Die OR (95% KI; kITT) betrug 0.3 (0.1, 0.9) für Ataluren 10/10/20 und 1.1 (0.5, 2.5) für Ataluren 20/20/40 jeweils verglichen mit Placebo und der CMH-Test zeigte einen signifikanten Unterschied ( $p=0.0423$ ).

Tabelle 4-39: Anteil der Patienten, deren 6MWD sich dauerhaft um mindestens 10% bis Woche 48 verschlechterte

Gruppe	N	n (%)	OR (Ataluren vs. Placebo)	95% KI	p-Wert
<i>ITT</i>					
Placebo	57	25 (43.9%)	-	-	-
Ataluren 10/10/20	57	15 (26.3%)	0.4	0.1, 0.9	0.0423 <sup>1</sup>
Ataluren 20/20/40	60	29 (48.3%)	1.1	0.5, 2.6	
<i>kITT</i>					
Placebo	57	25 (43.9%)	-	-	-

Ataluren 10/10/20	57	15 (26.3%)	<b>0.3</b>	<b>0.1, 0.9</b>	<b>0.0423<sup>1</sup></b>
Ataluren 20/20/40	60	29 (48.3%)	1.1	0.5, 2.5	

Quelle (Studienreport): S. 1008, S. 1009

Stratifizierungsfaktoren: Alter (<9 Jahre, ≥9 Jahre), Kortikosteroid-Anwendung zu BL und BL 6MWD (<350 m, ≥350 m)

<sup>1</sup> CMH-Test

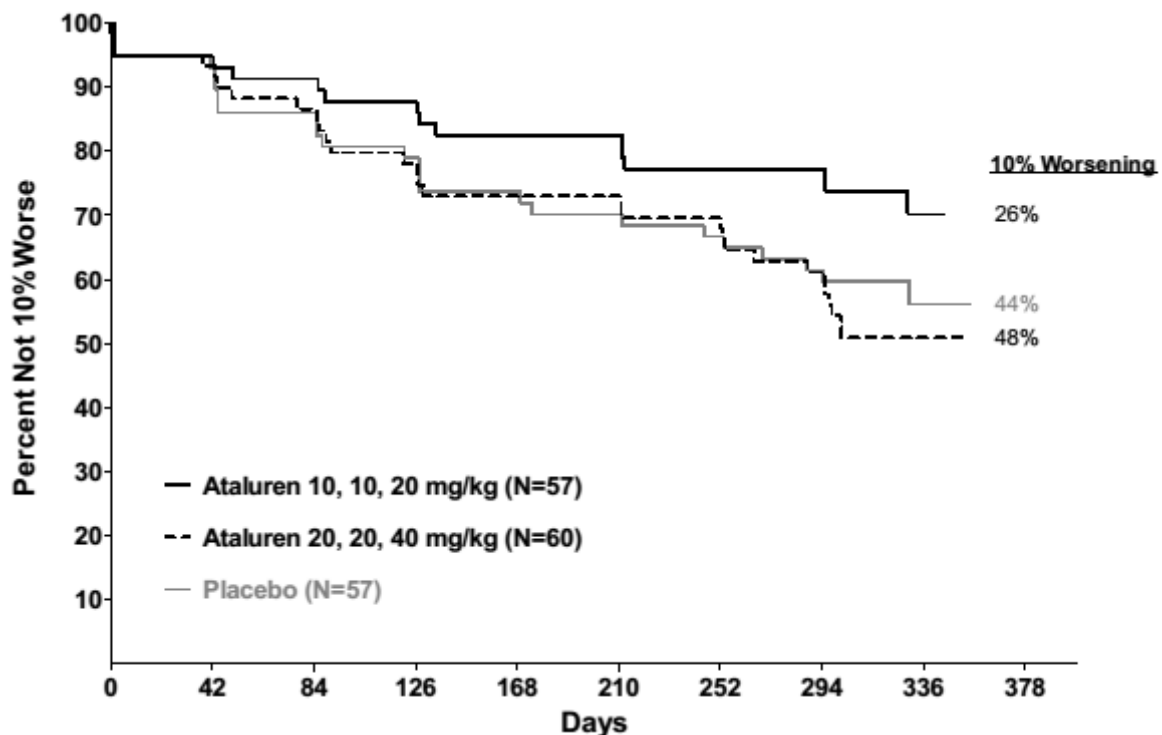


Abbildung 12: Zeit bis zu einer dauerhaften 10%igen Verschlechterung der 6MWD (kITT)

Quelle (Studienbericht): S. 102

Die Zeit bis zu einer dauerhaften mindestens 10%igen Verschlechterung der 6MWD von BL bis Woche 48 wurde in einem Cox-proportionalen Hazard-Modell analysiert. Die HR (95% KI; kITT) der Zeit bis zu einer mindestens 10%igen Verschlechterung der 6MWD in Woche 48 für den Vergleich von Placebo und Ataluren 10/10/20 betrug 0.508 (0.273, 0.946; nominal p=0.0326, adjustiert p=0.0652; Tabelle 4-40), was einer 48%igen Reduktion des Risikos unter Behandlung mit Ataluren 10/10/20 entsprach. Zwischen der Ataluren 20/20/40- und der Placebo-Gruppe ergab sich auch in dieser Auswertung kein Unterschied (kITT, HR 1.126, 95% KI 0.656, 1.933, nominal p= 0.6667), was eine inverse Dosis-Effekt-Beziehung nahelegt. Die Effekte waren in der ITT-Analyse und einer Analyse unter Ausschluss der beiden Patienten mit invaliden BL 6MWD-Werten vergleichbar. In der ITT-Analyse betrug die HR für den Vergleich von Ataluren 10/10/20 vs. Placebo 0.520 (95% KI 0.280-0.966; nominal p=0.0386, adjustiert p=0.078), was einer 48%igen Reduktion des Risikos einer dauerhaften mindestens 10%igen Verschlechterung der Gehstrecke im 6MWT entsprach. Wie aus Abbildung 12 ersichtlich war der Unterschied zwischen Ataluren 10/10/20 und Placebo

bereits zu einem frühen Zeitpunkt erkennbar und blieb über den gesamten Studienverlauf erhalten.

Zur klinischen Relevanz der mindestens 10%igen Verschlechterung der 6MWD siehe Kapitel 4.4.2.3.

Tabelle 4-40: Zeit bis zu einer mindestens 10%igen Verschlechterung der 6MWD in dem Cox-proportionalen Hazard-Modell

Gruppe	Schätzer	HR	95% KI	p-Wert (Wald)
<i>ITT</i>				
Ataluren 10/10/20 vs. Placebo	-0.654	0.520	0.280, 0.966	0.0386
Ataluren 20/20/40 vs. Placebo	0.142	1.153	0.672, 1.977	0.6055
<i>kITT</i>				
Ataluren 10/10/20 vs. Placebo	-0.677	0.508	0.273, 0.946	0.0326
Ataluren 20/20/40 vs. Placebo	0.119	1.126	0.656, 1.933	0.6667
<i>Ausschluss der Patienten 501-012 und 501-014</i>				
Ataluren 10/10/20 vs. Placebo	-0.696	0.499	0.268, 0.928	0.0280
Ataluren 20/20/40 vs. Placebo	0.194	1.214	0.706, 2.088	0.4821

Quelle (Studienreport): S. 101, S. 1019, S.1020, S. 1021

Die Kovariaten in dem Modell schlossen die Behandlung, Altersgruppen (<9 Jahre, ≥9 Jahre), Kortokosteroid-Behandlung und die BL 6MWD (<350 m, ≥350 m) ein.

Hinsichtlich des Anteils der Patienten, die auf die Therapie ansprachen (dauerhafte 10%ige Verbesserung der 6MWD bis Woche 48), ergab sich im Gruppenvergleich kein signifikanter Unterschied (kITT, 21% bzw. 17% Ataluren 10/10/20 bzw. 20/20/40, 9% Placebo), da die verglichenen Populationen sehr klein waren.

#### *Sensitivitätsanalyse Krankheitsprogression:*

Um die therapeutische Relevanz der Therapieunterschiede in Bezug auf die Krankheits-Progression unabhängig von dem 10% Kriterium (10%ige Gehstreckenverschlechterung) in einer Sensitivitätsanalyse zu untersuchen, wurde in Anlehnung an das Beratungsgespräch mit dem G-BA vom 30.10.2014 eine etablierte statistische Verteilungsmethode für eine Progressor-Analyse herangezogen. Für die Nutzenbewertung wurde die Responderanalyse für die kITT Population durchgeführt, in der das Responsekriterium für die Krankheitsprogression von „mindestens 0.5 x SD Verschlechterung der 6MWD bis Woche 48“ eingesetzt wurde. Die Standardabweichung ist definiert als SD der gepoolten BL-Werte für die gesamte Studienpopulation, und für die 0.5 x SD ergab sich ein Schwellenwert für die Progressor-Analyse von -47.5 Metern. Dieser Ansatz der Verwendung eines minimal relevanten Schwellenwertes ist analog zur Verwendung der Schwellenwerte für ein Hedges'g, unterscheidet sich jedoch hiervon, dass das Relevanzkriterium auf individueller Patientenbasis und nicht auf Populationsebene angewendet wird (18).

Die aus der logistischen Regressionsanalyse resultierenden Odds Ratios waren für den Vergleich von Ataluren 10/10/20 mit Placebo signifikant (p=0.0313 Auswertung der

beobachteten Werte;  $p=0.0224$  Auswertung in der Patienten mit fehlenden Werten als Progressor gewertet wurden), und zeigten eine therapeutische Überlegenheit von Ataluren 10/10/20 unter Verwendung des Schwellenwertes von  $-47.5$  Metern (Tabelle 4-41, Tabelle 4-42).

Im Vergleich zur Analyse der 10%igen Verschlechterung, in der ein relativer Wert für die Verschlechterung der 6MWD eingesetzt wurde, demonstrierte auch die Sensitivitätsanalyse unter Verwendung eines absoluten, etablierten Kriteriums (18) eine relevante therapeutische Überlegenheit von Ataluren 10/10/20 gegenüber der Vergleichsbehandlung.

Tabelle 4-41: Progressor Analyse (Verschlechterung der 6MWD um  $\geq 0.5 \times SD$  bis Woche 48, beobachtete Werte, kITT)

	<b>Ataluren 20/20/ 40</b>	<b>Ataluren 10/10/20</b>	<b>Placebo</b>	<b>p-Wert</b>
	(N = 60)	(N = 57)	(N = 57)	(Unterschied zwischen 3 Armen)
Progressors [1]	25/59 (42.4%)	14/55 (25.5%)	23/55 (41.8%)	0.0460 [3]
Odds Ratio (Ataluren vs. Placebo)	1.042	0.354		
95% Konfidenz Intervall	(0.440, 2.465)	(0.138, 0.911)		
p-Wert (Ataluren vs. Placebo)	0.9255 [3]	0.0313 [3]		
Logistische Regression Faktor/Kovariate [2]:	Alter < 9 vs. Alter $\geq 9$	Baseline 6MWD (Meter)	Kortikosteroide: Ja vs. Nein	
Odds Ratio	0.315	0.99	0.618	
95% Konfidenz Intervall	(0.147, 0.676)	(0.985, 0.994)	(0.275, 1.391)	

[1] Progressors = Patienten mit mindestens  $0.5 \times SD$  (Baseline 6MWD) Verschlechterung zu Woche 48 in der 6MWD

[2] Logistische Regression Faktoren/Kovariaten: Alter (< 9 Jahre oder  $\geq 9$  Jahre), Begleitbehandlung mit Kortikosteroiden bei BL (Ja oder Nein), 6-Minuten Gehstrecke (6MWD) bei BL (Meter)

[3] p-Werte der logistischen Regression.

Tabelle 4-42: Progressor Analyse (Verschlechterung der 6MWD um  $\geq 0.5 \times$  SD bis Woche 48, fehlende Werte imputiert, kITT)

	<b>Ataluren 20/20/ 40</b>	<b>Ataluren 10/10/20</b>	<b>Placebo</b>	<b>p-Wert</b>
	(N = 60)	(N = 57)	(N = 57)	(Unterschied zwischen 3 Armen)
Progressors [1,4]	26/60 (43.3%)	16/57 (28.1%)	25/57 (43.9%)	0.0363 [3]
Odds Ratio (Ataluren vs. Placebo)	1.008	0.34		
95% Konfidenz Intervall	(0.432, 2.348)	(0.135, 0.858)		
p-Wert (Ataluren vs. Placebo)	0.9861 [3]	0.0224 [3]		
Logistische Regression Faktor/Kovariate [2]:	Alter < 9 vs. Alter $\geq$ 9	Baseline 6MWD (Meter)	Kortikosteroide: Ja vs. Nein	
Odds Ratio	0.321	0.989	0.64	
95% Konfidenz Intervall	(0.151, 0.680)	(0.985, 0.994)	(0.288, 1.422)	

[1] Progressors = Patienten mit mindestens  $0.5 \times$  SD (Baseline 6MWD) Verschlechterung zu Woche 48 in der 6MWD

[2] Logistische Regression Faktoren/Kovariaten: Alter (< 9 Jahre oder  $\geq$  9 Jahre), Begleitbehandlung mit Kortikosteroiden bei BL (Ja oder Nein), 6-Minuten Gehstrecke (6MWD) bei BL (Meter)

[3] p-Werte der logistischen Regression.

[4] fehlende Werte wurden als Verschlechterung um mindestens  $0.5 \times$  SD (Baseline 6MWD) imputiert

#### 4.4.2.4.2.4 Subgruppenanalysen - RCT

##### *kITT – Population*

Die im Studienprotokoll geplanten Subgruppen-Analysen wurden hier nur in der bewertungsrelevanten kITT Population dargestellt. Für die Nutzenbewertung ist insbesondere der Einfluss der Stratifizierungsfaktoren auf die Änderung der 6MWD bis Woche 48 relevant. In den Subgruppen der Patienten deren Gehstrecke zu BL unter 350 m im 6MWT lag, und in den Subgruppen der Patienten unter 9 Jahren, sowie den Patienten die zu BL Kortikosteroide eingenommen hatten, waren die Gehstreckenunterschiede im 6MWT in der Ataluren 10/10/20-Gruppe gegenüber der Placebo-Gruppe statistisch signifikant und im Vergleich mit der Gesamtpopulation deutlich größer (Tabelle 4-43).

Insbesondere in der Subgruppe der Patienten, deren *Gehstrecke bei BL unter 350 m* lag (Stratifizierungsfaktor), zeigte sich ein beträchtlicher Unterschied von **59.81 m** für die Ataluren 10/10/20-Gruppe verglichen mit Placebo ( $p=0.0053$ ) in der Auswertung mittels des MMRM Models (Tabelle 4-43). Diese Primärauswertung ist konservativ, eine Auswertung der Mittelwertsunterschiede zu Woche 48 zeigte, dass der tatsächlich beobachtete Unterschied bei 68.2 m lag ( $p=0.0053$ , nominal) (siehe Abbildung 13, Abbildung 14).

Wie in Kapitel 4.4.2.3 beschrieben, befanden sich die Patienten mit Baseline-Werten > 350 m noch nicht in der Krankheitsphase des rapiden Mobilitätsverlustes. Auf Grund des deterministischen und progressiven Krankheitsverlaufes erreichen jedoch auch diese Patienten nach einigen Jahren die Phase des schnell fortschreitenden Gehstreckenverlustes (32), und sollten dann auch messbar von dem Therapieeffekt der Ataluren-Behandlung profitieren (in Bezug auf den Erhalt der Gehfähigkeit). Der therapeutische Effekt der Ataluren-Behandlung war daher in der 48-wöchigen Studie für Patienten mit Baseline-Werten > 350 m erwartungsgemäß geringer als für Patienten mit Baseline-Werten < 350 m.

Tabelle 4-43: Änderung der mittleren 6MWD über 48 Wochen in Abhängigkeit von den Stratifizierungsfaktoren (**kITT**)

Gruppe / Analyse	$\Delta$ LSM ( $\pm$ SE) in Woche 48 [95% KI] (m)	$\Delta$ ( $\pm$ SE) vs. Placebo, BL bis Woche 48 (m)	95% KI (m)	p-Wert nominal vs. Placebo
<i>Alter &lt;9 Jahre (Studienreport S. 560, S. 665)</i>				
Placebo (n=32)	-35.95 (12.69) [-61.01, -10.89]	-	-	-
ATL 10/10/ 20 (n=32)	7.05 (12.55) [-17.73, 31.84]	43.00 (17.81)	7.84, 78.16	0.0168
ATL 20/20/ 40 (n=34)	-33.93 (12.44) [-58.50, -9.36]	2.01 (17.64)	-32.82, 36.84	0.9093
<i>Alter <math>\geq</math>9 Jahre (Studienreport S. 560, S. 665)</i>				
Placebo (n=25)	-54.74 (14.37) [-83.12, -26.37]			
ATL 10/10/ 20 (n=25)	-38.76 (14.55) [-67.48, -10.04]	15.98 (20.32)	-24.13, 56.09	0.4327
ATL 20/20/ 40 (n=26)	-60.79 (14.01) [-88.44, -33.13]	-6.04 (19.93)	-45.39, 33.31	0.7622
<i>Kortikosteroide zu BL: ja (Studienreport S. 560, 614)</i>				
Placebo (n=40)	-46.61 (11.46) [-69.24, -23.99]	-	-	-
ATL 10/10/ 20 (n=41)	-7.10 (11.41) [-29.64, 15.44]	39.51 (16.18)	7.58, 71.45	0.0156
ATL 20/20/ 40 (n=43)	-38.70 (11.18) [-60.79, -16.62]	7.91 (16.01)	-23.70, 39.52	0.6219
<i>Kortikosteroide zu BL: nein (Studienreport S. 560, 614)</i>				
Placebo (n=17)	-39.21 (17.66) [-74.07, -4.35]	-	-	-
ATL 10/10/ 20 (n=16)	-26.22 (18.11) [-61.98, 9.54]	12.99 (25.31)	-36.98, 62.96	0.6085
ATL 20/20/ 40 (n=17)	-63.44 (17.60) [-98.19, -28.68]	-24.23 (24.95)	-73.49, 25.03	0.3329



Gruppe / Analyse	$\Delta$ LSM ( $\pm$ SE) in Woche 48 [95% KI] (m)	$\Delta$ ( $\pm$ SE) vs. Placebo, BL bis Woche 48 (m)	95% KI (m)	p-Wert nominal vs. Placebo
<i>BL 6MWD &lt;350 m (Studienreport S. 560, S.639)</i>				
Placebo (n=22)	-74.60 (17.56) [-109.3, -39.94]	-	-	-
ATL 10/10/ 20 (n=25)	-14.79 (17.21) [-48.77, 19.18]	<b>59.81 (21.16)</b>	<b>18.03, 101.59</b>	<b>0.0053</b>
ATL 20/20/ 40 (n=27)	-64.34 [16.49] [-96.90, -31.77]	10.26 (20.78)	-30.76, 51.29	0.6220
<i>BL 6MWD <math>\geq</math>350 m (Studienreport S. 560, S. 639)</i>				
Placebo (n=35)	-25.06 (13.40) [-51.52, 1.40]	-	-	-
ATL 10/10/ 20 (n=32)	-11.67 (13.91) [-39.14, 15.80]	13.39 (17.52)	-21.20, 47.98	0.4458
ATL 20/20/ 40 (n=33)	-30.76 (14.42) [-59.23, -2.30]	-5.71 (17.45)	-40.16, 28.75	0.7441

Die Analysen basieren auf einem gemischten Modell mit wiederholten Messungen der 6MWD als abhängige Variable und als unabhängige Variablen schloss das Modell die BL 6MWD, Behandlung, Visite, jeweiliger Stratifizierungsfaktor und Interaktionen zwischen Behandlung und Visite, BL 6MWD und Visite, jeweiligem Stratifizierungsfaktor und jeweiligem Stratifizierungsfaktor und Behandlung, Visite, jeweiligem Stratifizierungsfaktor und Behandlung und Visite ein.

