

# Dossier zur Nutzenbewertung gemäß § 35a SGB V

*Ataluren (Translarna™)*

PTC Therapeutics, 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>9</b>
<b>Abkürzungsverzeichnis</b> .....	<b>13</b>
<b>4 Modul 4 – allgemeine Informationen</b> .....	<b>17</b>
4.1 Zusammenfassung der Inhalte von Modul 4.....	18
4.2 Methodik .....	30
4.2.1 Fragestellung .....	30
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung .....	33
4.2.3 Informationsbeschaffung .....	34
4.2.3.1 Studien des pharmazeutischen Unternehmers .....	34
4.2.3.2 Bibliografische Literaturrecherche .....	34
4.2.3.3 Suche in Studienregistern .....	35
4.2.3.4 Selektion relevanter Studien .....	36
4.2.4 Bewertung der Aussagekraft der Nachweise.....	36
4.2.5 Informationssynthese und -analyse .....	38
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien .....	38
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	39
4.2.5.3 Meta-Analysen.....	39
4.2.5.4 Sensitivitätsanalysen .....	41
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren .....	42
4.2.5.6 Indirekte Vergleiche .....	46
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen .....	48
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	48
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel .....	48
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	48
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	49
4.3.1.1.3 Studien aus der Suche in Studienregistern .....	51
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	52
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	54
4.3.1.2.1 Studiendesign und Studienpopulationen .....	54
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	68
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	69
4.3.1.3.1 Endpunkte – RCT.....	69
4.3.1.3.1.1 Gehstrecke im 6MWT (primärer Endpunkt) und 10%ige Verschlechterung im 6MWD – RCT.....	69
4.3.1.3.1.2 Zeit-Funktionstests – RCT.....	82
4.3.1.3.1.3 North Star Ambulatory Assessment (NSAA) – RCT.....	91
4.3.1.3.1.4 Gesundheitsbezogene Lebensqualität (PRO) – RCT.....	95

4.3.1.3.1.5	Aktivitäten des täglichen Lebens und Krankheitssymptome (PRO) – RCT .....	99
4.3.1.3.1.6	Unerwünschte Ereignisse – RCT .....	103
4.3.1.3.2	Subgruppenanalysen – RCT .....	124
4.3.1.3.2.1	Gehstrecke im 6MWT - RCT .....	125
4.3.1.3.2.2	Zeit-Funktionstests - RCT .....	132
4.3.1.3.2.3	North Star Ambulatory Assessment (NSAA) – RCT .....	145
4.3.1.3.2.4	Gesundheitsbezogene Lebensqualität (PRO) – RCT .....	148
4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien .....	150
4.3.2	Weitere Unterlagen .....	150
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien .....	150
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche .....	150
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche .....	151
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen .....	151
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT .....	151
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT .....	154
4.3.2.2	Nicht randomisierte vergleichende Studien .....	154
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien .....	154
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien .....	154
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien .....	155
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien .....	155
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien .....	156
4.3.2.3	Weitere Untersuchungen .....	156
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen .....	157
4.3.2.3.2	Charakteristika der weiteren Untersuchungen .....	157
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen .....	157
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen .....	157
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen .....	158
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen .....	158
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens .....	158
4.4.1	Beurteilung der Aussagekraft der Nachweise .....	158
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß .....	159
4.4.2.1	Hintergrund und Fragestellung .....	159
4.4.2.2	Aussagekraft der Nachweise .....	162
4.4.2.3	Validität und Patientenrelevanz der Endpunkte .....	163
4.4.2.4	Rationale der Analysen der Subpopulationen nach BL-6MWD (<300, ≥300 bis <400, ≥400 m) .....	167
4.4.2.5	Meta-Analysen der RCTs PTC124-GD-007-DMD und PTC124-GD-020-DMD – RCT .....	175
4.4.2.5.1	Vergleichbarkeit der Studien 070 und 020 .....	175
4.4.2.5.2	Vergleichbarkeit der Analyse-Populationen der Meta-Analysen .....	192
4.4.2.5.3	Ergebnisse der Meta-Analysen .....	200
4.4.2.5.4	Zusammenfassung der Ergebnisse der Meta-Analysen .....	223

4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht .....	229
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte .....	230
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	230
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	230
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen .....	230
4.5.4	Verwendung von Surrogatendpunkten .....	230
4.6	Liste der eingeschlossenen Studien.....	231
4.7	Referenzliste.....	233
	<b>Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche .....</b>	<b>241</b>
	<b>Anhang 4-B : Suchstrategien – Suche in Studienregistern.....</b>	<b>243</b>
	<b>Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....</b>	<b>244</b>
	<b>Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....</b>	<b>245</b>
	<b>Anhang 4-E : Methodik der eingeschlossenen Studien – RCT .....</b>	<b>246</b>
	<b>Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten .....</b>	<b>256</b>

**Tabellenverzeichnis**

	<b>Seite</b>
Tabelle 4-1: Darstellung der patientenrelevanten Endpunkte .....	22
Tabelle 4-2: Übersicht der Ergebnisse für patientenrelevante Endpunkte aus PTC124-GD-007-DMD und PTC124-GD-020-DMD (gesamte Studienpopulationen, Meta-Analysen) .....	25
Tabelle 4-3: Übersicht der Ergebnisse für patientenrelevante Endpunkte aus PTC124-GD-007-DMD und PTC124-GD-020-DMD (Subgruppen, Meta-Analysen) .....	27
Tabelle 4-4: Kriterien zur Prüfung der Aussagesicherheit zu Subgruppeneffekten (39) .....	44
Tabelle 4-5: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel .....	49
Tabelle 4-6: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	49
Tabelle 4-7: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel .....	52
Tabelle 4-8: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	53
Tabelle 4-9: Analysepopulationen und Analysemethoden der Phase 2b Studie PTC124-GD-007-DMD in Meta-Analysen .....	54
Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel .....	56
Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel (44).....	58
Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel („as-treated“ Population).....	59
Tabelle 4-13: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	68
Tabelle 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel .....	69
Tabelle 4-15: Operationalisierung der Endpunkte 6MWD und 10%ige Verschlechterung im 6MWD .....	71
Tabelle 4-16: Bewertung des Verzerrungspotenzials für Endpunkt 6MWD in RCT mit dem zu bewertenden Arzneimittel.....	73
Tabelle 4-17: Änderung der 6MWD von BL bis Woche 48 für PTC124-GD-020-DMD (ITT, n=114 Placebo, n=114 Ataluren, Gesamtpopulation).....	74
Tabelle 4-18: Änderung der mittleren 6MWD über den Studienverlauf für PTC124-GD-020-DMD (ITT) .....	75
Tabelle 4-19: Änderung der mittleren 6MWD über 48 Wochen für PTC124-GD-020-DMD in prä-spezifizierten Sensitivitätsanalysen (ITT) .....	76
Tabelle 4-20: Operationalisierung von Endpunkt TFT („timed function tests“) .....	83

Tabelle 4-21: Bewertung des Verzerrungspotenzials für den Endpunkt TFTs in RCT mit dem zu bewertenden Arzneimittel.....	84
Tabelle 4-22: Änderung der Zeit für TFTs von BL bis Woche 48, PTC124-GD-020-DMD, (ITT; n=114 Placebo, n=114 Ataluren Gesamtpopulation) (8) .....	85
Tabelle 4-23: Sensitivitätsanalysen der Änderung der TFTs bis Woche 48 für PTC124-GD-020-DMD; Ataluren versus Placebo (8) .....	86
Tabelle 4-24: Operationalisierung von Endpunkt NSAA („North Star Ambulatory Assessment“).....	92
Tabelle 4-25: Bewertung des Verzerrungspotenzials für Endpunkt NSAA in RCT mit dem zu bewertenden Arzneimittel .....	94
Tabelle 4-26: Änderung des NSAA von BL bis Woche 48 (ITT; n=114 Placebo, n=114) (8) .....	94
Tabelle 4-27: Operationalisierung von Endpunkt PODCI.....	97
Tabelle 4-28: Bewertung des Verzerrungspotenzials für Endpunkt PODCI in RCT mit dem zu bewertenden Arzneimittel.....	97
Tabelle 4-29: Änderung des PODCI von BL bis Woche 48 (ITT; n=114 Placebo, n=114) (8) .....	98
Tabelle 4-30: Operationalisierung von Endpunkt Aktivitäten des täglichen Lebens und Krankheitssymptome (PRO) .....	100
Tabelle 4-31: Bewertung des Verzerrungspotenzials für Endpunkt Aktivitäten des täglichen Lebens und Krankheitssymptome (PRO) in RCT mit dem zu bewertenden Arzneimittel.....	100
Tabelle 4-32: Änderung der Alltagsaktivitäten und Krankheitssymptome in Woche 48 aus Sicht der Patienten/Betreuer (ITT) (8) .....	101
Tabelle 4-33: Operationalisierung des Endpunkts „Unerwünschte Ereignisse“ (UEs) .....	105
Tabelle 4-34: Bewertung des Verzerrungspotenzials für Endpunkt UEs in RCT mit dem zu bewertenden Arzneimittel .....	106
Tabelle 4-35: Ergebnisse für TEAEs aus PTC124-GD-020-DMD („as-treated“); Ataluren versus Placebo .....	107
Tabelle 4-36: Ergebnisse für TEAEs stratifiziert nach Organsystem, die bei >5% der Patienten der Gesamtpopulation über 48 Wochen auftraten aus PTC124-GD-020-DMD („as-treated“) .....	108
Tabelle 4-37: Ergebnisse für renale und hepatische TEAEs aus PTC124-GD-020-DMD („as-treated“); Ataluren versus Placebo .....	110
Tabelle 4-38: Ergebnisse für die mit der Behandlung in Zusammenhang stehende TEAEs, die bei mehr als einem Patienten auftraten, aus PTC124-GD-020-DMD („as-treated“), Ataluren versus Placebo .....	111
Tabelle 4-39: Inzidenz schwerer (Grad 3) TEAEs aus PTC124-GD-020-DMD („as-treated“), Ataluren versus Placebo .....	112
Tabelle 4-40: Inzidenz schwerwiegender TEAEs aus PTC124-GD-020-DMD („as-treated“).....	114