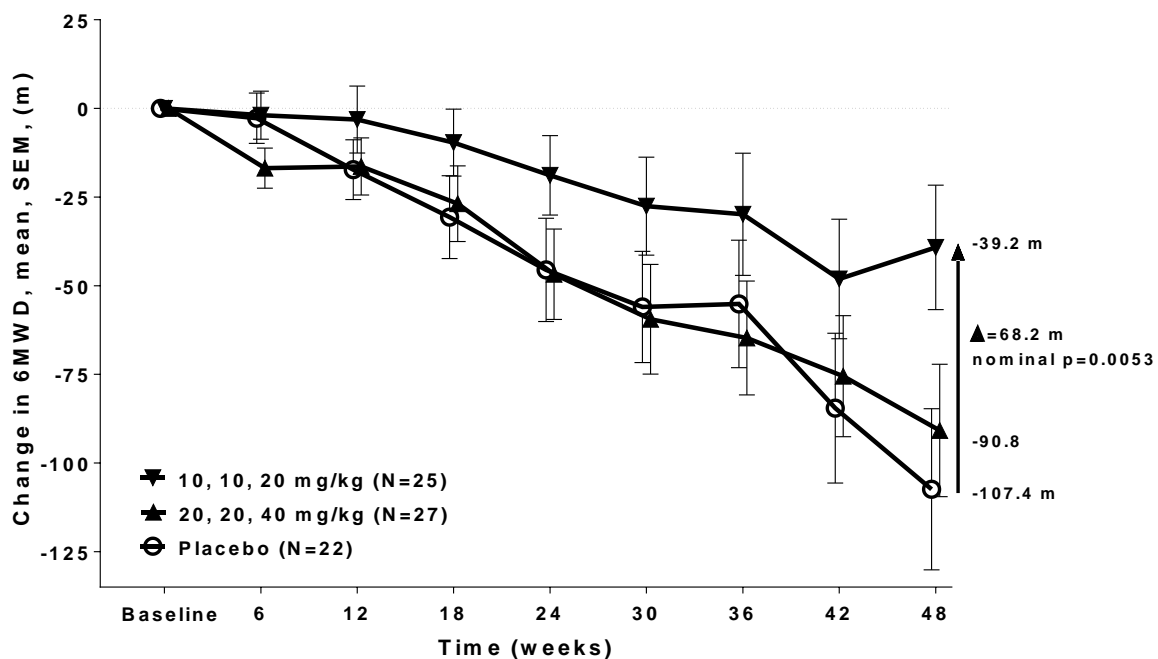


Abbildung 13: Mittlere Änderung der beobachteten 6MWD über den Studienverlauf in der Subgruppe BL 6MWD <350 m (kITT)

### Post-hoc Analysen

Die Analyse der Daten der Placebo-Gruppe sowie andere unabhängige longitudinale Studien (z.B. (27)) zeigte, dass Patienten über 7 Jahren mit einer BL 6MWD von  $< 350$  m insgesamt eine schlechter Prognose (Verlust der Gehfähigkeit) haben als Patienten, die diesen Kriterien nicht entsprechen. Basierend auf diesen Daten wurde post-hoc eine sogenannte "ambulatory decline phase"-Population definiert, d.h. Patienten im Alter von  $\geq 7$  bis  $\leq 16$  Jahren mit einer initialen 6MWD von  $\geq 150$  m und einer nach Alter und Größe erwarteten 6MWD von  $\leq 80\%$  unter stabiler Kortikosteroid-Therapie. Aufgrund des Wirkmechanismus (Stabilisierung der Muskelfunktion durch Dystrophin) wurde erwartet, dass die Effekte von Ataluren über ein Jahr bei Patienten mit hohem Risiko einer Krankheitsprogression deutlicher ausfallen würden. Eine post-hoc Analyse der Subgruppe der Patienten in der "ambulatory decline"-Phase bestätigte diese Annahme und der Unterschied zwischen Ataluren 10/10/20 und Placebo hinsichtlich der Änderung der Gehstrecke im 6MWT bis Woche 48 war ausgeprägter als in der Gesamtpopulation. In dieser Subgruppe betrug der Unterschied zugunsten von Ataluren 10/10/20 49.9 m ( $p=0.0096$  nominal) verglichen mit 31.3 m in der Gesamtpopulation (Abbildung 14).

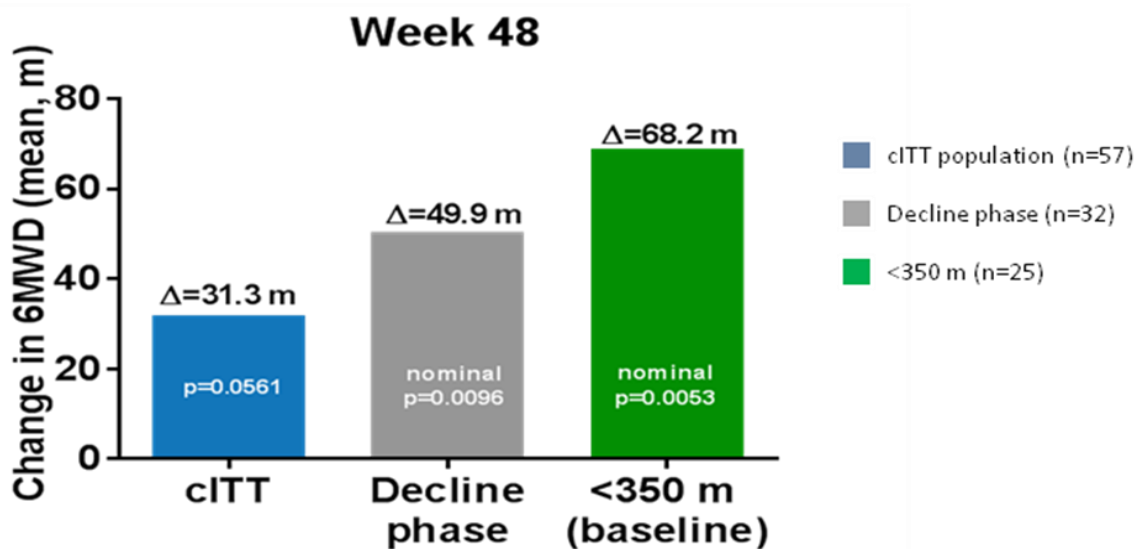


Abbildung 14: Mittlere Änderung der beobachteten 6MWD über den Studienverlauf in der Gesamtpopulation, "ambulatory decline phase"-Gruppe,  $< 350$  m BL-Gruppe (kITT)  
p-Werte sind nur für den Vergleich der Mittelwerte zu Woche 48 angegeben

Obwohl Patienten mit bereits eingeschränkter Gehfähigkeit (BL 6MWD  $< 350$  m) oder einem hohem Risiko einer Krankheitsprogression ("ambulatory decline group") am meisten von der Therapie mit Ataluren 10/10/20 zu profitieren scheinen, zeigte sich der Effekt von Ataluren 10/10/20 entlang des gesamten Krankheitsspektrums der eingeschlossenen Patienten. Der Schweregrad des Gehverlusts kann bei Jungen mit DMD auf Basis der prozentual bei gesunden Jungen vergleichbaren Alters und vergleichbarer Größe erwarteten 6MWD kategorisiert werden. In der Studie PTC124-GD-007-DMD profitierten alle Patienten inklusive

derjenigen Patienten mit milderer Krankheitsausprägung/früherem Progressionsstadium zu Studienbeginn (>70% der vorausgesagten prozentualen 6MWD) von der Therapie mit Ataluren 10/10/20 vs. Placebo über 48 Wochen (Abbildung 15) (31). Zum Zeitpunkt der Planung der Studie lagen keine Daten zur 6MWD bei DMD-Patienten relativ zu Gesunden vor, so dass diese Daten *post-hoc* evaluiert wurden.

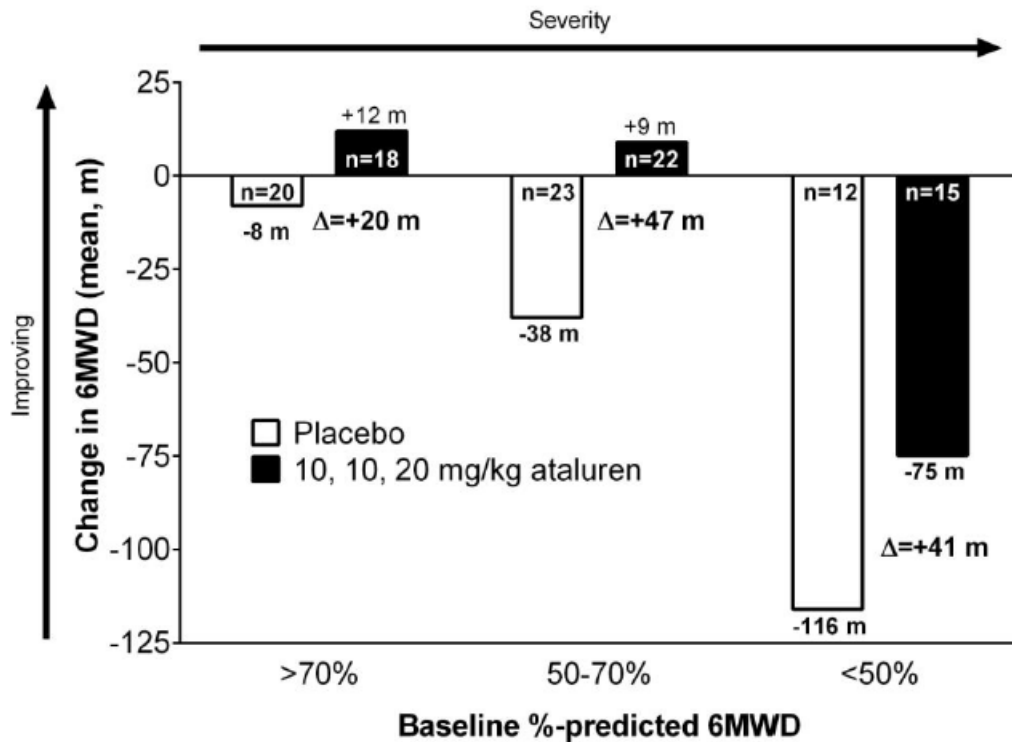


Abbildung 15: Änderung der mittleren 6MWD bis Woche 48 stratifiziert nach prozentual erwarteter 6MWD zu Beginn der Studie (31)

Die Longitudinaldaten (siehe Kapitel 4.4.2.3) sowie die Subgruppen-Analyse bei Patienten in der "ambulatory decline phase" der Studie PTC124-GD-007-DMD bildeten die Grundlage für das Design der noch andauernden konfirmatorischen, Placebo-kontrollierten Phase 3-Studie. In der konfirmatorischen Studie werden entsprechend der definierten Einschlusskriterien (z.B.  $\geq 7$  Jahre, BL 6MWD  $\leq 80\%$  des nach Alter und Größe erwarteten Wertes) Patienten in der "ambulatory decline phase" rekrutiert, für die nach vorliegenden Daten der Therapieeffekt von Ataluren am besten dargestellt werden kann. Aufgrund des Wirkmechanismus erscheint es plausibel, dass auch oder gerade Patienten mit noch milder Krankheitsausprägung langfristig von Ataluren profitieren könnten, allerdings wäre eine Placebo-kontrollierte Studie über viele Jahre ethisch nicht vertretbar (31).

### Fazit

Der Erhalt der Gehfähigkeit ist für Patienten unabhängig von der Grunderkrankung von unmittelbarer klinischer Relevanz und bei DMD-Patienten ist das Gehvermögen im 6MWT ein Prädiktor des Risikos in den folgenden zwei Jahren, die Gehfähigkeit zu verlieren (27).

Die Gehfähigkeit wird auch von der EMA (2013) als „Meilenstein“ in der Krankheitsprogression der DMD eingestuft (26).

In der korrigierten ITT Analyse (kITT), in der die nicht validen BL-Werte von zwei Patienten durch die Screening-Werte ersetzt wurden und ein zusätzlicher Interaktionsterm (BL-Visite) eingesetzt wurde, betrug der Unterschied der Gehstrecke im 6MWT zwischen Ataluren 10/10/20 und Placebo bis Studienende 31.7 m und erreichte das Signifikanzniveau (nominal  $p=0.0197$ ). Ein Re-Randomisierungstest, der die dynamische Randomisierung sowie Abweichungen von den zugrunde gelegten Annahmen wie z.B. der Normalverteilung der Daten adressierte, bestätigte, dass das Ergebnis der Änderung der 6MWD zugunsten von Ataluren 10/10/20 vs. Placebo nicht zufällig war. Der Unterschied zwischen Ataluren 10/10/20 und Placebo hinsichtlich der Änderung der 6MWD bis Woche 48 entsprach dem bei Studienplanung projizierten Wert (30 m). Der Behandlungseffekt von Ataluren 10/10/20 zeigte sich bereits früh in Woche 24-36 und der Unterschied zur Vergleichstherapie vergrößerte sich bis Studienende stetig. Das CHMP (2014) stuft die kITT-Analyse als methodisch akzeptabel ein und die Ergebnisse der kITT-Analysen flossen in die Bewertung durch das CHMP ein (9). Die kITT-Analyse wurde als klinisch sowie wissenschaftlich begründet eingeschätzt und die Bewertung von Ataluren bezieht sich auf diese Population.

Wie in Kapitel 4.4.2.3 detaillierter dargelegt, wird ein Unterschied von 30 m im 6MWT bei DMD-Patienten als klinisch relevant und für die Patienten bedeutsam eingestuft:

- Zwei verteilungsbasierte Methoden zeigten MCIDs von 28.5 bzw. 31.7 m im 6MWT bei nmDMD-Patienten (11).
- Eine Anker-basierte Methode, welche die 6MWD und ein QoL Instrument (PODCI) korrelierte, zeigte dass eine Änderung von 30 m im 6MWT über 48 Wochen für Patienten mit einer mittleren Gehfähigkeit wie in der Studie PTC124-GD-007-DMD als klinisch relevant eingestuft werden kann (13).
- Longitudinale Studien zum natürlichen Verlauf der DMD zeigten, dass jede Reduktion der BL 6MWD um 30 m das Risiko erhöht, innerhalb von zwei Jahren die Gehfähigkeit zu verlieren (siehe Abbildung 5; Mercuri et al. zitiert nach (30)).

Aufgrund der vorgenannten Fakten schätzt auch das CHMP (2014) in dem Beurteilungsbericht zu Ataluren den Unterschied von 30 m im 6MWT als klinisch bedeutsam ein (9).

Die Robustheit der primären Auswertung wird zusätzlich durch eine prä-spezifizierte Analyse der Zeit bis zum Fortschreiten der Erkrankung (mindestens 10% Verschlechterung der 6MWD) gestützt (9). Bei nur 26% der Patienten der Ataluren 10/10/20-Gruppe verglichen mit 44% der Patienten der Placebo-Gruppe schritt die Erkrankung im Studienverlauf fort (HR 0.508, 95% KI 0.273, 0.946). Die Longitudinal-Studie von Henricson et al. (2013) bei Jungen mit DMD zeigt, dass unter Bezug auf das QoL-Instrument PODCI eine Verschlechterung der 6MWD um 10% über einen breiten 6MWD-Bereich klinisch bedeutsam ist.

Sensitivitätsanalysen und Subgruppenanalysen bestätigten das Ergebnis der primären Analyse des 6MWT. In einer geplanten Subgruppenanalyse war der Unterschied zugunsten von Ataluren 10/10/20 vs. Placebo bei den Patienten, die zu BL weniger als 350 m im 6MWT gehen konnten, besonders deutlich (68.2 m). Eine post-hoc-Analyse stratifiziert nach Gehfähigkeit zu BL (% des erwarteten 6MWD-Wertes) zeigte, dass Ataluren 10/10/20 über das gesamte Spektrum der eingeschlossenen Patienten aktiv ist, inklusive derjenigen Patienten die sich zu Studienbeginn in einem frühen Stadium der Krankheitsprogression befanden (13).

In Übereinstimmung mit der Einschätzung des CHMPs (2014) wird ein Unterschied um etwa 30 m im 6MWT bis Woche 48 unter Behandlung mit Ataluren 10/10/20 vs. Placebo als klinisch bedeutsam eingestuft (9). Die Gehfähigkeit ist bei Erkrankungen wie der DMD, die mit einem fortschreitenden Verlust der Gehfähigkeit bis zur Rollstuhlpflichtigkeit einhergehen, von unmittelbarer Bedeutung und hoher Patientenrelevanz, zumal die 6MWD hinsichtlich der Zeit bis zum Verlust der Gehfähigkeit prognostisch relevant ist, d.h. sie stellt ein direktes Maß der Krankheitsprogression dar.