Tabelle 4-41: Änderung der 6MWD über 48 Wochen in prä-spezifizierten Subgruppen für PTC124-GD-020-DMD (ITT) - Interaktionstests .....	126
Tabelle 4-42: Zeit für TFTs von BL bis Woche 48 in Subpopulationen nach BL-6MWD (ITT) für PTC124-GD-020-DMD; Ataluren versus Placebo (8) .....	134
Tabelle 4-43: Änderung NSAA von BL bis Woche 48 (ITT; BL-6MWD-Subpopulationen) für PTC124-GD-020-DMD; Ataluren versus Placebo (8).....	146
Tabelle 4-44: Änderung des PODCI von BL bis Woche 48 nach BL-6MWD-Gruppe (ITT) für PTC124-GD-020-DMD; Ataluren versus Placebo (8) .....	149
Tabelle 4-45: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche .....	151
Tabelle 4-46: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	152
Tabelle 4-47: Operationalisierung von <Endpunkt xxx>.....	152
Tabelle 4-48: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche .....	153
Tabelle 4-49: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	153
Tabelle 4-50: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien .....	155
Tabelle 4-51: Operationalisierung von <Endpunkt xxx>.....	155
Tabelle 4-52: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien .....	156
Tabelle 4-53: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	157
Tabelle 4-54: Darstellung der patientenrelevanten Endpunkte .....	161
Tabelle 4-55: Entscheidungshilfen zum Einschluss klinischer Endpunkte in Phase 2/3-Studien bei DMD (24).....	163
Tabelle 4-56: Übersicht der Einschlusskriterien von Phase 2/3-Studien bei DMD-Patienten .....	173
Tabelle 4-57: Vergleich der in Studie 007 und 020 erhobenen patientenrelevanten Endpunkte.....	175
Tabelle 4-58: Demographische und andere Charakteristika der Patienten – Vergleich der RCTs 007 und 020.....	179
Tabelle 4-59: Demographische und andere Charakteristika der Patienten (Subgruppe BL-6MWD $\geq$ 300 bis <400 m) – Vergleich der RCTs 007 und 020.....	181
Tabelle 4-60: Ergebnisse des primären Endpunkts in Studie 007 und 020 (8,9).....	183
Tabelle 4-61: Ergebnisse des primären und relevanter sekundärer Endpunkte in der Gesamtpopulation aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo (8,9) .....	185
Tabelle 4-62: Ergebnisse des primären und relevanter sekundärer Endpunkte in den 6MWD-Subpopulationen (8,9).....	187

Tabelle 4-63: Ergebnisse des primären und relevanter sekundärer Endpunkte in den 6MWD-Subpopulationen (8,9).....	190
Tabelle 4-64: Vergleich der TEAEs in Studie 070 und 020 (8,9).....	192
Tabelle 4-65: Disposition der Patienten in den Analyse-Populationen der Meta-Analysen..	193
Tabelle 4-66: Begleitbehandlung in den Studien 070 und 020 – Gepoolte Analyse („as-treated“) (16) .....	194
Tabelle 4-67: Demographische und andere Charakteristika der Patienten – Meta-Analysen	196
Tabelle 4-68: Demographische und andere Charakteristika der Patienten – Meta-Analysen der 6MWD-Subpopulationen .....	198
Tabelle 4-69: Ergebnisse der Meta-Analysen für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	201
Tabelle 4-70: Ergebnisse der Meta-Analysen der Subgruppen (<300 m, ≥300 m bis <400 m, ≥400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	203
Tabelle 4-71: Ergebnisse der gepoolten Subgruppen für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	204
Tabelle 4-72: Ergebnisse der Meta-Analysen für Zeit bis zur 10%igen Verschlechterung im 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	204
Tabelle 4-73: Ergebnisse der Meta-Analysen für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	205
Tabelle 4-74: Ergebnisse der Meta-Analysen der Subgruppen (<300 m, ≥300 m bis <400 m, ≥400 m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	207
Tabelle 4-75: Ergebnisse der gepoolten Subgruppen für Zeit um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	208
Tabelle 4-76: Ergebnisse der Meta-Analysen für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	209
Tabelle 4-77: Ergebnisse der Meta-Analysen der Subgruppen (<300 m, ≥300 m bis <400 m, ≥400 m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	210
Tabelle 4-78: Ergebnisse der gepoolten Subgruppen für Zeit um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	211
Tabelle 4-79: Ergebnisse der Meta-Analysen für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	212
Tabelle 4-80: Ergebnisse der Meta-Analysen der Subgruppen (<300 m, ≥300 m bis <400 m, ≥400 m) für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	213



Tabelle 4-81: Patientenrelevanz der Therapieeffekte für Morbiditätsendpunkte (Hedges' g).....	215
Tabelle 4-82: Verträglichkeit – Ergebnisse der Meta-Analysen für PTC124-GD-007-DMD und PTC124-GD-020-DMD .....	217
Tabelle 4-83: Ausmaß des Zusatznutzens auf Endpunktebene aus PTC124-GD-007-DMD und PTC124-GD-020-DMD (Meta-Analysen); Ataluren versus Placebo .....	228
Tabelle 4-84: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	229
Tabelle 4-85 (Anhang): Studiendesign und -methodik für Studie PTC124-GD-020-DMD..	247
Tabelle 4-86 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PTC124-GD-020-DMD .....	257

**Abbildungsverzeichnis**

	<b>Seite</b>
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel .....	51
Abbildung 2: Kaplan-Meier-Plot der Zeit bis zu einer dauerhaften 10%igen Verschlechterung der 6MWD (ITT) für PTC124-GD-020-DMD; Ataluren versus Placebo ..	77
Abbildung 3: Meta-Analyse für 6MWT aus PTC124-GD-007-DMD (kITT, MMRM) und PTC124-GD-020-DMD (ITT, ANCOVA); Ataluren versus Placebo .....	78
Abbildung 4: Meta-Analyse für 6MWT aus PTC124-GD-007-DMD (ITT, MMRM) und PTC124-GD-020-DMD (ITT, ANCOVA); Ataluren versus Placebo .....	79
Abbildung 5: Meta-Analyse für 6MWT aus PTC124-GD-007-DMD (kITT, mMMRM) und PTC124-GD-020-DMD (ITT, ANCOVA); Ataluren versus Placebo .....	80
Abbildung 6: Meta-Analyse für 6MWT aus PTC124-GD-007-DMD (ITT, ANCOVA) und PTC124-GD-020-DMD (ITT, ANCOVA); Ataluren versus Placebo .....	80
Abbildung 7: Meta-Analyse für Zeit bis zur 10%igen Verschlechterung aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo .....	81
Abbildung 8: Meta-Analyse für Verlust der Gehfähigkeit aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo.....	81
Abbildung 9: Meta-Analyse für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo.....	87
Abbildung 10: Meta-Analyse für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo.....	87
Abbildung 11: Meta-Analyse für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo.....	88
Abbildung 12: Meta-Analyse für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo.....	88
Abbildung 13: Meta-Analyse für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo.....	89
Abbildung 14: Meta-Analyse für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo.....	89
Abbildung 15: Meta-Analyse für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo .....	90
Abbildung 16: Meta-Analyse für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo .....	90
Abbildung 17: Meta-Analyse für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo .....	91
Abbildung 18: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	95
Abbildung 19: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	98

Abbildung 20: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	103
Abbildung 21: Meta-Analyse für schwerwiegende UE aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	117
Abbildung 22: Meta-Analyse für schwere UE (Grad $\geq 3$ ) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	118
Abbildung 23: Meta-Analyse für UE, die zum Studienabbruch führten aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	118
Abbildung 24: Meta-Analyse für UE der SOC „Erkrankungen des Gastrointestinaltrakts“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	118
Abbildung 25: Meta-Analyse für UE „Erbrechen“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	118
Abbildung 26: Meta-Analyse für UE „Diarrhöe“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	119
Abbildung 27: Meta-Analyse für UE „Abdominalschmerz“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	119
Abbildung 28: Meta-Analyse für UE „Flatulenz“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	119
Abbildung 29: Meta-Analyse für UE „Abdominalbeschwerden“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	119
Abbildung 30: Meta-Analyse für UE „Nausea“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	120
Abbildung 31: Meta-Analyse für UE der SOC „Erkrankungen des Nervensystems“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	120
Abbildung 32: Meta-Analyse für UE „Kopfschmerzen“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	120
Abbildung 33: Meta-Analyse für UE der SOC „Erkrankungen der Niere und der Harnwege“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	120
Abbildung 34: Meta-Analyse für UE „Enuresis“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	121
Abbildung 35: Meta-Analyse für UE der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	121
Abbildung 36: Meta-Analyse für UE „Fieber“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	121
Abbildung 37: Meta-Analyse für UE „Progression einer Erkrankung“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	121
Abbildung 38: Meta-Analyse für UE der SOC „Infektionen und parasitaere Erkrankungen“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	122

Abbildung 39: Meta-Analyse für UE „Gastroenteritis“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	122
Abbildung 40: Meta-Analyse für UE „Rhinitis“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	122
Abbildung 41: Meta-Analyse für UE der SOC „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	122
Abbildung 42: Meta-Analyse für UE „Sturz“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	123
Abbildung 43: Meta-Analyse für UE „Kontusion“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	123
Abbildung 44: Meta-Analyse für UE der SOC „Stoffwechsel- und Ernährungsstörungen“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	123
Abbildung 45: Meta-Analyse für UE „Appetit vermindert“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	123
Abbildung 46: Meta-Analyse für UE der SOC „Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	124
Abbildung 47: Meta-Analyse für UE „Rueckenschmerzen“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	124
Abbildung 48: Meta-Analyse der Subgruppen (<300, ≥300 bis <400, ≥400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	129
Abbildung 49: Paarweiser Interaktionstest der Subgruppen (<300 m versus ≥300 bis <400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	130
Abbildung 50: Paarweiser Interaktionstest der Subgruppen (≥400 m versus ≥300 bis <400m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	130
Abbildung 51: Paarweiser Interaktionstest der Subgruppen (<300 m versus ≥400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	131
Abbildung 52: Gesamteffektschätzer der gepoolten Subgruppen (≥300 bis <400 m; <300 m plus ≥400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	131
Abbildung 53: Meta-Analyse der Subgruppen (<300, ≥300 bis <400, ≥400 m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	138
Abbildung 54: Paarweiser Interaktionstest der Subgruppen (<300 m versus ≥300 bis <400 m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	139