#### 4.4.2.4.3 TFTs (sekundärer Endpunkt) - RCT

In der kITT-Analyse mit BL-Adjustierung der Patienten 501-012 und 501-014 zeigten die Patienten der Ataluren-Gruppen insgesamt eine geringere Abnahme der Muskelfunktion bis Woche 48, als Patienten, die mit Vergleichstherapie behandelt wurden.

Der Übersichtlichkeit halber werden hier nur die signifikanten Ergebnisse der TFTs diskutiert (Zeit um vier Treppenstufen zu ersteigen).

Die Zeit, um vier Treppenstufen zu ersteigen, nahm bis Woche 48 in der Ataluren 10/10/20-Gruppe deutlich weniger stark zu als in der Placebo-Gruppe (-2.57 sec. vs. Placebo), und der Unterschied zwischen der Ataluren 10/10/20- und der Placebo-Gruppe war signifikant ( $p=0.0252$  nominal,  $p=0.0467$  adjustiert). Insgesamt unterstützen die Ergebnisse dieses TFTs die Primäranalyse des 6MWT. Die anderen TFTs (Zeit um vier Treppenstufen hinabzugehen, Zeit um 10 m zu gehen oder zu rennen, Zeit bis zum Aufstehen aus der Rückenlage) zeigten positive Trends zugunsten von Ataluren 10/10/20 verglichen mit Placebo.

Tabelle 4-44: Änderung der Dauer der TFTs von BL bis Woche 48 (kITT)

Gruppe	n BL/ Woche 48	M ( $\pm$ SD) BL (sec)	M ( $\pm$ SD) Woche 48 (sec)	$\Delta$ BL bis Woche 48, M ( $\pm$ SD) (sec) p-Wert	LSM ( $\pm$ SE) Woche 48 (sec)	$\Delta$ vs. Placebo, LSM ( $\pm$ SE) [95% KI] (sec)	p-Wert vs. Placebo Nominal Dunnetts
<i>Zeit bis zum Aufstehen aus Rückenlage</i>							
Placebo	57/56	11.50 (11.437)	14.60 (12.327)	3.24 (7.254)	14.85 (0.83)	-	-
ATL 10/10/ 20	57/57	10.80 (9.924)	14.03 (11.408)	3.23 (5.761) $p=1.0000^1$	14.79 (0.83)	-0.06 (1.16) [-2.35, 2.23]	0.9580 0.9980
ATL 20/20/ 40	60/59	12.25 (11.191)	15.39 (11.823)	3.00 (5.686) $p= 0.9696^1$	14.71 (0.81)	-0.15 (1.15) [-2.41, 2.12]	0.8985 0.9881

Gruppe	n BL/ Woche 48	M (±SD) BL (sec)	M (±SD) Woche 48 (sec)	Δ BL bis Woche 48, M (±SD) (sec) p-Wert	LSM (± SE) Woche 48 (sec)	Δ vs. Placebo, LSM (±SE) [95% KI] (sec)	p-Wert vs. Placebo Nominal Dunnetts
<i>Zeit, um 10 m zu gehen oder zu rennen</i>							
Placebo	57/56	6.72 (2.675)	9.85 (8.382)	3.17 (6.556)	10.95 (0.77)	-	-
ATL 10/10/ 20	57/57	7.45 (4.373)	9.13 (6.936)	1.68 (5.617) p=0.2980 <sup>1</sup>	9.59 (0.77)	-1.37 (1.07) [-3.48, 0.75]	0.2045 0.3393
ATL 20/20/ 40	60/59	7.44 (4.360)	10.20 (7.946)	2.74 (5.341) p=0.8901 <sup>1</sup>	10.28 (0.75)	-0.67 (1.07) [-2.78, 1.43]	0.5291 0.7521
<i>Zeit, um 4 Treppenstufen hinaufzusteigen</i>							
Placebo	57/56	6.01 (5.671)	10.81 (11.292)	4.82 (7.933)	12.11 (0.81)	-	-
ATL 10/10/ 20	57/57	6.94 (6.474)	9.33 (9.396)	2.39 (4.618) p=0.0905 <sup>1</sup>	9.54 (0.81)	<b>-2.57</b> (1.14) <b>[-4.82, -0.32]</b>	<b>0.0252</b> 0.0467
ATL 20/20/ 40	60/59	7.48 (7.463)	11.21 (10.838)	3.65 (6.665) p=0.5300 <sup>1</sup>	10.52 (0.80)	-1.59 (1.13) [-3.83, 0.65]	0.1626 0.2751
<i>Zeit, um 4 Treppenstufen hinabzusteigen</i>							
Placebo	57/56	5.49 (5.755)	9.56 (10.708)	4.06 (7.819)	10.55 (0.89)	-	-
ATL 10/10/ 20	57/57	6.08 (5.985)	8.49 (9.334)	2.41 (6.162) p=0.3605 <sup>1</sup>	8.80 (0.88)	-1.74 (1.24) [-4.20, 0.71]	0.1626 0.2752
ATL 20/20/ 40	60/59	6.66 (7.202)	9.73 (10.611)	3.05 (7.220) p=0.6619 <sup>1</sup>	9.29 (0.87)	-1.25 (1.24) [-3.69, 1.19]	0.3120 0.4935

<sup>1</sup> Dunnetts-Test vs. Placebo

Quelle (Studienbericht): S. 1155, S. 1215-1219, S. 1240, S. 1299-1303, S. **1329**, S. 1394-1398, S. 1419, S. 1478-1482

Die Analysen der LSM-Unterschiede basieren auf einem gemischten, augmentierten Modell (d.h. plus BL-Visite-Interaktion) für wiederholte Messungen der TFT-Parameter als abhängige Variable, als unabhängige Variablen schloss das Modell den BL-Wert, Behandlung, Visite, die Interaktion zwischen Behandlung und Visite, die Stratifizierungsfaktoren und die Interaktion zwischen dem BL-TFT-Wert und der Visite ein.

Auch im *post-hoc* Permutationstest basierend auf der MMRM-Analyse unter Berücksichtigung der Interaktion zwischen BL-Wert und Visite war der Unterschied zwischen Ataluren 10/10/20 und Placebo hinsichtlich der Zeit, die benötigt wurde um vier Stufen zu ersteigen, signifikant (LSM Differenz -2.5733 sec, nominal p=0.0169, adjustiert p=0.0358). Hinsichtlich der anderen TFTs zeigten sich mit dem obigen MMRM-Modell keine signifikanten Gruppenunterschiede.

Analog der 6MWD wurden 1000 Monte Carlo-Analysen *post-hoc* ausgeführt, um die Robustheit der primären Analyse der TFTs zu bestimmen (siehe Kapitel 4.4.2.4.2.2). Auch hinsichtlich der TFTs zeigte sich, dass es unwahrscheinlich war, dass die Ergebnisse zufällig oder durch Ausreißer bedingt waren (Abbildung 16; nur Daten zu TFT Treppen steigen gezeigt) (31).

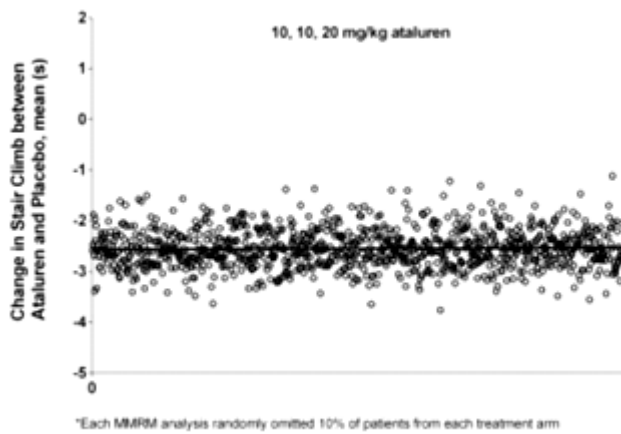


Abbildung 16: MMRM Monte Carlo-Analysen des TFTs (vier Treppenstufen ersteigen) in der Gesamtpopulation (31)

Jeder Kreis entspricht einer von insgesamt 1000 Monte Carlo-Analysen. „Ambulatory decline“-Gruppe: Patienten im Alter von  $\geq 7$  bis  $\leq 16$  Jahren mit einer initialen 6MWD von  $\geq 150$  m und einer nach Alter und Größe erwarteten 6MWD von  $\leq 80\%$  unter stabiler Kortikosteroid-Therapie.

### *Subgruppenanalysen*

In *post-hoc* Subgruppen-Analysen der TFTs wurde der Effekt von Ataluren 10/10/20 vs. Placebo bei Patienten mit einer BL 6MWD unter 350 m und *post-hoc* in der "ambulatory decline phase"-Gruppe evaluiert. In Übereinstimmung mit den Ergebnissen der 6MWD-Analysen zeigte sich bei den Subgruppen ein deutlich größerer Effektunterschied von Ataluren 10/10/20 vs. Placebo (Abbildung 17) (31).

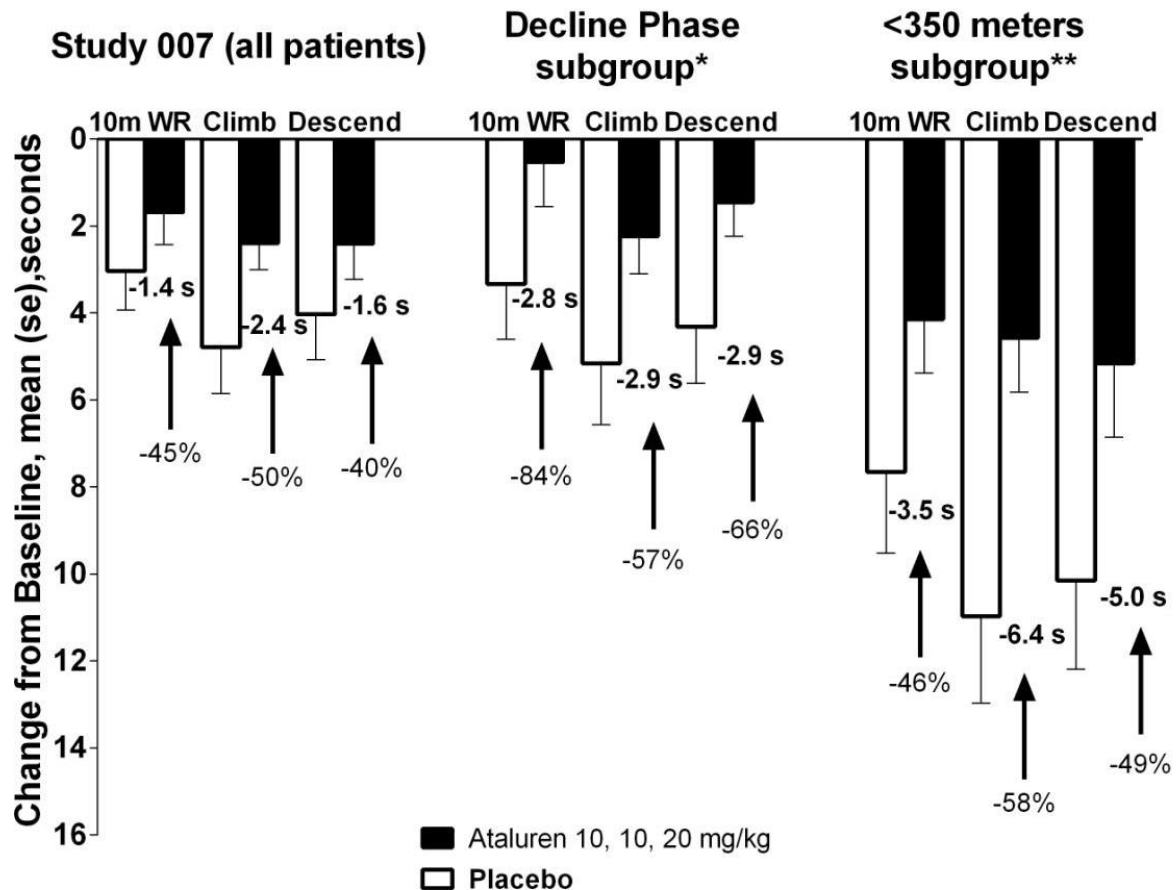


Abbildung 17: Ergebnisse der TFTs in der Gesamtpopulation, in der "ambulatory decline phase"-Gruppe und bei Patienten mit einer BL 6MWD <350 m (31). Werte sind Mittelwert-Unterschiede von Baseline zu Woche 48

Abk.: 10m WR, 10-meter-walk-run test (Zeit um 10 m zu gehen oder zu rennen); Climb, Zeit um vier Treppenstufen hinaufzugehen; Descend, Zeit um vier Treppenstufen hinabzugehen

### Fazit

Nach Einschätzung des CHMPs (2014) spiegeln TFTs typische Aktivitäten wider, die für DMD-Patienten wichtig sind und Prädiktoren für den Zeitpunkt darstellen, wann ein Patient seine Gehfähigkeit verliert (9). Der Zeit-Unterschied zwischen der Ataluren 10/10/20- und Placebo-Gruppe (2.57 sec), um vier Treppenstufen zu ersteigen, ist größer als der von McDonald et al. (2013) mittels zweier verteilungsbasierter statistischer Verfahren ermittelten MCIDs (2.1 bzw. 2.2 sec) und der Gruppenunterschied war statistisch signifikant (11). Entsprechend der von Escolar et al. (17) für diesen Endpunkt vorgeschlagenen MCID (0.4 log sec entsprechend 1.5 sec) ist der Unterschied zwischen der Ataluren 10/10/20- und der Placebo-Gruppe ebenfalls klinisch relevant.



In ihrer Ausrichtung stützen die Ergebnisse der TFTs (exklusive Aufstehen aus der Rückenlage) die Analyse des primären Endpunkts und demonstrieren in ihrer Gesamtheit eine bessere Muskelfunktion unter Ataluren-Behandlung verglichen mit der Placebo-Behandlung.

#### **4.4.2.4.4 Andere sekundäre Endpunkte - RCT**

Die Patientenrelevanz anderer Wirksamkeitsendpunkte der Studie wie z.B. die Zahl der unbeabsichtigten Stürze oder der Benutzung eines Rollstuhls erscheint selbsterklärend, selbst bei Vorliegen nur minimaler Änderungen. Es liegen allerdings keine validen Daten vor, welche Änderungen dieser Parameter für DMD-Patienten klinisch relevant sein könnten. Daher werden diese Endpunkte nicht bei der Bewertung des Ausmaßes des Zusatznutzens von Ataluren berücksichtigt und lediglich zur Information kurz zusammenfassend dargestellt.

Hinsichtlich anderer sekundärer Endpunkte d.h.

- Patienten/Betreuer-berichtete gesundheitsbezogene Lebensqualität ermittelt im PedsQL
- der Zahl der Tage, an denen ein Rollstuhl nach Angabe der Patienten/Betreuer benutzt wurde
- Zahl der versehentlichen Stürze
- Muskelstärke gemessen mittels Myometrie (Beugung und Streckung der Ellbogen und Knie, Schulter-Abduktion),
- Aktivität beim Treppensteigen im Alltag gemessen mittels eines Pedometer-ähnlichen Instruments,

ergaben sich in der ITT-Analyse mehrheitlich positive Trends zugunsten von Ataluren verglichen mit Placebo, die in der Ataluren 10/10/20-Gruppe tendenziell stärker ausgeprägt waren als in der Ataluren 20/20/40-Gruppe. Das Signifikanzniveau wurde allerdings in keinem Fall erreicht.

In der Patienten-berichteten physischen Dimension des HRQoL Instruments PedsQL zeigte sich eine numerische Überlegenheit von Ataluren 10/10/20 gegenüber Vergleichstherapie in der ITT Population (Abbildung 18). Die Therapieeffekte waren in der „decline phase“ Gruppe und der Subpopulation von Patienten mit BL-Werten < 350 m größer, als in der Gesamtpopulation. Die Therapieunterschiede waren jedoch nicht signifikant; die wahrscheinlichsten Ursachen hierfür sind die heterogene Patientenpopulation, und dass das HRQoL Instrument in dieser Patientenpopulation nicht sehr sensitiv ist. Die eindeutigen Trends in der physischen Domäne des PedsQL unterstützen jedoch die Resultate für den Wirksamkeitsendpunkt.

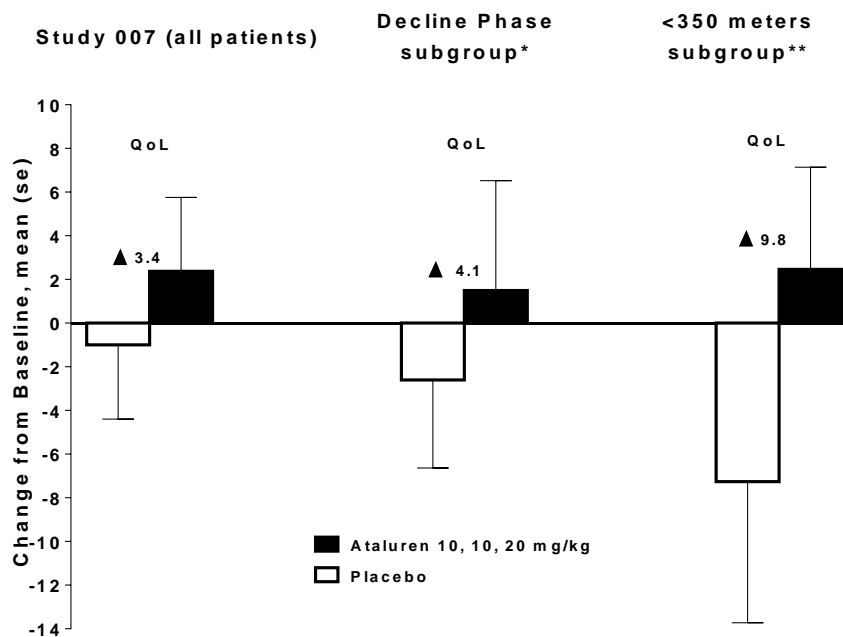


Abbildung 18: Mittelwertsveränderungen in der physischen Domäne des pedSQL in der Gesamtpopulation, in der "ambulatory decline phase"-Gruppe, und bei Patienten mit einer BL 6MWD <350 m

Die von den Patienten berichtete mittlere Prozentzahl von Tagen, an denen ein Rollstuhl verwendet wurde vergrößerte sich signifikant in der Placebo - Gruppe (11.5%; KI: 4.36-18.54) und der Ataluren 20/20/40 – Gruppe (9.0%; KI: 0.70-17.38), in der Ataluren 10/10/20 – Gruppe war keine signifikante Erhöhung messbar (4%; KI: -2.77-10.68).