Abbildung 55: Paarweiser Interaktionstest der Subgruppen ( $\geq 400$ m versus $\geq 300$ bis $< 400$ m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	139
Abbildung 56: Paarweiser Interaktionstest der Subgruppen ( $< 300$ m versus $\geq 400$ m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	140
Abbildung 57: Gesamteffektschätzer der gepoolten Subgruppen ( $\geq 300$ bis $< 400$ m; $< 300$ m plus $\geq 400$ m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	140
Abbildung 58: Meta-Analyse der Subgruppen ( $< 300$ , $\geq 300$ bis $< 400$ , $\geq 400$ m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo.....	141
Abbildung 59: Paarweiser Interaktionstest der Subgruppen ( $< 300$ m versus $\geq 300$ bis $< 400$ m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	142
Abbildung 60: Paarweiser Interaktionstest der Subgruppen ( $\geq 400$ m versus $\geq 300$ bis $< 400$ m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	143
Abbildung 61: Paarweiser Interaktionstest der Subgruppen ( $< 300$ m versus $\geq 400$ m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	143
Abbildung 62: Gesamteffektschätzer der gepoolte Subgruppen ( $\geq 300$ bis $< 400$ m; $< 300$ m plus $\geq 400$ m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	144
Abbildung 63: Meta-Analyse der Subgruppen ( $< 300$ , $\geq 300$ bis $< 400$ , $\geq 400$ m) für Zeit, um 10 m zu gehen/laufen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo .....	145
Abbildung 64: Lineare NSAA-Skala bei DMD (25) .....	166
Abbildung 65: Magnetresonanztomographie des Oberschenkels von Patienten mit DMD im Alter von 5-14 Jahren (62) .....	170
Abbildung 66: Fettgehalt des <i>M. vastus lateralis</i> von Patienten mit DMD im Alter von 5-14 Jahren (62).....	171
Abbildung 67: Beziehung zwischen Fettfraktion im <i>M. vastus lateralis</i> und dem Verlust der Gehfähigkeit (62) .....	171
Abbildung 68: Beziehung zwischen Alter und Verlust der Gehstrecke im 6MWT (62).....	172
Abbildung 69: Patientenfluss der Studie PTC124-GD-020-DMD (8).....	255

**Abkürzungsverzeichnis**

<b>Abkürzung</b>	<b>Bedeutung</b>
%	Prozent
$\Delta$	Delta
6MWD	Six-Minute Walk Distance
6MWT	Six-Minute Walk Test
ACC	American College of Cardiology
ACE	Angiotensin-Converting Enzyme
ADP	Ambulatory Decline Phase
AHA	American Heart Association
AM-NutzenV	Arzneimittel-Nutzenverordnung
ANCOVA	Kovarianzanalyse
ARB	Angiotensin-Rezeptorblocker
ATC	Anatomical Therapeutic Chemical
ATS	American Thoracic Society
BGBI	Bundesgesetzblatt
BL	Baseline
CHMP	Committee for Medicinal Products for Human Use
cm	Zentimeter
CONSORT	Consolidated Standards of Reporting Trials
CRO	Contract Research Organisation
CTCAE	Common Terminology Criteria of Adverse Events
CTD	Common Technical Document
CYP2C8	Cytochrom 2C8
CYP2C9	Cytochrom 2C9
d	Day(s)
d.h.	das heißt
DBMD	Duchenne-Becker Muskeldystrophie
df	Degrees of Freedom
DIMDI	Deutsches Institut für Medizinische Dokumentation
DMC	Data Monitoring Committee
DMD	Duchenne Muskeldystrophie

EG	Europäische Gemeinschaft
EKG	Elektrokardiogramm
EMA	Europäische Arzneimittelagentur
EU	Europäische Union
FDA	Food and Drug Administration
FVC	Forced Vital Capacity
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
GdS	Grad der Schädigungsfolge
GSGC	Gait, Stairs, Gowers, Chair
h	Hour(s) (Stunde[n])
HDL	High-Density Lipoprotein
HR	Hazard-Ratio
I <sup>2</sup>	Maß der Heterogenität
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ISE	Integrated Summary of Efficacy
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
LDL	Low-Density Lipoprotein
LOCF	Last Observation Carried Forward
LSM	Least Square Mean
m	Meter
M	Mittelwert
m <sup>2</sup>	Quadratmeter
M(C)ID	Minimal (Clinically) Important Difference
MedDRA	Medical Dictionary for Regulatory Activities)
mg	Milligramm
min	Minute(n)

mmHg	Millimeter Quecksilbersäule
mmol	Millimol
mMMRM	Modifiziertes MMRM Model ( <i>post-hoc</i> )
MMRM	Mixed-Model Repeated-Measures
mOsm	Milliosmolalität
MRS	Magnetresonanzspektroskopie
MTC	Mixed Treatment Comparison
n	Anzahl
NICE	National Institute for Health and Care Excellence
nmDMD	Duchenne Muskeldystrophie infolge einer Nonsense-Mutation
NSAA	North Star Ambulatory Assessment
p	Probability
p.o.	<i>per os</i> (oral)
PedQL	Pediatric Quality of Life Inventory
PK	Pharmakokinetik
PODCI	Pediatric Outcomes Data Collection Instrument
PP	per Protokoll
PRO	Patient-Reported Outcome
PT	Preferred Term
RANCOVA	Repeated-Measures Analysis of Covariance
RCT	Randomised Controlled Trial
REML	Restricted Maximum Likelihood
RR	relatives Risiko
S.	Seite
SAP	Statistischer Analyseplan
SD	Standardabweichung
SE	Standard Error of the Mean
sec	Sekunde(n)
SGB	Sozialgesetzbuch
SGB V	Sozialgesetzbuch Fünftes Buch
SOC	System Organ Class
Std.	standardisiert
STE	Surrogate Threshold Effects



STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TEAE	Treatment-Emergent Adverse Event
TFT	Timed Function Test
TREAT-NMD	Translational Research in Europe for the Assessment and Treatment of Neuromuscular Disease
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
u.a.	unter Anderem
UE	unerwünschtes Ereignis
USA	United States of Amerika
VerfO	Verfahrensordnung
vs.	versus
WHO	World Health Organization
z.B.	zum Beispiel
$\alpha$	alpha

#### 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 vom Typ Duchenne (DMD) aufgrund einer Nonsense-Mutation im Dystrophin-Gen bei gehfähigen Patienten im Alter von fünf Jahren und älter zugelassen. Bei nicht gehfähigen Patienten ist die Wirksamkeit nicht erwiesen. Die Nonsense-Mutation des Dystrophin-Gens muss mittels genetischer Methoden nachgewiesen werden (1). Bei der DMD handelt es sich um eine lebensbedrohliche X-chromosomal vererbte Form der Muskeldystrophie, bei der nicht funktionales Dystrophin exprimiert wird, welche letztendlich zu fortschreitendem Muskelabbau führt (1,2). Bei einem Anteil 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 (nmDMD). 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.

Translarna wurde im Jahr 2014 nach Artikel 14(7) der EU-Verordnung Nr. 726/2004 mit der Auflage zugelassen, die Ergebnisse einer konfirmatorischen Wirksamkeits- und Sicherheitsstudie mit Translarna vorzulegen (1). Die der auflagengebundenen Zulassung zugrunde liegende Phase 2b-Studie wurde bereits in einer früheren Nutzenbewertung durch den Gemeinsamen Bundesausschuss (G-BA) bewertet. Die Geltungsdauer des Beschlusses war bis zum 1. Juni 2016 befristet (3,4). In diesem Dossier wird das Ausmaß des Zusatznutzens einer Therapie mit Translarna in der zugelassenen Indikation auf Basis der konfirmatorischen Phase 3-Studie, sowie eine Zusammenfassung der Ergebnisse der Phase 2b und der Phase 3 Studie auf Endpunktebene in Meta-Analysen dargestellt.

Bei der DMD handelt es sich um ein seltenes Leiden. Am 27. Mai 2005 wurde dem Wirkstoff Ataluren nach Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 der Orphan-Drug Status zugesprochen (5). 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 (6).

Für die frühe Nutzenbewertung von Translarna ist daher keine zweckmäßige Vergleichstherapie zu benennen. Die Bewertung von Translarna vs. Placebo („Best supportive care“) basiert auf den patientenrelevanten Zielgrößen Morbidität, Gesundheitsbezogene Lebensqualität und Nebenwirkungen (7).

## Datenquellen

Die Phase 2b-Studie PTC124-GD-007-DMD, die die Grundlage der Zulassung von Translarna mit Auflagen bildete, wurde in einem früheren Dossier zur Nutzenbewertung dargestellt und wurde bereits durch den G-BA bewertet (Beschluss am 21. Mai 2015 in Kraft getreten) (3,4). Hinsichtlich der Studie PTC124-GD-007-DMD wird auf das frühere Verfahren verwiesen.

Die in diesem Dossier berücksichtigten Studien (PTC124-GD-007-DMD und PTC124-GD-020-DMD) sind randomisierte, doppelt-verblindete Placebo-kontrollierte Studien mit einer Studiendauer von 48 Wochen, in die gehfähige Patienten mit nonsense-Mutation-bedingter DMD (nmDMD) eingeschlossen wurden. Im vorliegenden Dossier werden Studiendaten aus den Studienberichten der Placebo-kontrollierten Phase 2b-Studie PTC124-GD-007-DMD und der Phase 3-Studie PTC124-GD-020-DMD berücksichtigt (8,9). Für dieses Dossier wurden für die Studie PTC124-GD-007-DMD Ergebnisse aus der kITT Population unter Verwendung des prädefinierten MMRM Modells berechnet (10–13) (s. Kapitel 4.3.1.2.1). Für die Meta-Analysen werden zusätzliche Auswertungen aus den CTD-Modulen 2.7.3 und 2.7.4 der Studie PTC124-GD-020-DMD sowie deren Anhänge herangezogen (14–16). Die Quantifizierung des Zusatznutzens von Translarna in dem zugelassenen Anwendungsgebiet basiert schwerpunktmäßig auf den Ergebnissen der Meta-Analysen beider pivotalen Studien.

Für Arzneimittel zur Behandlung seltener Leiden sind nur die Abschnitte 4.4.2 und 4.4.3 auszufüllen (17); daher sind eine systematische Literaturrecherche sowie Recherchen in Patientenregistern nicht erforderlich. Die Ergebnisse der patientenrelevanten Endpunkte sind jedoch in Abschnitt 4.3.1.3 dargestellt und werden in Kapitel 4.4.2 weiter diskutiert.

## 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 (17). Somit entfällt 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, werden in der folgenden Nutzenbewertung nur randomisierte kontrollierte Studien berücksichtigt (Studie PTC124-GD-007-DMD siehe frühere Nutzenbewertung; (3)).

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

Die hier gegenständliche zulassungsrelevante randomisierte, kontrollierte klinische Studie PTC124-GD-020-DMD (Studie 020) wurde gemäß der 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 (für Studie PTC124-GD-007-DMD [Studie 007] siehe frühere Nutzenbewertung; (3)). Die Bewertung der Aussagekraft der Nachweise erfolgte für die eingeschlossene, kontrollierte Studie entsprechend der Vorgaben der Verfahrensordnung (VerfO) des G-BA (6) zur Beurteilung des Verzerrungspotentials endpunktspezifisch und endpunktübergreifend

(Abschnitt 4.3, Anhang 4-F). In Anhang 4-F wurde das Verzerrungspotential nur für diejenigen Endpunkte bewertet, die in einer prospektiv geplanten Meta-Analyse ausgewertet wurden (14). 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 (6). Der primäre Endpunkt war in beiden zulassungsrelevanten Studien die Änderung der Gehstrecke im 6-Minuten-Gehtest ("6-minute walk test", 6MWT) von Baseline (BL) bis Woche 48 unter oraler Therapie mit Ataluren vs. Placebo. 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 deterministische, irreversible Krankheitsprogression angesehen (1). Da für den 6MWT keine skalenspezifischen Irrelevanzschwellen vorlagen, wurde für die Relevanzbewertung der Therapieeffekte auf ein allgemeines statistisches Maß in Form von standardisierten Mittelwertsdifferenzen (Hedges'  $g$ ) zurückgegriffen (Kapitel 4.4.2.5.3).

Als Sekundärendpunkte wurden die Zeit bis zu einer 10%igen Verschlechterung der 6MWD und 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 erhoben. Diese Tests sind einfach anzuwenden und reproduzierbar, sie sind prädiktiv in Bezug auf klinische bedeutsame „Meilensteine“ der Krankheitsprogression (21), und zeigen eine verlässliche Response auf eine Intervention mit Kortikosteroiden (22,23). TFTs werden als Maß angesehen, inwieweit ein DMD-Patient in der Lage ist, wichtige Alltagstätigkeiten zu verrichten (1). Für die TFTs wurden MCIDs publiziert, die auf statistischen Verteilungsmethoden basieren (18,20,24). Für die Relevanzbewertung der Therapieeffekte aus den Treppensteigttests wurde aus diesem Grund auf das allgemeine statistische Maß in Form von standardisierten Mittelwertsdifferenzen (Hedges'  $g$ ) zurückgegriffen (s. Kapitel 4.4.2.5.3).

Das „North Star Ambulatory Assessment“ (NSAA), welches in der Studie 020 eingesetzt wurde, ist ein für die DMD validiertes Instrument zur Messung verschiedener körperlicher Funktionen, welches von Experten als für die Patienten und ihre Familien sehr wichtig eingestuft wird, da es Aktivitäten des täglichen Lebens erfasst (24). Die für das NSAA vorgeschlagenen MCIDs basieren auf indirekten Methoden (18,25), sodass sie noch mit einer gewissen Unsicherheit behaftet sind.

Der 6MWT, die TFTs und das NSAA sind laut Experteneinschätzung die einzigen Instrumente, die alle Kriterien erfüllen, die an ein Messinstrument in Phase 2/3-Studien bei gehfähigen DMD-Patienten zu stellen sind, und sie sind klinisch sinnvoll (24). Die Änderung der Gehstrecke im 6MWT, die TFTs und das NSAA sind der Zielgröße Verbesserung des Gesundheitszustands zuzuordnen.

Aktivitäten des täglichen Lebens und Krankheitssymptome („patient-reported outcome“, PRO) wurden in der Studie 020 mittels eines vom pharmazeutischen Unternehmer speziell entwickelten Surveys erfasst. Die EMA empfiehlt die Erfassung der Aktivitäten des täglichen Lebens in der Richtlinie zur klinischen Entwicklung bei DBMD (2) und die Patientenrelevanz von Alltagsaktivitäten wurde jüngst während eines internationalen Meetings bestätigt (24). Auch der G-BA stuft die Partizipation im Alltag als für die Betroffenen von großer Bedeutung ein (3). Es liegen allerdings keine Informationen zur Validität oder MCIDs für das in der Studie 020 eingesetzte Instrument vor.