Die Zahl versehentlicher Stürze wurde in einem Patienten-/Betreuer-Tagebuch erfasst. Zu BL betrug die Zahl der verseentlichen Stürze (Mittelwert  $\pm$  SD) 0.40 (0.597), 0.27 (0.480) und 0.54 (0.943) pro Tag in der Ataluren 20/20/40- (n=54), 10/10/20- (n=48) bzw. Placebo-Gruppe (n=48). Die mittlere Zahl der Stürze (SD) pro Tag stieg bis Woche 48 um 0.20 in der Placebo-Gruppe (1.282) an und sank um -0.06 (0.501) bzw. -0.10 (0.466) in der Ataluren 10/10/20- (p=0.2309 vs. Placebo, Dunnetts) bzw. 20/20/40-Gruppe (p=0.1252, ITT). In der kITT-Population war der Unterschied zwischen Ataluren 10/10/20 vs. Placebo in Woche 48 signifikant (LSM Differenz  $\pm$  SE -0.35 $\pm$ 0.160, 95% KI -0.66, -0.03, p=0.0319), während sich im Dunnetts-Test ein Trend zeigte (p=0.0582). Der Unterschied (LSM Differenz  $\pm$  SE) zwischen Ataluren 20/20/40 und Placebo in Woche 48 war in beiden Tests signifikant (-0.36 $\pm$ 0.155, 9% KI -0.66, -0.05; p=0.0229, Dunnetts p=0.0422). In der prä-spezifizierten GEE-Analyse ("generalised estimation equation") unter Einsatz von GENMOD in SAS betrug das relative Risiko (95% KI) der geschätzten Sturzrate in Woche 48 0.38 (0.16, 0.94) für den Vergleich von Ataluren 10/10/20 vs. Placebo und 0.41 (0.20, 0.85, p=0.0356) für den Vergleich von Ataluren 20/20/40 vs. Placebo (p=0.0171). Da 24 Ausgangswerte fehlten, wurde eine Sensitivitätsanalyse ausgeführt, in der die BL-Werte und Werte aus Woche 6 als mittlere BL-Werte eingegeben wurden, und wodurch sich die Zahl der fehlenden Werte auf 14 reduzierte. In dieser Analyse betrug das relative Sturzrisiko 0.49 (p=0.0229) bzw. 0.55

( $p=0.0400$ ) für den Vergleich von Ataluren 10/10/20 bzw. 20/20/40 mit Placebo. Zwischen der Zahl der unbeabsichtigten Stürze und der BL 6MWD ergab sich eine schwache Korrelation (Pearson Korrelationskoeffizient  $-0.09$ ), so dass diese Endpunkte wahrscheinlich unabhängig sind.

Die Zahl der Stufen, die die Patienten unter Alltagsbedingungen passierten, wurde mittels eines Pedometer-ähnlichen Instruments erfasst und es wurde die relative Zeit, während der der Patient keine (0 Stufen/min), niedrige (1-15 Stufen/min), mittlere (16-30 Stufen/min) oder hohe Aktivität ( $>30$  Stufen/min) zeigte, dokumentiert. Es zeigte sich ein nicht signifikanter Trend, dass die Patienten der Ataluren 10/10/20-Gruppe weniger Zeit ohne Aktivität und mehr Zeit in mittlerer Aktivität ( $<2\%$  Unterschied) verbrachten als die Patienten der Placebo-Gruppe.

### Fazit

Aufgrund von Unsicherheiten hinsichtlich der klinischen Relevanz einiger der oben dargestellten Parameter und der Tatsache, dass sich meist kein oder ein nur marginaler, positiver Effekt von Ataluren zeigte, werden diese Endpunkte in der vorliegenden Bewertung nicht berücksichtigt. Das CHMP erachtet die Häufigkeit akzidenteller Stürze als bei DMD relevanten Endpunkt, weist aber auf die hohe Zahl fehlender BL-Werte in der Studie und das Ungleichgewicht der BL-Werte milderer Ausprägung in der Ataluren 10/10/20-Gruppe hin. Das CHMP kommt zu dem Schluss, dass das Ergebnis wahrscheinlich keinen tatsächlichen Effekt von Ataluren widerspiegelt. Nach Einschätzung des CHMPs ist die Häufigkeit der Benutzung eines Rollstuhls für DMD-Patienten von hoher Bedeutung, da dies nicht nur auf die Gehfähigkeit selbst sondern auch auf Folgekomplikationen wie die respiratorische Funktion und das Risiko einer Skoliose Einfluss haben kann. Da in der Studie lediglich die Zahl der Tage, an denen ein Rollstuhl benutzt wurde, nicht aber die Dauer der Benutzung eines Rollstuhls dokumentiert wurde, wurden nach Einschätzung des CHMPs die Variabilität und die tatsächliche Nutzung eines Rollstuhls nicht vollständig erfasst. Das CHMP schätzt die Ergebnisse hinsichtlich der Aktivität im Alltag (Zahl der passierten Treppenstufen) als nicht eindeutig und das Ausmaß der Änderung (relative Zeit, in der die Patienten nicht aktiv waren:  $<2\%$ ) als nicht substantiiert ein (9).

Zusammenfassend haben die oben dargestellten z.T. sicherlich patientenrelevanten Endpunkte keinen Einfluss auf die Bewertung des Ausmaßes des Zusatznutzens von Ataluren 10/10/20.

#### **4.4.2.4.5 Sicherheit und Verträglichkeit - RCT**

Hinsichtlich der Bewertung der UEs ist zu berücksichtigen, dass zur Behandlung der DMD bislang keine zweckmäßige Vergleichstherapie zugelassen ist. Ein Zusatznutzen im Sinne einer Vermeidung schwerwiegender oder anderer Nebenwirkungen ist dementsprechend für Ataluren nicht belegbar. Zudem wird Ataluren nur unter streng kontrollierten klinischen Bedingungen von spezialisierten Ärzten mit Erfahrung in der Behandlung der DMD verabreicht.

In die Sicherheitspopulation (“as-treated”) wurden alle randomisierten Patienten, die irgendeines der Prüfpräparate entsprechend ihres Randomisierungscode erhalten hatten, eingeschlossen. Die Mehrzahl der Patienten aller Gruppen entwickelte mindestens ein behandlungsspezifisches UE (Placebo 98.2%, Ataluren 10/10/20 96.5%, Ataluren 20/20/40 95.0%). Die behandlungsspezifischen UEs waren mehrheitlich milden oder mäßigen Schweregrads. Ein Fortschreiten der Erkrankung war das häufigste schwere (Grad 3) UE (10.5% Placebo, 7.0% Ataluren 10/10/20, 8.3% Ataluren 20/20/40). Die Prüfarzte waren angehalten, bei Patienten, die ihre Gehfähigkeit während der Studie verloren, eine schwere (Grad 3) Progression der Erkrankung zu dokumentieren. Die Einschätzung des Zusammenhangs der UEs mit den Prüfpräparaten durch die Prüfarzte war in den Gruppen insgesamt vergleichbar. Nur 7 Patienten (n=3 Placebo, n=2 jeweils Ataluren 10/10/20 und 20/20/40) entwickelten ein schwerwiegendes behandlungsspezifisches UE. Keiner der Patienten brach die Studie aufgrund eines behandlungsspezifischen UEs vorzeitig ab und lebensbedrohliche oder fatale behandlungsspezifische UEs wurden nicht berichtet (Tabelle 4-45). Insgesamt 44 Patienten (n=9 bzw. n=22 Ataluren 20/20/40 bzw. 10/10/20, n=13 Placebo) brachen die Behandlung aufgrund eines oder mehrerer meist gastrointestinaler UEs temporär ab.

Tabelle 4-45: Überblick der behandlungsspezifischen UEs aus RCT (Sicherheitspopulation)

UEs, n(%)	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
Patienten mit $\geq 1$ UE	56 (98.2)	55 (96.5)	57 (95.0)
UEs nach Schwere <sup>1</sup>			
Grad 1 (mild)	21 (36.8)	16 (28.1)	20 (33.3)
Grad 2 (mäßig)	26 (45.6)	31 (54.4)	27 (45.0)
Grad 3 (schwer)	9 (15.8)	8 (14.0)	10 (16.7)
Grad 4 (lebensbedrohlich)	0	0	0
UEs nach Zusammenhang mit den Prüfpräparaten			
Kein Zusammenhang	14 (24.6)	8 (14.0)	11 (18.3)
Unwahrscheinlich	16 (28.1)	17 (29.8)	13 (21.7)
Möglich	20 (35.1)	25 (43.9)	29 (48.3)
Wahrscheinlich	6 (10.5)	5 (8.8)	4 (6.7)
Abbruch aufgrund eines UEs	0	0	0
Schwerwiegende UEs	3 (5.3)	2 (3.5)	2 (3.3)
Todesfälle	0	0	0

Quelle (Studienreport): S. 134

<sup>1</sup> Die Zahl (%) der Patienten basiert auf der schwersten durch den Prüfarzt berichteten Kategorie.

Die häufigsten behandlungsspezifischen UEs, die bei mindestens 5% der Patienten auftraten, sind in Tabelle 4-46 dargestellt. Die häufigsten behandlungsspezifischen UEs, die bei 20% und mehr Patienten in einer der Gruppen auftraten, waren Erbrechen (insgesamt 46.6%), Kopfschmerzen (29.3%), Diarrhoe (24.1%), Nasopharyngitis (20.7%), Pyrexie (19.0%), Husten (19.0%) und oberer Abdominalschmerz (17.8%). Behandlungsspezifische UEs, die bei mindestens 10% der Patienten in einer Gruppe auftraten und deren Inzidenz von Placebo zu Ataluren 10/10/20 respektive 20/20/40 zunahm, waren Übelkeit (12.3%, 14.0% bzw. 16.7%), Abdominalschmerz (7.0%, 12.3% bzw. 16.7%), Schmerz in den Extremitäten

(10.5%, 12.3% bzw. 13.3%), Flatulenz (7.0%, 8.8% bzw. 11.7%) und Verstopfung der Nase (7.0%, 8.8% bzw. 10.0%). Hinsichtlich der Verteilung der UEs innerhalb der verschiedenen Systemorganklassen (SOKs) ergaben sich keine auffälligen Gruppenunterschiede. Die meisten der behandlungsspezifischen UEs traten innerhalb der SOKs „Gastrointestinale Erkrankungen“ (Erbrechen, Diarrhoe und Abdominalschmerz), „Infektionen (und parasitäre Erkrankungen)“ (Nasopharyngitis und Infektionen der oberen Atemwege), „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ (Stürze bedingt durch die DMD und durch die Biopsien bedingter prozeduraler Schmerz), „Muskuloskeletale Erkrankungen“ (Schmerz in den Extremitäten und Rückenschmerzen) und „Erkrankungen des Nervensystems“ (Kopfschmerzen) auf. Eine Analyse der Inzidenzen behandlungsspezifischer UEs per 12-Wochen-Interval erbrachte keine Hinweise auf eine Zunahme der UEs während des Studienverlaufs (Daten nicht gezeigt).

Tabelle 4-46: Behandlungsspezifische UEs stratifiziert nach Organsystem aus RCT, die bei mehr als 5% der Patienten der Gesamtpopulation über 48 Wochen auftraten (Sicherheitspopulation)

SOK / UE MedDRA PT, n (%) <sup>1</sup>	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
<i>Erkrankungen des Gastrointestinaltrakts</i>	37 (64.9)	42 (73.7)	44 (73.3)
Erbrechen	22 (38.6)	32 (56.1)	27 (45.0)
Diarrhoe	14 (24.6)	11 (19.3)	17 (28.3)
Oberer Abdominalschmerz	9 (15.8)	9 (15.8)	13 (21.7)
Übelkeit	7 (12.3)	8 (14.0)	10 (16.7)
Abdominalschmerz	4 (7.0)	7 (12.3)	10 (16.7)
Flatulenz	4 (7.0)	5 (8.8)	7 (11.7)
Magenbeschwerden	0	4 (7.0)	5 (8.3)
<i>Allgemeine Erkrankungen und Erkrankungen am Applikationsort</i>	21 (36.8)	23 (40.4)	20 (33.3)
Pyrexie	12 (21.1)	14 (24.6)	7 (11.7)
Fortschreiten der Erkrankung	6 (10.5)	4 (7.0)	5 (8.3)
Asthenie	2 (3.5)	3 (5.3)	4 (6.7)
<i>Infektionen und parasitäre Erkrankungen</i>	43 (75.4)	38 (66.7)	39 (65.0)
Nasopharyngitis	13 (22.8)	13 (22.8)	10 (16.7)
Infektionen der oberen Atemwege	10 (17.5)	9 (15.8)	11 (18.3)
Influenza	8 (14.0)	6 (10.5)	7 (11.7)
Gastroenteritis	4 (7.0)	9 (15.8)	3 (5.0)
Rhinitis	2 (3.5)	6 (10.5)	3 (5.0)
Ohrinfektion	3 (5.3)	3 (5.3)	4 (6.7)
Virale Gastroenteritis	3 (5.3)	4 (7.0)	3 (5.0)
<i>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</i>	26 (45.6)	28 (49.1)	31 (51.7)
Sturz	7 (12.3)	11 (19.3)	6 (10.0)
Prozeduraler Schmerz	7 (12.3)	6 (10.5)	8 (13.3)
Kontusion	3 (5.3)	6 (10.5)	4 (6.7)
Verstauchung des Gelenks	1 (1.8)	4 (7.0)	4 (6.7)
<i>Untersuchungen</i>	4 (7.0)	10 (17.5)	6 (10.0)

SOK / UE MedDRA PT, n (%) <sup>1</sup>	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
Körpergewicht reduziert	1 (1.8)	5 (8.8)	3 (5.0)
<i>Stoffwechsel- und Ernährungsstörungen</i>	3 (5.3)	7 (12.3)	6 (10.0)
Verminderter Appetit	2 (3.5)	5 (8.8)	5 (8.3)
<i>Sklelettmuskulatur- und Bindegeweberkrankungen</i>	19 (33.3)	25 (43.9)	28 (46.7)
Schmerz in den Extremitäten	6 (10.5)	7 (12.3)	8 (13.3)
Rückenschmerz	5 (8.8)	9 (15.8)	6 (10.0)
Arthralgie	2 (3.5)	2 (3.5)	6 (10.0)
Muskelspasmen	5 (8.8)	3 (5.3)	1 (1.7)
Muskelschwäche	1 (1.8)	3 (5.3)	5 (8.3)
<i>Erkrankungen des Nervensystems</i>	17 (29.8)	25 (43.9)	18 (30.0)
Kopfschmerzen	14 (24.6)	22 (38.6)	15 (25.0)
Schwindel („dizziness“)	4 (7.0)	3 (5.3)	3 (5.0)
<i>Erkrankungen der Atemwege, des Brustraums und Mediastinums</i>	18 (31.6)	20 (35.1)	22 (36.7)
Husten	11 (19.3)	9 (15.8)	13 (21.7)
Nasenverstopfung	4 (7.0)	5 (8.8)	6 (10.0)
Schmerzen im Oropharynx	4 (7.0)	6 (10.5)	4 (6.7)
Rhinorrhoe	6 (10.5)	4 (7.0)	0
<i>Erkrankungen der Haut und des Unterhautzellgewebes</i>	18 (31.6)	19 (33.3)	14 (23.3)
Ausschlag („rash“)	5 (8.8)	4 (7.0)	8 (13.3)
Narbe	3 (5.3)	4 (7.0)	5 (8.3)

Quelle (Studienbericht): S. 138-139

<sup>1</sup> Patienten, bei denen das gleiche UE mehrmals auftrat, wurden nur einmal genannt.

Die häufigsten aus Sicht der Prüfarzte möglicherweise oder wahrscheinlich mit der Behandlung in Zusammenhang stehenden UEs ( $\geq 5\%$  über alle Behandlungsgruppen) waren Erbrechen (17.8%), Diarrhoe (12.6%), Schmerzen im oberen Abdomen (10.3%), Flatulenz (8.6%), Übelkeit (8.0%), Kopfschmerzen (7.5%) und verminderter Appetit (6.3%). Insgesamt war die Inzidenz der mit der Behandlung in Zusammenhang stehenden UEs zwischen den Gruppen vergleichbar, allerdings nahmen die UEs Erbrechen, Flatulenz, Magenbeschwerden und Ermüdung von Placebo über Ataluren 10/10/20 respektive 20/20/40 zu (Tabelle 4-47). Bei nur einem Patienten der Ataluren 20/20/40-Gruppe wurde aufgrund eines möglicherweise Behandlungs-bedingten UEs (Grad 2 „rash“) die Behandlung zeitweise unterbrochen und die Dosis zeitweise reduziert.

Tabelle 4-47: Inzidenz der behandlungsspezifischen mit der Behandlung in Zusammenhang stehenden UEs aus RCT, die bei mehr als einem Patienten auftraten (Sicherheitspopulation)

SOK / UE MedDRA PT, n (%) <sup>1</sup>	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
<i>Patienten mit <math>\geq 1</math> mit der Behandlung in Zusammenhang stehendem UE</i>	26 (45.6)	30 (52.6)	33 (55.0)
<i>Erkrankungen des Gastrointestinaltrakts</i>	18 (31.6)	21 (36.8)	28 (46.7)
Abdominale Beschwerden	1 (1.8)	0	2 (3.3)

<b>SOK / UE MedDRA PT, n (%)<sup>1</sup></b>	<b>Placebo (n=57)</b>	<b>Ataluren 10/10/20 (n=57)</b>	<b>Ataluren 20/20/40 (n=60)</b>
Abdominalschmerz	3 (5.3)	1 (1.8)	4 (6.7)
Oberer Abdominalschmerz	5 (8.8)	5 (8.8)	8 (13.3)
Abdominale Druckempfindlichkeit	1 (1.8)	0	2 (3.3)
Aerophagie	2 (3.5)	0	1 (1.7)
Konstipation	1 (1.8)	1 (1.8)	2 (3.3)
Diarrhoe	7 (12.3)	3 (5.3)	12 (20.0)
Flatulenz	3 (5.3)	5 (8.8)	7 (11.7)
Häufiger Stuhlgang / -drang	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)
<i>Allgemeine Erkrankungen und Erkrankungen am Applikationsort</i>	2 (3.5)	2 (3.5)	5 (8.3)
Ermüdung („fatigue“)	0	1 (1.8)	2 (3.3)
Pyrexie	0	1 (1.8)	1 (1.7)
<i>Untersuchungen</i>	3 (5.3)	6 (10.5)	1 (1.7)
Farbe des Urins nicht normal	1 (1.8)	1 (1.8)	0
Körpergewicht vermindert	1 (1.8)	4 (7.0)	1 (1.7)
<i>Stoffwechsel- und Ernährungsstörungen</i>	2 (3.5)	5 (8.8)	4 (6.7)
Verminderter Appetit	2 (3.5)	5 (8.8)	4 (6.7)
<i>Skelettmuskulatur- und Bindegeweberkrankungen</i>	0	3 (5.3)	2 (3.3)
Schmerz in den Extremitäten	0	2 (3.5)	1 (1.7)
<i>Erkrankungen des Nervensystems</i>	4 (7.0)	8 (14.0)	5 (8.3)
Schwindel („dizziness“)	1 (1.8)	1 (1.8)	1 (1.7)
Kopfschmerzen	2 (3.5)	7 (12.3)	4 (6.7)
<i>Erkrankungen der Niere und der Harnwege</i>	2 (3.5)	6 (10.5)	1 (1.7)
Enuresis	0	2 (3.5)	0
Pollakisurie	1 (1.8)	1 (1.8)	1 (1.7)
Nierenzyste	1 (1.8)	1 (1.8)	0
<i>Erkrankungen der Atemwege, des Brustraums und Mediastinums</i>	0	4 (7.0)	3 (5.0)
Husten	0	2 (3.5)	2 (3.3)
Epistaxis	0	1 (1.8)	1 (1.7)
<i>Erkrankungen der Haut und des Unterhautzellgewebes</i>	2 (3.5)	1 (1.8)	2 (3.3)
Erythem	0	1 (1.8)	1 (1.7)
Ausschlag („rash“)	1 (1.8)	0	2 (3.3)
<i>Gefäßerkrankungen</i>	2 (3.5)	3 (5.3)	0
Hypertension	1 (1.8)	3 (5.3)	0

Quelle (Studienbericht PTC124-GD-007-DMD): S. 148-149

<sup>1</sup> UEs, die von den Prüfarzten als möglicherweise oder wahrscheinlich mit der Behandlung in Zusammenhang stehend klassifiziert wurden. UEs, die mit der Behandlung in Zusammenhang standen und bei einem Patienten mehrmals auftraten, wurden jeweils nur einmal gelistet.

Die Inzidenz schwerer UEs (Grad 3) war in den Gruppen vergleichbar (15.8% Placebo, 14.0% bzw. 16.7% Ataluren 10/10/20 bzw. 20/20/40; Tabelle 4-48). Bei den meisten der schweren UEs handelte es sich um Einzelfälle und eine Krankheitsprogression wurde in allen Gruppen als häufigste Grad 3-UE berichtet.

Tabelle 4-48: Inzidenz der behandlungsspezifischen schweren UEs aus RCT (Sicherheitspopulation)

<b>SOK / UE MedDRA PT, n (%)<sup>1</sup></b>	<b>Placebo (n=57)</b>	<b>Ataluren 10/10/20 (n=57)</b>	<b>Ataluren 20/20/40 (n=60)</b>
<i>Patienten mit <math>\geq 1</math> schwerem UE (Grad 3)</i>	9 (15.8)	8 (14.0)	10 (16.7)
<i>Erkrankungen des Ohres und Labyrinths</i>	0	0	1 (1.7)
Ohrenschmerzen	0	0	1 (1.7)
<i>Erkrankungen des Gastrointestinaltrakts</i>	1 (1.8)	2 (3.5)	1 (1.7)
Abdominalschmerzen	0	1 (1.8)	0
Konstipation	0	1 (1.8)	0
Diarrhoe	0	0	1 (1.7)
Ileus	0	1 (1.8)	0
Erbrechen	1 (1.8)	0	1 (1.7)
<i>Allgemeine Erkrankungen und Erkrankungen am Applikationsort</i>	6 (10.5)	4 (7.0)	5 (8.3)
Fortschreiten der Erkrankung	6 (10.5)	4 (7.0)	5 (8.3)
<i>Infektionen und parasitäre Erkrankungen</i>	0	0	1 (1.7)
Virale Gastroenteritis	0	0	1 (1.7)
<i>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</i>	1 (1.8)	2 (3.5)	2 (3.3)
Kontusion	0	1 (1.8)	0
Sturz	0	1 (1.8)	0
Fraktur des Femurs	1 (1.8)	0	0
Fraktur des Fußes	0	1 (1.8)	0
Unterschenkelfraktur	0	0	1 (1.7)
Handgelenksfraktur	0	0	1 (1.7)
<i>Skellettmuskulatur- und Bindegeweberkrankungen</i>	0	2 (3.5)	1 (1.7)
Muskuloskeletaler Brustschmerz	0	1 (1.8)	0
Schmerzen in den Extremitäten	0	1 (1.8)	0
Osteoporose	0	0	1 (1.7)
<i>Erkrankungen des Nervensystems</i>	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)
<i>Chirurgische und medizinische Eingriffe</i>	0	1 (1.8)	0
Wurzelbehandlung	0	1 (1.8)	0



Quelle (Studienbericht PTC124-GD-007-DMD): S. 141-142

UEs, die bei einem Patienten mehrmals auftraten, wurden jeweils nur einmal gelistet. Die Schweregrade wurden von den Prüfern auf Basis der CTCAE oder protokollgemäßer Definitionen bestimmt.

In der Placebo-Gruppe wurden zwei möglicherweise mit der Behandlung in Zusammenhang stehende schwerwiegende UEs berichtet (Grand-Mal-Anfall [re-klassifiziert: unwahrscheinlicher Zusammenhang] und Schmerzen im Abdomen). Jeweils zwei Patienten beider Ataluren-Gruppen entwickelten ein schwerwiegendes UE (Tabelle 4-49), die von den Prüfern als nicht mit der Behandlung in Zusammenhang stehend klassifiziert wurden oder ein Zusammenhang unwahrscheinlich war. Der Fall einer paroxysmalen supraventrikulären Tachykardie unter Ataluren 20/20/40 wurde als wahrscheinlich nicht mit der Behandlung in Zusammenhang stehend bewertet, da der Patient bereits früher ähnliche Episoden hatte.

Tabelle 4-49: Inzidenz der behandlungsspezifischen schwerwiegenden UEs aus RCT (Sicherheitspopulation)

SOK / UE MedDRA PT, n (%) <sup>1</sup>	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
<i>Patienten mit ≥1 schwerwiegendem UE</i>	3 (5.3)	2 (3.5)	2 (3.3)
<i>Herzerkrankungen</i>	0	0	1 (1.7)
Supraventrikuläre Tachykardie	0	0	1 (1.7)
<i>Erkrankungen des Gastrointestinaltrakts</i>	1 (1.8)	0	0
Schmerz im Abdomen	1 (1.8)	0	0
<i>Infektionen und parasitäre Erkrankungen</i>	1 (1.8)	1 (1.8)	0
Appendizitis	0	1 (1.8)	0
Influenza	1 (1.8)	0	0
Varicella	1 (1.8)	0	0
<i>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</i>	1 (1.8)	0	1 (1.7)
Fraktur des Femurs	1 (1.8)	0	0
Unterschenkelfraktur	0	0	1 (1.7)
<i>Stoffwechsel- und Ernährungsstörungen</i>	0	1 (1.8)	0
Dehydration	0	1 (1.8)	0
<i>Erkrankungen des Nervensystems</i>	1 (1.8)	0	0
Grand-Mal-Anfall	1 (1.8)	0	0

Quelle (Studienbericht): S. 2556

<sup>1</sup> Patienten, die zwei oder mehr UEs desselben PTs berichteten, wurden nur einmal gelistet. Ein Patient, der/die zwei oder mehr UEs mit verschiedenen PTs innerhalb derselben SOK berichtete, wurde innerhalb der SOK (gesamt) nur einmal gezählt. Ein Patient, der/die zwei oder mehr UEs in verschiedenen SOKs berichtete, wurde hinsichtlich der Zahl der Patienten nur einmal gezählt.

Da Patienten mit DMD häufig mit Kortikosteroiden behandelt werden und auch in der RCT die Mehrzahl der Patienten Kortikosteroide als Begleitmedikation erhielten (71.3%), wurde die Inzidenz von UEs in Abhängigkeit von der Behandlung mit Kortikosteroiden analysiert. Hinsichtlich dieser Analyse ist zu beachten, dass die Gruppengrößen nicht ausgeglichen

waren (Kortikosteroide: ja / nein) und die Patienten unterschiedliche Kortikosteroide und Regimes erhielten. Der deutlichste Unterschied zeigte sich beim Auftreten von Kopfschmerzen, welche bei 35.5% vs. 14.0% der Patienten, die Kortikosteroide erhielten bzw. nicht erhielten, auftrat. Die mit Kortikosteroiden behandelten Patienten berichteten öfter über Rückenschmerzen (15.3% vs. 2.0%), Schmerzen in den Extremitäten (13.7% vs. 8.0%) und Muskelkrämpfe (6.5% vs. 2.0%). Hypertension trat als UE nur bei den Kortikosteroid-behandelten Patienten auf (4.0% vs. 0%). Folgende UEs traten bei den Patienten, die keine Kortikosteroide erhielten, häufiger auf als unter Begleitmedikation mit Kortikosteroiden: Pyrexie (26.0% vs. 16.1%), Fortschreiten der Erkrankung (12.0% vs. 7.3%), Rhinitis (10.0% vs. 4.8%), virale Gastroenteritis (10.0% vs. 4.0%), Rhinorrhoe (8.0% vs. 4.8%), Lymphadenopathie (6.0% vs. 0.8%), Sinusitis (6.0% vs. 3.2%), virale Infektionen (6.0% vs. 3.2%), Haut-Läsionen (6.0% vs. 0%) und saisonale Allergien (10.0% vs. 0%).

Da die Mehrzahl der Patienten (etwa 90%) kaukasischer Herkunft war, lassen die Daten keine Rückschlüsse hinsichtlich des Nebenwirkungsprofils von Ataluren in Abhängigkeit von der Herkunft zu. Zwischen den Altersgruppen und dem BL 6MWD-Status (<350 m vs. ≥350 m) ergaben sich keine auffälligen Unterschiede hinsichtlich des Musters der UEs.

### Fazit

Das UE-Profil von Ataluren war hinsichtlich Inzidenz, Art, Schwere und möglichem Kausalzusammenhang mit den Prüfpräparaten insgesamt demjenigen unter Placebo vergleichbar. Übelkeit, Abdominalschmerzen, Schmerzen in den Extremitäten, Flatulenz und Verstopfung der Nase traten unter Ataluren verglichen mit Placebo etwas häufiger auf. Bei keinem der wenigen schwerwiegenden UEs wurde ein Zusammenhang mit der Therapie mit Ataluren vermutet und keiner der Patienten brach die Therapie aufgrund von UEs ab. Lebensbedrohliche oder fatale UEs wurden nicht berichtet. In dem Beurteilungsbericht zu Ataluren schloss das CHMP (2014), dass die häufigsten mit der Behandlung in Zusammenhang stehenden UEs (Erbrechen, Diarrhoe, Schmerz im oberen Abdomen, Flatulenz, Nausea, Kopfschmerzen und verminderter Appetit) bei einer äußerst belastenden und lebensbedrohlichen Erkrankung wie der DMD keine bedeutenden Sicherheitsbedenken hervorrufen. Nach Ansicht des CHMPs (2014) sind die identifizierten Risiken von Ataluren akzeptabel (9).

Da bislang in Deutschland keine Referenztherapie bei DMD zugelassen ist, ist es nicht möglich, einen Nutzen von Ataluren im Sinne einer Vermeidung (schwerwiegender) Nebenwirkungen oder deren Ausmaß zu belegen. Es ergibt sich für Ataluren 10/10/20 bei der Zielpopulation kein relevantes Schadenspotential, welches sich u.U. negativ auf die Bewertung des Ausmaßes des Zusatznutzens auswirken könnte. Zudem waren die unter Ataluren berichteten UEs mehrheitlich milder bis mittelschwerer Ausprägung und keines der schwerwiegenden UEs stand mit der Ataluren-Behandlung in Zusammenhang. Demgegenüber steht ein Zusatznutzen von Ataluren 10/10/20 hinsichtlich der Abschwächung eines schwerwiegenden bzw. schweren Symptoms (Erhalt der Gehfähigkeit).

#### 4.4.2.4.6 Zusammenfassung der Ergebnisse aus RCT

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 (Morbidität), d.h. eine deutliche Verbesserung des therapielevanten 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.

Tabelle 4-50: Übersicht der Ergebnisse der patientenrelevanten Endpunkte aus RCT: Ataluren 10/10/20 vs. Placebo (kITT)

Patientenrelevanter Endpunkt	Ataluren 10/10/20 vs. Placebo	p-Wert
Änderung der 6MWD (MMRM, m)	Mittelwert 31.7 m (95% KI 5.1, 58.3)	0.0197 (nominal) 0.0367 (Dunnetts)
Änderung der 6MWD (Permutationstest)		0.0281 (nominal) 0.0561 (adjustiert)
Zeit bis zu einer 10%ige Verschlechterung der 6MWD (Cox-proportionales Hazard-Modell)	HR 0.508 (95% KI 0.273, 0.946)	p=0.0326 (nominal) p=0.0652 (adjustiert)
Anteil der Patienten mit $\geq 10\%$ iger Verschlechterung der 6MWD (CMH-Test)	OR 0.3 (95% KI 0.1-0.9)	p=0.0423
Anteil der Patienten mit $\geq 0.5 \times$ SD Verschlechterung der 6MWD (logistische Regression)	OR 0.34 (95% KI 0.135-0.858)	p=0.0224
Zeit, die benötigt wird um 4 Treppenstufen zu überwinden (sec)		
(MMRM)	LSM Differenz -2.57 s (95% KI -4.82, -0.32)	p=0.0252 (nominal) p=0.0467 (adjustiert)
MMRM Permutationstest	LSM-Differenz -2.5733	p=0.0169 (nominal) p=0.0358 (adjustiert)

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

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

#### 4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-52: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Behandlung der genetisch gesicherten nmDMD bei gehfähigen Patienten im Alter von mindestens 5 Jahren	Beträchtlich

#### 4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

##### 4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht zutreffend.

#### **4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen**

*Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.*

Nicht zutreffend

#### **4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen**

*Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.*

Nicht zutreffend

#### **4.5.4 Verwendung von Surrogatendpunkten**

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005<sup>13</sup>, Molenberghs 2010<sup>14</sup>). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006<sup>15</sup>) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten

<sup>13</sup> Burzykowski T (Ed.): *The evaluation of surrogate endpoints*. New York: Springer; 2005.

<sup>14</sup> Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: *A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials*. *Stat Methods Med Res* 2010; 19(3): 205-236.

<sup>15</sup> Burzykowski T, Buyse M. *Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation*. *Pharm Stat* 2006; 5(3): 173-186.

Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006<sup>16</sup>) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

*Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.*

Nicht zutreffend. Die in der Studie PTC124-GD-007-DMD erhobenen Surrogatendpunkte wie z.B. die Aktivität der Kreatinkinase im Serum oder die muskuläre Dystrophinexpression werden nicht zur Bewertung des Ausmaßes des Zusatznutzens von Ataluren herangezogen.

#### **4.6 Liste der eingeschlossenen Studien**

*Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).*

Atkinson L, Elfring G, Reha A, Miller LL. A Phase 2b Efficacy and Safety Study of PTC124 in Subjects with Nonsense-Mutation-Mediated Duchenne and Becker Muscular Dystrophy. Sponsor: PTC Therapeutics, South Plainfield, USA, Clinical Protocol PTC124-GD-007-DMD Version 3.0, 04 March 2009.

Bushby K, Finkel R, Wong B, Barohn R, Campbell C, Comi GP, et al. Ataluren treatment of patients with nonsense mutation dystrophinopathy. *Muscle Nerve*. 2014 Oct;50(4):477–87

Elfring G, Atkinson L, Northcutt VJ, Miller LL. A Phase 2b Efficacy and Safety Study of PTC124 in Subjects with Nonsense-Mutation-Mediated Duchenne and Becker Muscular

---

<sup>16</sup> Weir CJ, Walley RJ. *Statistical evaluation of biomarkers as surrogate endpoints: a literature review*. *Stat Med* 2006; 25(2): 183-203.

Dystrophy. Sponsor: PTC Therapeutics, South Plainfield, USA, Statistical Analysis Plan PTC124-GD-007-DMD Version 3.0, 14 January 2010.

PTC Therapeutics. Phase 2b Study of PTC124 in Duchenne/Becker Muscular Dystrophy (DMD/BMD). NCT00592553 Registereintrag [Internet]. 2013 Sep 06 update [Zugriff 2014 Sep 19];[3 Seiten].

URL:<http://clinicaltrials.gov/ct2/show/study/NCT00592553?term=ataluren+007&rank=2>

PTC Therapeutics, Inc. EudraCT Protokoll [Internet]. 17 March 2008 [Zugriff 2014 Sep 19];[1 Seite].

URL:<https://www.clinicaltrialsregister.eu/ctr-search/search?query=Ataluren+2007-005478-29>

Wong B, Barth J. A. ATALUREN (PTC124). Phase 2b efficacy and safety study of ptc124 in subjects with nonsense-mutation-mediated Duchenne and Becker muscular dystrophy. Sponsor: PTC Therapeutics, South Plainfield, USA, Clinical Study Report PTC124-GD-007-DMD, 26 February 2011 (revised 31 Aug 2012).

#### 4.7 Referenzliste

*Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.*

1. Bushby K, Finkel R, Birnkrant DJ, Case LE, Clemens PR, Cripe L, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. *Lancet Neurol*. 2010 Jan;9(1):77–93.
2. Davies K, Perkins. Recent advances in Duchenne muscular dystrophy. *Degener Neurol Neuromuscul Dis* [Internet]. 2012 Oct [cited 2014 Sep 21];141. Available from: <http://www.dovepress.com/recent-advances-in-duchenne-muscular-dystrophy-peer-reviewed-article-DNND>
3. European Medicines Agency. Public summary of opinion on orphan designation 3-[5-(2-fluoro-phenyl)-[1,2,4]oxadiazole-3-yl]-benzoic acid for the treatment of Duchenne muscular dystrophy [Internet]. 2005 [cited 2014 Sep 26]. Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Orphan\\_designation/2009/10/WC500005403.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Orphan_designation/2009/10/WC500005403.pdf)
4. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [Internet]. 2014. Available from: [https://www.g-ba.de/downloads/62-492-873/VerfO\\_2014-03-20.pdf](https://www.g-ba.de/downloads/62-492-873/VerfO_2014-03-20.pdf)
5. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden. Entwurf für Version 4.2 vom 18.06.2014 [Internet]. 2014. Available from: [https://www.iqwig.de/download/IQWiG\\_Methoden\\_Entwurf-fuer-Version-4-2.pdf](https://www.iqwig.de/download/IQWiG_Methoden_Entwurf-fuer-Version-4-2.pdf)

6. Atkinson L, Elfring G, Reha R, Miller L. A Phase 2b Efficacy and Safety Study of PTC124 in Subjects with Nonsense-Mutation-Mediated Duchenne and Becker Muscular Dystrophy. Sponsor: PTC Therapeutics, South Plainfield, USA, Clinical Protocol PTC124-GD-007-DMD Version 3.0. 2009.
7. Elfring G, Atkinson L, Northcutt V, Miller L. A Phase 2b Efficacy and Safety Study of PTC124 in Subjects with Nonsense-Mutation-Mediated Duchenne and Becker Muscular Dystrophy. Sponsor: PTC Therapeutics, South Plainfield, USA, Statistical Analysis Plan PTC124-GD-007-DMD Version 3.0. 2010.
8. Wong B, Barth JA. ATALUREN (PTC124). Phase 2b efficacy and safety study of ptc124 in subjects with nonsense-mutation-mediated Duchenne and Becker muscular dystrophy. Sponsor: PTC Therapeutics, South Plainfield, USA, Clinical Study Report PTC124-GD-007-DMD, 26 February 2011 (revised 31 Aug 2012). 2011.
9. European Medicines Agency. Translarna, INN-ataluren - Assessment report for initial marketing authorisation application [Internet]. 2014 [cited 2014 Sep 26]. Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/EPAR\\_-\\_Public\\_assessment\\_report/human/002720/WC500171816.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002720/WC500171816.pdf)
10. Gemeinsamer Bundesausschuss. Erstellung und Einreichung eines Dossiers zur Nutzenbewertung gemäß § 35a SGB V. Format und Gliederung des Dossiers, einzureichende Unterlagen, Vorgaben für technische Standards [Internet]. 2013. Available from: [https://www.g-ba.de/downloads/17-98-3518/2013-04-18\\_An12\\_1\\_Erstellung-Einreichung-Dossier.pdf](https://www.g-ba.de/downloads/17-98-3518/2013-04-18_An12_1_Erstellung-Einreichung-Dossier.pdf)
11. McDonald CM, Henricson EK, Abresch RT, Florence J, Eagle M, Gappmaier E, et al. The 6-minute walk test and other clinical endpoints in duchenne muscular dystrophy: reliability, concurrent validity, and minimal clinically important differences from a multicenter study. *Muscle Nerve*. 2013 Sep;48(3):357–68.
12. McDonald CM, Henricson EK, Abresch RT, Florence JM, Eagle M, Gappmaier E, et al. The 6-minute walk test and other endpoints in Duchenne muscular dystrophy: longitudinal natural history observations over 48 weeks from a multicenter study. *Muscle Nerve*. 2013 Sep;48(3):343–56.
13. Henricson E, Abresch R, Han JJ, Nicorici A, Goude Keller E, de Bie E, et al. The 6-minute walk test and person-reported outcomes in boys with duchenne muscular dystrophy and typically developing controls: longitudinal comparisons and clinically-meaningful changes over one year. *PLoS Curr*. 2013;5.
14. McDonald CM, Abresch RT, Carter GT, Fowler WM, Johnson ER, Kilmer DD, et al. Profiles of neuromuscular diseases. Duchenne muscular dystrophy. *Am J Phys Med Rehabil Assoc Acad Physiatr*. 1995 Oct;74(5 Suppl):S70–92.
15. Mendell JR, Moxley RT, Griggs RC, Brooke MH, Fenichel GM, Miller JP, et al. Randomized, double-blind six-month trial of prednisone in Duchenne's muscular dystrophy. *N Engl J Med*. 1989 Jun 15;320(24):1592–7.



16. Griggs RC, Moxley RT, Mendell JR, Fenichel GM, Brooke MH, Pestronk A, et al. Prednisone in Duchenne dystrophy. A randomized, controlled trial defining the time course and dose response. Clinical Investigation of Duchenne Dystrophy Group. Arch Neurol. 1991 Apr;48(4):383–8.
17. Escolar DM, Hache LP, Clemens PR, Cnaan A, McDonald CM, Viswanathan V, et al. Randomized, blinded trial of weekend vs daily prednisone in Duchenne muscular dystrophy. Neurology. 2011 Aug 2;77(5):444–52.
18. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003 May;41(5):582–92.
19. U.S. National Institute of Health. Phase 2b Study of PTC124 in Duchenne/Becker Muscular Dystrophy (DMD/BMD) - Full Text View - ClinicalTrials.gov [Internet]. [cited 2014 Sep 26]. Available from: <http://clinicaltrials.gov/show/NCT00592553>
20. EU Clinical Trials Register. Clinical trials for Ataluren 2008-007648-32 [Internet]. [cited 2014 Sep 21]. Available from: <https://www.clinicaltrialsregister.eu/ctr-search/search?query=Ataluren+2008-007648-32>
21. Search of: NCT01826487 - List Results - ClinicalTrials.gov [Internet]. [cited 2014 Sep 21]. Available from: <http://clinicaltrials.gov/ct2/results?term=NCT01826487&Search=Search>
22. Bushby K, Finkel R, Wong B, Barohn R, Campbell C, Comi GP, et al. Ataluren treatment of patients with nonsense mutation dystrophinopathy. Muscle Nerve. 2014 Oct;50(4):477–87.
23. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med. 2002 Jul 1;166(1):111–7.
24. McDonald CM, Henricson EK, Han JJ, Abresch RT, Nicorici A, Elfring GL, et al. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. Muscle Nerve. 2010 Apr;41(4):500–10.
25. Goemans N, Kirschner J, Mercuri E. New Perspectives in the Management of Duchenne Muscular Dystrophy [Internet]. 2014 [cited 2014 Sep 26]. Available from: <http://www.touchneurology.com/articles/new-perspectives-management-duchenne-muscular-dystrophy>
26. European Medicines Agency. Draft Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy [Internet]. 2013 [cited 2014 Sep 26]. Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2013/03/WC500139508.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/03/WC500139508.pdf)

27. Mazzone ES, Pane M, Sormani MP, Scalise R, Berardinelli A, Messina S, et al. 24 month longitudinal data in ambulant boys with Duchenne muscular dystrophy. *PloS One*. 2013;8(1):e52512.
28. McDonald CM, Henricson EK, Abresch RT, Han JJ, Escolar DM, Florence JM, et al. The cooperative international neuromuscular research group Duchenne natural history study-- a longitudinal investigation in the era of glucocorticoid therapy: design of protocol and the methods used. *Muscle Nerve*. 2013 Jul;48(1):32–54.
29. Peay HL, Hollin I, Fischer R, Bridges JFP. A community-engaged approach to quantifying caregiver preferences for the benefits and risks of emerging therapies for Duchenne muscular dystrophy. *Clin Ther*. 2014 May;36(5):624–37.
30. McDonald CM. Outcome measures in ambulatory boys with DMD (London, UK 21 June 2013) [Internet]. 2013. Available from: [http://www.treat-nmd.eu/downloads/file/meetings/2013/workshop/Session1/McDonald\\_AmbulatoryEndpoints.pdf](http://www.treat-nmd.eu/downloads/file/meetings/2013/workshop/Session1/McDonald_AmbulatoryEndpoints.pdf)
31. PTC Therapeutics, Ltd. Ataluren (PTC124) for Nonsense Mutation Duchenne Muscular Dystrophy. Grounds for Reexamination of the CHMP Negative Opinion for the Conditional Marketing Authorisation of Translarna. Interner Bericht. 2014.
32. Pane M, Mazzone ES, Sivo S, Sormani MP, Messina S, D Amico A, et al. Long term natural history data in ambulant boys with duchenne muscular dystrophy: 36-month changes. *PloS One*. 2014;9(10):e108205.
33. Bushby K, Connor E. Clinical outcome measures for trials in Duchenne muscular dystrophy: report from International Working Group meetings. *Clin Investig*. 2011 Sep;1(9):1217–35.
34. McDonald CM. Natural history data argues against strength and function as co-primary endpoints in Duchenne muscular dystrophy [Internet]. 2013 [cited 2014 Sep 26]. Available from: [http://www.treat-nmd.eu/downloads/file/meetings/2013/workshop/Session2/McDonald\\_Strength\\_Function.pdf](http://www.treat-nmd.eu/downloads/file/meetings/2013/workshop/Session2/McDonald_Strength_Function.pdf)
35. European Medicines Agency. ZUSAMMENFASSUNG DER MERKMALE DES ARZNEIMITTELS [Internet]. 2014 [cited 2014 Sep 26]. Available from: [http://www.ema.europa.eu/docs/de\\_DE/document\\_library/EPAR\\_-\\_Product\\_Information/human/002720/WC500171813.pdf](http://www.ema.europa.eu/docs/de_DE/document_library/EPAR_-_Product_Information/human/002720/WC500171813.pdf)

**Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche**

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

<b>Datenbankname</b>	EMBASE	
<b>Suchoberfläche</b>	Ovid	
<b>Datum der Suche</b>	08.12.2010	
<b>Zeitsegment</b>	1980 to 2010 week 50	
<b>Suchfilter</b>	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle <sup>17</sup> ] – Strategy minimizing difference between sensitivity and specificity	
<b>#</b>	<b>Suchbegriffe</b>	<b>Ergebnis</b>
1	Meglitinide/	848
2	Nateglinide/	1686
3	Repaglinide/	2118
4	(glinid* or meglitinid* or nateglinid* or repaglinid*).ab,ti.	1069
5	(starlix or novonorm or novo norm or prandin).ab,ti.	32
6	(105816-04-4 or 135062-02-1).rn.	2854
7	or/1-6	3467
8	Diabetes mellitus/	224164
9	Non Insulin dependent Diabetes mellitus/	91081
10	(diabet* or niddm or t2dm).ab,ti.	379777
11	or/8-10	454517
12	(random* or double-blind*).tw.	650136
13	placebo*.mp.	243550
14	or/12-13	773621
15	and/7,11,14	719

<sup>17</sup> Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

**Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel**

Nicht zutreffend

**Anhang 4-A2: Suche nach RCT für indirekte Vergleiche**

Nicht zutreffend

**Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien**

Nicht zutreffend

**Anhang 4-A4: Suche nach weiteren Untersuchungen**

Nicht zutreffend

**Anhang 4-B: Suchstrategien – Suche in Studienregistern**

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jedes durchsuchte Studienregister ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters (z. B. [clinicaltrials.gov](http://www.clinicaltrials.gov)), die Internetadresse, unter der das Studienregister erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

<b>Studienregister</b>	clinicaltrials.gov
<b>Internetadresse</b>	<a href="http://www.clinicaltrials.gov">http://www.clinicaltrials.gov</a>
<b>Datum der Suche</b>	08.12.2010
<b>Suchstrategie</b>	(Starlix OR Novonorm OR Prandin OR Nateglinid OR Repaglinid) [ALL-FIELDS] AND ("Phase II" OR "Phase III" OR "Phase IV") [PHASE]
<b>Treffer</b>	23

**Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel**

Nicht zutreffend

**Anhang 4-B2: Suche nach RCT für indirekte Vergleiche**

Nicht zutreffend

**Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien**

Nicht zutreffend

**Anhang 4-B4: Suche nach weiteren Untersuchungen**

Nicht zutreffend

**Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)**

*Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der / den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.*

**Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel**

Nicht zutreffend

**Anhang 4-C2: Suche nach RCT für indirekte Vergleiche**

Nicht zutreffend

**Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien**

Nicht zutreffend

**Anhang 4-C4: Suche nach weiteren Untersuchungen**

Nicht zutreffend

**Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern)**

*Listen Sie nachfolgend die durch die Studienregistersuche(n) identifizierten, aber ausgeschlossenen Studien auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.*

**Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel**

Nicht zutreffend

**Anhang 4-D2: Suche nach RCT für indirekte Vergleiche**

Nicht zutreffend

**Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien**

Nicht zutreffend

**Anhang 4-D4: Suche nach weiteren Untersuchungen**

Nicht zutreffend

**Anhang 4-E: Methodik der eingeschlossenen Studien – RCT**

*Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-53 inklusive eines Flow-Charts für den Patientenfluss.*

*Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-53 inklusive eines Flow-Charts für den Patientenfluss.*



Tabelle 4-53 (Anhang): Studiendesign und -methodik für Studie PTC124-GD-007-DMD

Item <sup>a</sup>	Charakteristikum	Studieninformation
<b>Studienziel</b>		
<b>2 b</b>	Genaue Ziele, Fragestellung und Hypothesen	Primäres Ziel: Effekt von Ataluren auf die Gehfähigkeit bei gehfähigen männlichen Patienten mit genetisch gesicherter nmDBD im Alter von 5 Jahren und älter
<b>Methoden</b>		
<b>3</b>	Studiendesign	Internationale, Phase 2b, multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte, Dosisfindungsstudie, GCP-konform
<b>3a</b>	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Parallelgruppenstudie; Randomisierung im Verhältnis 1:1:2:2: Placebo (niedrige Dosis), Placebo (hohe Dosis), Ataluren 10, 10, 20 mg/kg KG oder 20, 20, 40 mg/kg KG und Tag
<b>3b</b>	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p><u>Relevante Änderungen:</u></p> <p><i>Amendment 1.1</i> (20. Dezember 2007; vor Einschluss des ersten Patienten im Februar 2008): -</p> <p><i>Amendment 2.0</i> (28. April 2008):</p> <ul style="list-style-type: none"> <li>- Spezifizierung des erforderlichen phänotypischen Nachweises einer DMD → Nur Einschluss von BMD-Patienten mit einem schwereren Erscheinungsbild, um einen möglichen Ceiling-Effekt zu verringern</li> <li>- Zuteilung von Zwillingspaaren zu derselben Behandlungsgruppe, um das Risiko einer Entblindung und von Dosierungsfehlern innerhalb einer Familie zu verringern; zusätzliche Sensitivitätsanalyse, um den Einfluss dieser Änderung auf die Primäranalyse zu erfassen</li> <li>- Definition eines Behandlungs-bedingten UEs</li> <li>- Korrektur der Zahl der Patienten in der Interim-Wirksamkeits- und Sicherheits-Analyse: n=100 auf n=90 entsprechend der O'Brien Fleming Prozedur</li> </ul> <p><i>Amendment 2.1</i> (27. Oktober 2008):</p> <ul style="list-style-type: none"> <li>- Aktualisierung der Screening-Laborwerte, um Einschluss von Patienten mit klinisch insignifikanten Laborabnormalitäten zu ermöglichen und unnötige Blutabnahmen zu vermeiden.</li> </ul> <p><i>Amendment 3</i> (4. März 2009):</p> <ul style="list-style-type: none"> <li>- Spezifizierung des Eignungskriteriums während des Screenings: Gehstrecke von ≥75 m im 6MWT; Gehstrecken von &lt;75 m zulässig während der nachfolgenden 6MWTs</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>- Änderung des Analyseplans der SAM-Daten</p> <p>- Änderung des Stratifizierungskriteriums 6MWD von &lt;270 m vs. ≥270 m auf &lt;350 m vs. ≥350 m. Eine erste Analyse der 6MWT-Daten vor Behandlungsbeginn zeigte eine mittlere Gehstrecke von etwa 350-360 m.</p> <p>- Revision des Analyseplans der 6MWD: MMRM (RANCOVA) anstelle der RANOVA, da vermutet wurde, dass das Ausmaß der Änderung mit der BL 6MWD in Zusammenhang steht und die MMRM wahrscheinlich eine höhere Power besitzt als die RANOVA</p> <p><u>Änderung des initialen Studienreports (26. Februar 2011) und der primären Analysepopulation (relevante Änderungen):</u></p> <p>- Bei zwei Patienten desselben Zentrums wurden die BL 6MWTs, die aufgrund von Verletzungen der unteren Extremitäten als nicht valide hätten eingestuft werden müssen, unkorrekter Weise als valide klassifiziert. Da diese Patienten trotz invalider 6MWT-Werte in die ITT-Population eingeschlossen worden waren, wurde die Definition der Analysepopulation klargestellt.</p> <p>- Ergänzung einer Methode zur Adjustierung für multiple Vergleiche in Permutationstests; post-hoc-Analyse der Gehstrecke im 6MWT in Prozent des erwarteten Wertes</p>
<b>4</b>	Probanden / Patienten	
<b>4a</b>	Ein-/Ausschlusskriterien der Probanden / Patienten	<p><u>Einschlusskriterien:</u></p> <ol style="list-style-type: none"> <li>1. Nachweis einer unterschriebenen und datierten informierten Einwilligungserklärung des Patienten (und/oder seiner Eltern oder des legalen Vormunds) nach Aufklärung über alle relevanten Aspekte der Studie</li> <li>2. Männliches Geschlecht</li> <li>3. Alter ≥5 Jahre</li> <li>4. Phänotypischer Nachweis einer DMD auf Basis des Auftretens charakteristischer klinischer Symptome oder Zeichen (d.h. proximale Muskelschwäche, schwankender Gang und Gowers Manöver; siehe 3b) im Alter von 9 Jahren, erhöhte Serum-Kreatinkinase und anhaltende Gehschwierigkeiten</li> <li>5. Dokumentation einer Nonsense-Punktmutation im Dystrophin-Gen durch Gensequenzierung durch ein Labor, welches durch das „College of American Pathologists“, „Clinical Laboratory Improvement Act/Amendment“ oder eine</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>äquivalente Organisation zertifiziert wurde</p> <p>6. Dokumentation, dass eine Blutprobe zur Bestätigung der Nonsense-Mutation im Dystrophin-Gen entnommen wurde</p> <p>7. Fähigkeit während des Screening-6MWTs <math>\geq 75</math> m ohne Unterstützung zu gehen</p> <p>8. Überprüfung, ob bestimmte Screening-Laborwerte (hepatisch, renal, Plasma-ACTH, und Serumelektrolyte) innerhalb der spezifizierten Bereiche des Zentrallabors lagen</p> <p>8. Bei sexuell aktiven Patienten Einwilligung während der Einnahme der Studienmedikation und eines Follow-ups über sechs Wochen auf Geschlechtsverkehr zu verzichten oder eine Barriere- oder medizinische Verhütungsmethode anzuwenden</p> <p>9. Bereitschaft und Fähigkeit, die vorgesehenen Visiten, den Einnahmeplan, die Studienverfahren, die Laboranalysen und Studienrestriktionen einzuhalten</p> <p><u>Ausschlusskriterien:</u></p> <ol style="list-style-type: none"> <li>1. Behandlung mit systemischen Aminoglykosid-Antibiotika innerhalb von drei Monaten vor Beginn der Behandlung</li> <li>2. Einleitung einer systemischen Kortikosteroid-Therapie innerhalb von sechs Monaten vor Behandlungsbeginn oder Änderung der systemischen Kortikosteroid-Therapie innerhalb von drei Monaten vor Behandlungsbeginn</li> <li>3. Jede Änderung der Prophylaxe/Behandlung der kongestiven Herzinsuffizienz innerhalb von drei Monaten vor Behandlungsbeginn</li> <li>4. Behandlung mit Warfarin innerhalb von einem Monat vor Behandlungsbeginn</li> <li>5. Vorherige Therapie mit Ataluren</li> <li>6. Bekannte Überempfindlichkeit gegenüber einem der Inhaltsstoffe oder Hilfsstoffe der Studienmedikation</li> <li>7. Exposition mit einem anderen investigativen Arzneimittel innerhalb von zwei Monaten vor Behandlungsbeginn</li> <li>8. Vorgeschichte eines großen chirurgischen Eingriffs innerhalb von 30 d vor Behandlungsbeginn</li> <li>9. Andauernde immunsuppressive Therapie (ausgenommen Glukokortikoide)</li> <li>10. Fortlaufende Teilnahme in einer anderen klinischen, therapeutischen Studie</li> <li>11. Erwartung eines großen chirurgischen Eingriffs (z.B. Skoliose-Operation) während</li> </ol>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>der 12-monatigen Behandlungsphase der Studie</p> <p>12. Notwendigkeit einer Beatmungshilfe während des Tages</p> <p>13. Klinische Symptome oder Zeichen einer kongestiven Herzinsuffizienz (ACC/AHA Stadium C oder D) oder Nachweis einer klinisch signifikanten Myopathie im EKG</p> <p>14. Vorherige oder fortbestehende Beschwerden (z.B. Begleiterkrankungen, psychiatrische Beschwerden, Verhaltensauffälligkeiten, Alkoholismus, Substanzabusus), medizinische Vorgeschichte, physikalische Befunde, EKG-Befunde oder auffällige Laborbefunde, die nach Auffassung des Prüfarztes die Sicherheit des Patienten negativ beeinflussen könnten, es unwahrscheinlich erscheinen lassen, dass die Behandlung oder das Follow-up abgeschlossen werden können oder die die Erfassung der Studienergebnisse beeinflussen könnten</p>
<b>4b</b>	Studienorganisation und Ort der Studiendurchführung	
<b>5</b>	<p>Interventionen</p> <p>Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.</p>	<p>Placebo (niedrige Dosis)</p> <p>Placebo (hohe Dosis)</p> <p>Ataluren 10, 10, 20 mg/kg KG (gesamt 40 mg/d)</p> <p>Ataluren 20, 10, 20 mg/kg KG (gesamt 80 mg/d)</p> <p>Die Studienpräparate sollten in den angegebenen Einzeldosen jeweils dreimal täglich morgens, mittags und abends oral eingenommen werden. Wenn möglich sollten Dosierungsintervalle von 6 h (<math>\pm</math> 1 h) zwischen der Einnahme am Morgen und Mittag bzw. Mittag und Abend und von 12 h (<math>\pm</math> 1 h) zwischen der Einnahme am Abend und Morgen eingehalten werden. Eine Einnahme innerhalb von 30 min nach einer Mahlzeit wurde empfohlen. Die geplante Behandlungsdauer betrug 48 Wochen. Die Prüfpräparate (Pulver) wurden in 30-60 ml stillem Wasser oder Milch aufgelöst, mindestens 15 sec gerührt, und eingenommen. Eventuelle Rückstände wurden durch erneute Zugabe von Flüssigkeit suspendiert und ebenfalls eingenommen.</p>
<b>6</b>	Zielkriterien	
<b>6a</b>	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden	<p><u>Primäres Zielkriterium:</u></p> <p>Änderung der 6MWD im 6MWT von BL bis Woche 48: Der Test wurde entsprechend</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
	(z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p>eines standardisierten Verfahrens, welches in der „University California Davis“ entwickelt wurde (22), ausgeführt (siehe unten Training der Prüferärzte).</p> <p><u>Sekundäre Zielkriterien:</u></p> <p><i>Körperliche Funktion</i></p> <ul style="list-style-type: none"> <li>- „Timed function tests“ (TFT; zu Screening, BL sowie alle 6 Wochen während der Behandlung): Zeit, die benötigt wird, um <ul style="list-style-type: none"> <li>• aus der Rückenlage aufzustehen</li> <li>• 10 m zu gehen oder zu rennen</li> <li>• 4 Stufen standardisierter Größe zu ersteigen</li> <li>• 4 Stufen standardisierter Größe herabzusteigen</li> </ul> </li> <li>- Myometrie <ul style="list-style-type: none"> <li>→ Änderung der Kraft während des Beugens und Streckens der Knie und Ellbogen und Schulter-Abduktion</li> </ul> </li> </ul> <p>Gemessen zum Screening, BL sowie alle 6 Wochen während der Behandlung mittels Myometer entsprechend standardisierter Verfahren (37,38)</p> <ul style="list-style-type: none"> <li>- Aktivität im Lebensbereich</li> </ul> <p>Die Zahl der pro Zeiteinheit bewältigten Treppenstufen wurden kontinuierlich mittels eines Pedometers („StepWatch<sup>®</sup> Activity Monitor“, SAM), welches am Fußgelenk getragen wurde, erfasst. Der SAM wurde während der Visiten durch das Klinikpersonal kalibriert und die Patienten/Betreuer wurden hinsichtlich der Handhabung des SAM instruiert. Die Patienten verließen die Klinik mit dem SAM und dieser sollte über <math>\geq 9</math> d ohne Unterbrechung während jeder der 6-wöchigen Behandlungsphasen getragen werden.</p> <p>Eine „Clinical Evaluator (CE) Training Group“ war verantwortlich, um standardisierte Verfahren für den 6MWT, die TFTs sowie die Myometrie zu entwickeln und den SAM zu kalibrieren. Ein CE Manual, in dem diese Verfahren beschrieben wurden, wurde allen Studienzentren zur Verfügung gestellt. Die CE Trainingsgruppe beaufsichtigte ein zentralisiertes Training der Studien-CEs, welches aus einem initialen Training vor Studienbeginn und einem Auffrischungstraining etwa ein Jahr nach Studienbeginn bestand.</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p><i>PROs („patient [parent/caregiver]-reported outcomes“)</i></p> <ul style="list-style-type: none"> <li>- „Pediatric Quality of Life Inventory“ (PedsQL): Änderung der gesundheitsbezogenen Lebensqualität im PedsQL (Domänen: körperlich, emotional, sozial und Funktion in der Schule sowie Fatigue-Gesamtscore)</li> <li>- „Treatment Satisfaction Questionnaire for Medication“ (TSQM): Änderung der Zufriedenheit mit der Behandlung</li> <li>- Änderung der Rate unbeabsichtigter Stürze auf Basis eines Patienten-/Betreuer-Tagebuchs</li> </ul> <p><i>Kognitive Funktion</i></p> <ul style="list-style-type: none"> <li>- Änderung des verbalen Gedächtnisses und der Aufmerksamkeit im „digit span test“; erhoben zum Screening, BL und alle 12 Wochen unter Behandlung</li> </ul> <p><i>Kardiale Funktion</i></p> <p>Änderung der Herzrate vor, während und nach dem 6MWT während des Screenings, zu BL und alle 6 Wochen unter Behandlung</p> <p>Wie oben beschrieben unterzogen sich sie CEs einem Training, um das Testprotokoll und die Logistik zu harmonisieren.</p> <p><i>Pharmakodynamik</i></p> <ul style="list-style-type: none"> <li>- Muskel-Dystrophin Expression (Bizeps) via Immunfluoreszenz</li> <li>- Serum Kreatinkinase</li> </ul> <p><i>Verträglichkeit und Sicherheit</i></p> <p><u>UEs:</u> Dokumentation aller UEs inklusive Schwere und Schweregrad, maximale Schwere des UEs, Zusammenhang mit der Prüfmedikation (wahrscheinlich, möglich, unwahrscheinlich oder kein Zusammenhang), Datum und Zeit bis zum Auftreten des UEs, Datum und Zeit bis zur Resolution des Ereignisses oder Tod, Maßnahmen im Zusammenhang mit dem Ereignis und Ausgang des Ereignisses</p> <p><u>Klinische Laborparameter:</u></p> <p>Blutproben zur Bestimmung hämatologischer, biochemischer und von Nebennieren-Parametern (Screening, vor der ersten Einnahme der Prüfpräparate, alle 6 Wochen während der Behandlung und nach der Behandlung)</p> <p>Urinanalyse: Screening, vor der ersten Einnahme der Prüfpräparate, alle 6 Wochen während der Behandlung und nach der Behandlung; zusätzliche Urinproben zur</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Bestimmung explorativer renaler Biomarker</p> <p>Abgesehen von dem Urin-Dipstick-Test wurden alle Analysen in einem zentralen Labor ausgeführt. Hepatische, renale und adrenale Effekte wurden speziell überwacht und es wurden Maßnahmen im Falle relevanter Änderungen dieser Laborparameter festgelegt; die Maßnahmen im Falle anderer Laborabweichungen lagen im Ermessen des Prüfarztes.</p> <p>Vitalparameter: Puls, Atemfrequenz, SBP und DBP und Temperatur (Screening, vor der ersten Dosis, alle 6 Wochen und Follow-up) EKGs (zum Screening, alle 12 Wochen und Follow-up): Den Prüfzentren wurden kalibrierte EKG-Geräte zur Verfügung gestellt und die EKG-Daten wurden direkt an ein zentrales EKG-Labor übertragen. Dort wertete ein Kardiologe die Daten aus und das Zentrum erhielt innerhalb von 24 h einen Bericht. EKG-Berichte, die den EKG-„Alert“-Kriterien entsprachen, waren unmittelbar weiterzuleiten.</p> <p>Ultraschall der Niere: Entsprechend der lokalen Standards (Screening und nach 48 Wochen Behandlung)</p> <p>Ärztliche Untersuchung, Größe und KG: Screening, BL, alle 12 Wochen und Follow-up</p>
<b>6b</b>	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine
<b>7</b>	Fallzahl	
<b>7a</b>	Wie wurden die Fallzahlen bestimmt?	<p>Die Fallzahlschätzung beruht auf dem primären Endpunkt 6MWD. Zum Zeitpunkt der Planung der Studie lagen keine Daten zum 6MWT bei Jungen mit DMD vor. Basierend auf Daten im 10MWT wurde eine Gehstrecke von ca. 270 m im 6MWT angenommen. Ausgehend von den Ergebnissen bei anderen chronischen Erkrankungen wurde in der Studie eine Verlängerung der Gehstrecke um 10% (entsprechend 30 m) oder mehr als Target definiert. Unter der Annahme einer Änderung von mindestens 30 m in einem der Ataluren-Arme, einer erwarteten SD von etwa 50 m in allen Armen und einer 1:1:1 Randomisierung waren 50 Patienten in jedem der drei Arme erforderlich (n=150), um einen Unterschied von 30 m mit einer &gt;85%igen Power unter Einsatz des Dunnetts T-Test bei einem Signifikanzniveau von 0,042 (unter Berücksichtigung des Alpha in der Interim-Analyse) zu detektieren. Unter der Annahme, dass etwa 10% die Studie vorzeitig</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		abbrechen, war eine Gesamtzahl von etwa 160 Patienten geplant. Die Schätzung beruhte auf 10000 Monte-Carlo-Simulationen unter Anwendung von Normalnull-Verteilungen des Mittelwerts $\pm$ SD ( $0\pm 50$ m in allen drei Armen, $0\pm 50$ m in der Placebo-Gruppe und $30\pm 50$ m in den Ataluren-Armen). Die SD („within-patient“) sollte kontinuierlich überwacht werden (Anpassung der Fallzahl wenn vorausgesagte SD $>50$ m).
<b>7b</b>	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Es fanden zwei geplante Interim-Analysen durch ein unabhängiges DMC statt: 1) Nach Einschluss von $\geq 45$ Patienten, die 12 Wochen oder mehr behandelt worden waren (Sicherheitsanalyse); 2) Nach Einschluss von $\geq 90$ Patienten, die 24 Wochen oder mehr behandelt worden waren (in Amendment 2.0 des Studienprotokolls von 100 auf 90 Patienten reduziert; Sicherheitsanalyse, auf die geplante Wirksamkeitsanalyse wurde verzichtet)
<b>8</b>	Randomisierung, Erzeugung der Behandlungsfolge	
<b>8a</b>	Methode zur Generierung der zufälligen Zuteilung	„Interactive Voice Response/Interactive Web Response“ (IVR/IWR)-System
<b>8b</b>	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Stratifizierung nach Alter ( $<9$ vs. $\geq 9$ Jahre), Einnahme von Glukokortikoiden (ja / nein) und BL 6MWD ( $\geq 350$ m vs. $<350$ m; initial $<270$ m vs. $\geq 270$ m, siehe 3b) Bei der Randomisierung wurde eine adaptive Methode unter Einsatz von Minimierungstechniken angewendet <sup>1</sup> .
<b>9</b>	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Die Zuteilung erfolgte doppelblind mittels eines IVR/IWR-Systems. Die Geheimhaltung war bis zur Zuteilung gewährleistet.
<b>10</b>	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Das IVR/IWR-System wurde durch eine unabhängige CRO erstellt. Ein Vertreter des Studienteams gab die notwendigen Kriterien aus der Screening-Untersuchung in das IVR/IWR-System ein und das IVR/IWR-System teilte den Patienten einem Behandlungsarm zu.
<b>11</b>	Verblindung	
<b>11a</b>	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Es handelte sich um eine Doppelblindstudie. Die Patienten, Eltern/Betreuer, das Personal am Studienzentrum, die Angestellten des Sponsors sowie andere Studienbeteiligte waren bis zum Schließen der Datenbank verblindet. Die Prüfpräparate waren



Item <sup>a</sup>	Charakteristikum	Studieninformation
		hinsichtlich ihres äußerlichen Erscheinungsbildes, Geschmacks, Geruchs, der Verpackung, Beschriftung und Anwendungsart nicht unterscheidbar. Eine Entblindung war nur im Falle einer Notfallsituation, auf Verlangen des DMCs zum Zeitpunkt der Interim-Analysen und nach Beendigung der Studie möglich. Während der Studie war die Gruppenzuteilung lediglich dem unabhängigen Statistiker und dem DMC zugänglich.
<b>11b</b>	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht zutreffend
<b>12</b>	Statistische Methoden	
<b>12a</b>	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p><u>Änderung des 6MWD</u>: MMRM-Analyse der Kovarianz (RANCOVA); Faktoren: BL 6MWD, Alter (&lt;9 oder ≥9 Jahre), Kortikosteroid-Anwendung (ja / nein), Visite und die Behandlung-Visite-Interaktion; ITT-Analyse (siehe unten); Analyse der log- und anschließend rank-transformierten Daten im Falle einer Nicht-Normalverteilung (analysiert mittels Shapiro-Wilks W-Test auf einem Signifikanzniveau von 0.05); siehe auch Kapitel 4.3.1.2.1.</p> <p>Die Analysen des 6MWT und der TFTs wurden post-hoc nach Ausschluss von zwei Patienten mit nicht validen BL 6MWT-Werten in der kITT-Population wiederholt.</p> <p><u>Sekundäre Endpunkte (wenn nicht anders vermerkt)</u>: MMRM-Analyse unter Berücksichtigung der entsprechenden BL-Variable als Kovariate; unabhängige Variablen: Behandlung, Visite, Behandlung-Visite und die Stratifikationsfaktoren Alter (&lt;9 vs. ≥9 Jahre), BL Kortikosteroidanwendung (ja vs. nein) und BL 6MWD (≥350 m vs. &lt;350 m); ITT- und PP-Analyse; siehe auch Kapitel 4.3.1.2.1.</p>
<b>12b</b>	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p><u>Supportive Analysen</u>:</p> <ul style="list-style-type: none"> <li>- Änderungen der 6MWD von BL bis zu jeder der geplanten Visiten</li> <li>- Einfluss unabhängiger Variablen als Kovariaten: Stratifizierungsfaktoren (Alter, Kortikosteroide), Altersverteilung (z.B. 5-6, 7-8, 9-11, ≥12 Jahre). Typ des Nonsense Stopcodons und Lokalisation des Nonsense-Exons. Auf die Analysen nach Größe und KG zu BL und Prophylaxe/Behandlung einer kongestiven Herzinsuffizienz wurde verzichtet.</li> <li>- MMRM-Modell mit BL 6MWD als Kovariate, um den Effekt von</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Interaktionen zwischen der Behandlung und der Stratifizierung (&lt;9 Jahre vs. ≥9 Jahre und ≥350 m vs. &lt;350 m) sowie einer Behandlung mit Kortikosteroiden (ja / nein) zu evaluieren</p> <ul style="list-style-type: none"> <li>- Anteil der Patienten mit einer 10%igen Verschlechterung oder Besserung der 6MWD in Woche 48 (ITT und PP; Cochran-Mantel-Haenszel Chi Quadrat)</li> <li>- Zeit bis zu einer andauernden 10%igen Verschlechterung oder Besserung der 6MWD in Woche 48 (Kaplan-Meier und stratifizierter log-rank-Test)</li> </ul> <p><u>Sensitivitätsanalysen:</u></p> <ul style="list-style-type: none"> <li>- Fehlende Compliance mit dem Protokoll,</li> <li>- Einschluss von Zwillingspaaren,</li> <li>- Fehlende Werte (Analyse nach multipler Imputation) und</li> <li>- Dynamische Randomisierung im MMRM (Permutationstest)</li> <li>- LOCF-Imputation für fehlende Werte unter Anwendung eines ANCOVA-Modells unter Berücksichtigung der im SAP definierten Kovariaten</li> </ul> <p><u>Post-hoc-Analysen:</u></p> <ul style="list-style-type: none"> <li>- Analyse des primären und der TFT Endpunkte in der kITT-Population, in der invalide BL 6MWD-Werte (aufgrund von Verletzungen der unteren Extremitäten) bei zwei Patienten durch die Screening-Werte ersetzt wurden (nur valide BL 6MWD-Werte durften in die ITT-Analyse einfließen)</li> <li>- MMRM-Analyse unter Einsatz eines zusätzlichen BL-Visite Interaktionsterms, da die Beziehung zwischen BL 6MWD und post-BL 6MWD über die Zeit variieren kann</li> <li>- Permutationstests zur Analyse des Effekts der dynamischen Randomisierung unter Anwendung des ANCOVA-LOCF basierend auf den nicht transformierten Daten</li> <li>- Analyse der 6MWD über 48 Wochen als Prozent des erwarteten Wertes, da im Verlauf der Studie normative Daten bei DMD-Patienten verfügbar wurden</li> <li>- Einschluss des BL-Visite Interaktionsterms in die geplante GEE-Analyse der unbeabsichtigten Stürze; zusätzlicher Permutationstest zur Validierung des Modells</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		- Analyse der Beziehung zwischen den Ataluren-Plasmakonzentration und den Wirksamkeitsparametern
<b>Resultate</b>		
<b>13</b>	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
<b>13a</b>	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	Placebo: a) 57, b) 57, c) 57 Ataluren 10/10/20: a) 57, b) 57, c) 57 Ataluren 20/20/40: a) 60, b) 60, c) 60
<b>13b</b>	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Placebo: Keine Ataluren 10/10/20: Keine Ataluren 20/20/40: n=1 vorzeitiger Abbruch aufgrund von Nicht-Compliance mit dem Protokoll
<b>14</b>	Aufnahme / Rekrutierung	
<b>14a</b>	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	28. Februar 2008 (erste Patientenvisite) bis 17. Dezember 2009 (letzte Patientenvisite)
<b>14b</b>	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde protokollgemäß beendet.
<b>a: nach CONSORT 2010.</b>		

<sup>1</sup> Folgende Charakteristika wurden der adaptiven Randomisierungsmethode zugrunde gelegt:

- Das angestrebte Verhältnis zwischen den Gruppen betrug 1.1:2:2 für die Placebo-Niedrig-Dosis bzw. - Hoch-Dosis-Gruppe und die Ataluren 10/10/20 mg/kg bzw. 20/20/40 mg/kg Gruppe.
- Alle Stratifizierungsfaktoren sollten gleichermaßen gewichtet werden.
- Imbalancen wurden mittels der Varianz-Imbalanzfunktion kalkuliert.
- Die Wahrscheinlichkeit der Zuordnung zu einer Gruppe ist in nachstehender Tabelle beschrieben.
- Bei Vorliegen von gebundenen Paaren ("ties") erfolgte die Zuteilung zufällig.
- Ausnahme: Hatte ein Patient einen Zwilling, der bereits randomisiert worden war, wurde der Patient derselben Behandlungsgruppe zugeteilt.

Gruppe	Wahrscheinlichkeit einer Akzeptanz	Wahrscheinlichkeit einer Ablehnung			
		Niedrig-Dosis-Placebo	Hoch-Dosis-Placebo	Ataluren 10/10/20	Ataluren 20/20/40
Niedrig-Dosis-Placebo	0.770	-	0.046	0.092	0.092
Hoch-Dosis-Placebo	0.770	0.046	-	0.092	0.092
Ataluren 10/10/20	0.816	0.046	0.046	-	0.092
Ataluren 10/20/40	0.816	0.046	0.046	0.092	-

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

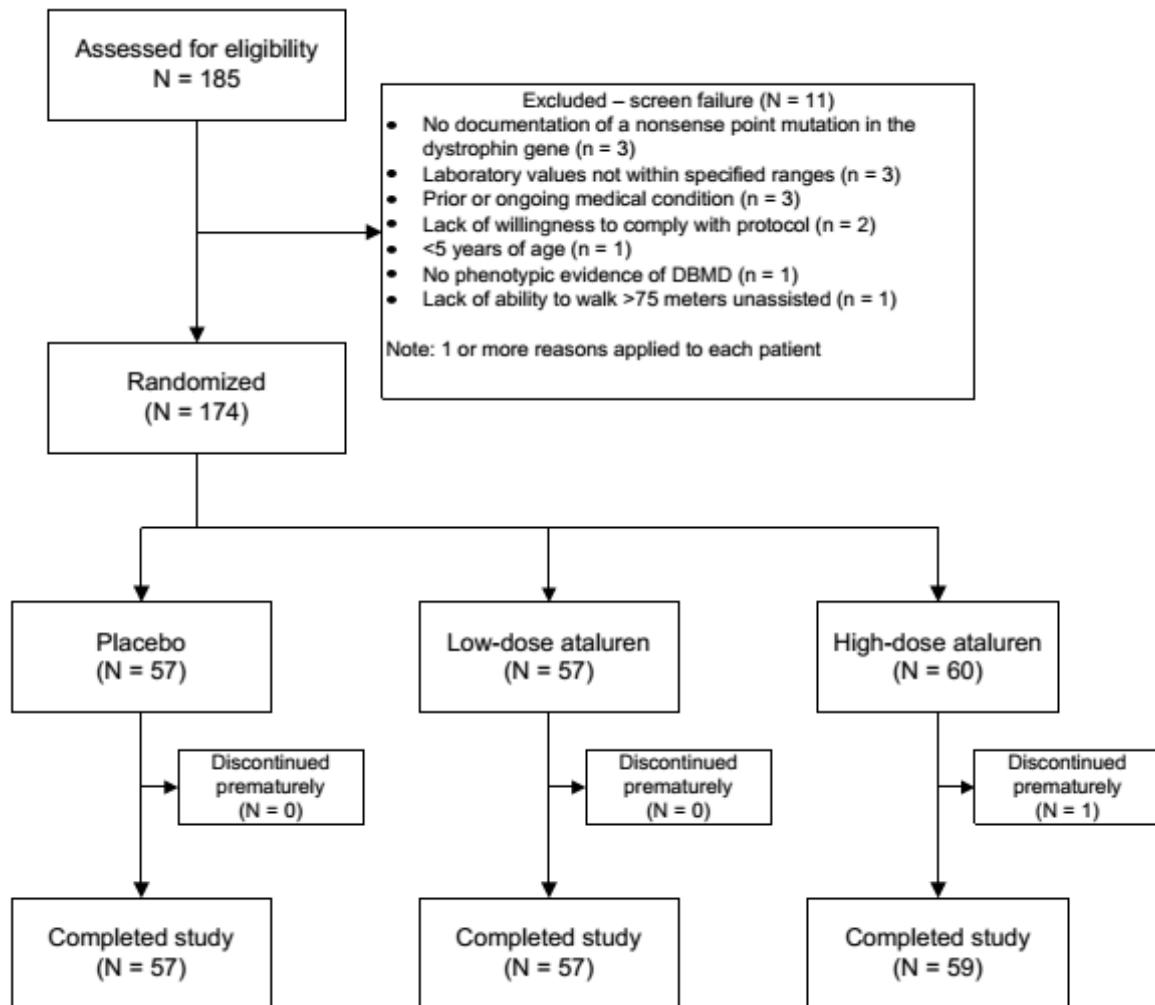


Abbildung 19: Patientenfluss in der Studie PTC124-GD-007-DMD

**Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten**

Der nachfolgend dargestellte Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.

- Verzerrungsaspekte auf Endpunktebene. In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich sollen die Bögen studienbezogen ausgefüllt werden. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

*Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.*

Tabelle 4-54 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PTC124-GD-007-DMD

**Studie: PTC124-GD-007-DMD**

**Tabelle: Liste der für die Bewertung herangezogenen Quellen**

Genauere Benennung der Quelle	Kürzel
PTC124-GD-007-DMD Klinischer Studienbericht (8)	A
PTC124-GD-007-DMD Studienprotokoll (6)	B
PTC124-GD-007-DMD Statistischer Analyseplan (7)	C
PTC124-GD-007-DMD CHMP Beurteilungsbericht Ataluren (9)	D
PTC124-GD-007-DMD Gründe für eine erneute Begutachtung von Ataluren durch das CHMP (31)	E

#### A Verzerrungsaspekte auf Studienebene:

##### Einstufung als randomisierte Studie

**ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

**nein** → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Es handelt sich um eine randomisierte Studie (A,B,C).

1.

##### für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

**ja**       **unklar**       **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gruppenzuteilung erfolgte randomisiert mittels IWR oder IVR unter Einsatz einer adaptiven Prozedur und die Randomisierung ist detailliert im Protokoll und SAP beschrieben (A,B).

##### für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

**ja**       **unklar**       **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

##### für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

**ja**       **unklar**       **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Die Gruppenzuteilung erfolgte zentral mittels IVR/TWR und die Prüfpräparate waren äußerlich nicht unterscheidbar (A,B).

---

**für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren**

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

**3. Verblindung von Patienten und behandelnden Personen**

**Patient:**

ja       unklar       nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

Es handelte sich um eine doppelblinde Studie (A,B,C).

---

**behandelnde bzw. weiterbehandelnde Personen:**

ja       unklar       nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

Die Behandler / Weiterbehandler waren ebenfalls verblindet, da die Behandlungsarme durch das doppelblinde Design nicht zu identifizieren waren (A,B,C).

---

**4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte**

ja       unklar       nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es ergeben sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung der relevanten Endpunkte. Die kITT-Analyse (post-hoc; A), die in diesem Nutzendossier als primär angesehen wird, wird als klinisch sowie methodisch begründet eingestuft. Auch das CHMP (2014) bestätigt in dem Beurteilungsbericht zu Ataluren, dass die kITT-Analyse von einem methodischen Standpunkt trotz ihres post-hoc Charakters als akzeptabel angesehen werden könne (D). Es wurden diverse post-hoc-Analysen (E) durchgeführt, primär mit dem Ziel, die Robustheit der primären Analyse zu bestätigen oder Patientenkollektive zu identifizieren, die mit hoher Wahrscheinlichkeit innerhalb eines Jahres auf die Therapie mit Ataluren gut ansprechen (im Hinblick auf die noch laufende confirmatorische Phase 3-Studie). Diese zusätzlichen Sensitivitätsanalysen tragen dazu bei, Unsicherheiten hinsichtlich der Validität der Ergebnisse zu vermindern und die Ergebnissicherheit zu erhöhen; eine ergebnisgesteuerte Berichterstattung ist aus diesen Analysen nicht ableitbar.

---

**5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es existieren keine Hinweise auf sonstige (endpunktübergreifende) Aspekte, die das Verzerrungspotenzial beeinflussen könnten (A,B,C). Siehe Punkt 4 zur Analyse der kITT-Population.

---

**Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Es liegen keine Hinweise vor, dass systematische Verzerrungen zu den Ergebnissen der Studie beigetragen haben.

---

**B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:****Endpunkt: 6MWD inklusive Progressor-Analysen****1. Verblindung der Endpunkterheber** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

Die Erhebung und Auswertung des Endpunktes erfolgte verblindet.

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Das ITT-Prinzip wurde für diesen Endpunkt adäquat umgesetzt (Nichtberücksichtigungsanteil: 0%). Die kITT-Analyse (post-hoc; A), die in diesem Nutzendossier als primär angesehen wird, wird als klinisch und medizinisch begründet eingestuft. Per SAP (D) durften in der ITT-Population nur Patienten mit validen BL 6MWD-Werten eingeschlossen werden. Der Ersatz von zwei verletzungsbedingt invaliden BL-Werten, die inkorrekt Weise der ITT-Population zugeordnet worden waren, durch die Screening-Werte der 6MWD erscheint methodisch gerechtfertigt. Auch das CHMP (2014) bestätigt in dem Beurteilungsbericht zu Ataluren, dass die kITT-Analyse von einem methodischen Standpunkt trotz ihres post-hoc-Charakters als akzeptabel angesehen werden könne (D). Das Ergebnis der protokollgemäßen ITT-Analyse ist in seiner Ausrichtung und hinsichtlich der Effektstärke der kITT-Analyse vergleichbar (A). Verschiedene Sensitivitätsanalysen bestätigen die Robustheit des Ergebnisses und zeigen, dass die Effekte nicht zugunsten der zu evaluierenden Behandlung verzerrt sind (A,E)

---



**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es existieren keine Hinweise auf eine ergebnisgesteuerte Berichterstattung dieses Endpunkts.

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja       neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es existieren keine Hinweise auf sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotential beeinflussen.

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig       hoch

Begründung für die Einstufung:

---

Es liegen keine Hinweise auf eine systematische Verzerrung der Ergebnisse des Endpunktes vor.

---

**Endpunkt: TFTs****1. Verblindung der Endpunkterheber** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

Die Erhebung und Auswertung des Endpunktes erfolgte verblindet.

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja       unklar       neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Das ITT-Prinzip wurde für diesen Endpunkt adäquat umgesetzt (Nichtberücksichtigungsanteil: 0%). Die kITT-Analyse (post-hoc; A), die in diesem Nutzendossier als primär angesehen wird, wird als klinisch und medizinisch begründet eingestuft. Auch das CHMP (2014) bestätigt in dem Beurteilungsbericht zu Ataluren, dass die kITT-Analyse von einem methodischen Standpunkt trotz ihres post-hoc-Charakters als akzeptabel angesehen werden könne (D).

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja     unklar     neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es existieren keine Hinweise auf eine ergebnisgesteuerte Berichterstattung dieses Endpunkts.

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können** ja     neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es liegen keine Hinweise auf eine systematische Verzerrung der Ergebnisse des Endpunktes vor.

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):** niedrig     hoch

Begründung für die Einstufung:

---

Es existieren keine Hinweise auf systematische, die das Verzerrungspotential beeinflussen.

---

**Endpunkt: Unerwünschte Ereignisse****1. Verblindung der Endpunkterheber** ja     unklar     neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

---

Die Erhebung und Auswertung des Endpunktes erfolgte verblindet.

---

**2. Adäquate Umsetzung des ITT-Prinzips** ja     unklar     neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Das ITT-Prinzip wurde für diesen Endpunkt adäquat umgesetzt und alle randomisierten Patienten wurden in der Sicherheitsanalyse berücksichtigt.

---

**3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine** ja     unklar     nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

Es existieren keine Hinweise auf eine ergebnisgesteuerte Berichterstattung dieses Endpunkts.

---

**4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können**

ja       nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

Es existieren keine Hinweise auf sonstige (endpunktspezifische) Aspekte, die das Verzerrungspotential beeinflussen.

---

**Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):**

niedrig       hoch

Begründung für die Einstufung:

---

Es liegen keine Hinweise auf eine systematische Verzerrung der Ergebnisse des Endpunktes vor.

---

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

**Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)**

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

**A Verzerrungsaspekte auf Studienebene:****Einstufung als randomisierte Studie**

**ja** → Bewertung der Punkte 1 und 2 für randomisierte Studien

**nein:** Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

---

---

1.

**für randomisierte Studien:****Adäquate Erzeugung der Randomisierungssequenz**

**ja:** Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

**unklar:** Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

**nein:** Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**für nicht randomisierte Studien:****Zeitliche Parallelität der Gruppen**

**ja:** Die Gruppen wurden zeitlich parallel verfolgt.

**unklar:** Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

**nein:** Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

**für randomisierte Studien:****Verdeckung der Gruppenzuteilung („allocation concealment“)** **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

 **unklar:** Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau. **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**für nicht randomisierte Studien:****Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren** **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

 **unklar:** Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau. **nein:** Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

**3. Verblindung von Patienten und behandelnden Personen****Patient:** **ja:** Die Patienten waren verblindet. **unklar:** Es finden sich keine diesbezüglichen Angaben. **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---



---

**behandelnde bzw. weiterbehandelnde Personen:**

**ja:** Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

**unklar:** Es finden sich keine diesbezüglichen Angaben.

**nein:** Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---



---

**4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte**

*Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).*

*Beispiele zu a und b:*

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

*Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter*

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

*Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:*

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

*Zulässige Gründe sind:*

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*

- Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---



---

##### 5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
  - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
  - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
  - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
  - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

**ja**

**nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---



---

##### Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

**niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

**hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

---



---

## B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: \_\_\_\_\_

### 1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

**ja:** Der Endpunkt wurde verblindet erhoben.

**unklar:** Es finden sich keine diesbezüglichen Angaben.

**nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---



---

### 2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

**ja:** Eines der folgenden Merkmale trifft zu:

- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und



es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

**unklar:** Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

**nein:** Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

*Beachte die Hinweise zu Punkt A.4!*

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

---

---

### 4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

*z. B.*

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

**ja**

**nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

---

---

### Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

*Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung*

*bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.*

**niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

**hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

---

---