Daten zur Lebensqualität wurden in der Studie 020 mittels PODCI (und in der Studie 007 mittels PedsQL) erhoben. Die Domänen Transfer/Basismobilität und Sport/körperliche Funktion des PODCI, die in der Studie 020 erfasst wurden, zeigten bei DMD-Patienten signifikante Assoziationen mit dem Alter (und entsprechend der Krankheitsprogression) sowie klinischen Endpunkten (z.B. TFTs oder 6MWT) (20,26). Auch die EMA stuft gewisse Domänen des PODCI als nützliche Instrumente bei gehfähigen DMD-Patienten ein (2). Das PODCI hat sich bei DMD-Patienten als sensitiver erwiesen als das in der Studie 007 eingesetzte PedsQL-Instrument und eine MCID wurde vorgeschlagen (20). Die offensichtlichen Schwierigkeiten, bei Kindern die patientenberichtete Lebensqualität zu erfassen, erklären zumindest teilweise die Tatsache, dass für die DMD derzeit noch kein validiertes Instrument existiert, das die Veränderung der gesundheitsbezogenen Lebensqualität innerhalb eines Jahres verlässlich erfassen kann (27). Ein vor kurzem publizierter Review kommt zu der Schlussfolgerung, dass in den bislang durchgeführten Studien in DMD Patienten nur die physischen Domänen der Lebensqualitäts-Instrumente mit dem Krankheitsstadium korrelieren. Für psychosoziale Domänen ist dies nicht der Fall. Auch ist die Übereinstimmung der Patientenbewertung mit der Bewertung der Eltern generell sehr schlecht (28). Die klassischen Methoden zur Erfassung der gesundheitsbezogenen Lebensqualität erscheinen derzeit nicht ausreichend geeignet, um Unterschiede in der Krankheitsprogression - geschweige denn Therapieeffekte - erfassen zu können. Zumindest in der frühen Nutzenbewertung sollte jedoch der Endpunktkategorie „Morbidity“ für diese schwerwiegende Erkrankung ein deutlich höheres Gewicht zugemessen werden als der „gesundheitsbezogenen Lebensqualität“, da eine therapeutische Verlangsamung der Krankheitsprogression für die Patienten von vorrangiger Bedeutung ist.

Die patientenrelevante Zielgröße Mortalität konnte in den 48-Wochenstudien auf Grund der zeitlich eingeschränkten Studiendauer und angesichts der Seltenheit der Erkrankung nicht evaluiert werden.

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

Tabelle 4-1: Darstellung der patientenrelevanten Endpunkte

Patientenrelevante Zielgrößenkategorie	Endpunktparameter
Krankheitsbedingte Symptome (Morbidität)	<ul style="list-style-type: none"> <li>– Gehstrecke im 6MWT bis Woche 48</li> <li>– Zeit bis zu einer mindestens 10%igen 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> <li>– Körperliche Funktion mittels „North Star Ambulatory Assessment“ (NSAA)<sup>1</sup> inclusive Zeit, die benötigt wird, um aus der Rückenlage aufzustehen</li> <li>– Aktivitäten des täglichen Lebens und Krankheitssymptome<sup>1</sup></li> </ul>
Gesundheitsbezogenen Lebensqualität	<ul style="list-style-type: none"> <li>– PODCI<sup>1</sup></li> </ul>
Nebenwirkungen	<ul style="list-style-type: none"> <li>– Schwerwiegende (bzw. schwere) Symptome (bzw. Folgekompli-kationen) und Nebenwirkungen</li> <li>– Nicht schwerwiegende (bzw. nicht schwere) Symptome (bzw. Folgekompli-kationen) und Nebenwirkungen</li> </ul>

<sup>1</sup> Das NSAA und PODCI wurden nur in der hier gegenständlichen Studie 020 erfasst und das Instrument zur Erfassung von Alltagsaktivitäten unterschied sich zwischen den Studien deutlich. Diese Endpunkte konnten daher nicht meta-analytisch ausgewertet werden. Die Zeit, um aus dem Liegen aufzustehen, wurde in Studie 020 als Teil des NSAA erfasst und in Studie 007 als TFT.

In dem vorliegenden Nutzendossier werden für die patientenrelevanten Endpunkte, die in den beiden eingeschlossenen Studien 007 und 020 gemeinsam erhoben wurden, die Ergebnisse von Meta-Analysen dargestellt. Ob eine statistische Zusammenfassung der jeweiligen Effektschätzer sinnvoll war wurde mittels Heterogenitätstests geprüft. In der Regel wurde von einer statistischen Zusammenfassung abgesehen, falls der Heterogenitätstest einen p-Wert unter 0.05 liefert (7). Meta-Analysen aus mindestens zwei homogenen Studien können je nach qualitativer Ergebnissicherheit eine sehr hohe Evidenz bis hin zu einem Beleg liefern (IQWiG Allgemeine Methoden – Version 4.2) (7). Für die beiden eingeschlossenen Studien ergibt sich auf Studienebene sowie auf Endpunktebene ein geringes Verzerrungspotential. Formell liefern die Ergebnisse der vorgelegten Meta-Analysen für den zu bewertenden Wirkstoff Belege auf der höchsten Evidenzstufe 1a.

Bei der Planung beider Translarna-Studien stand der Intention, eine möglichst homogene Patientengruppe zu rekrutieren, selbstverständlich das Anliegen gegenüber, Patienten ohne andere Therapieoptionen eine Behandlung nicht aufgrund enger Einschlusskriterien vorzuenthalten oder die Rekrutierung unnötig zu verzögern. Die prospektiv geplanten Subgruppen-Analysen sollten der Tatsache Rechnung tragen, dass gehfähige Patienten in nahezu allen Krankheitsstadien in die Studien eingeschlossen waren, sowie den Hinweisen aus der Phase 2b Studie, dass Therapieeffekte nicht in allen Krankheitsstadien mit vergleichbarer Sensitivität erhoben werden können. Der Hinweis, dass die Assay-Sensitivität in der Phase 2b Studie mit den unterschiedlichen Krankheitsphasen der Erkrankung variiert, hat sich mit den Studiendaten der Phase 3 Studie eindrücklich bestätigt. In der Studie 020 waren Subgruppen-Analysen nach 6MWD zu BL prädefiniert (u.a. für die Kategorien <300

m,  $\geq 300$  bis  $< 400$  m,  $\geq 400$  m). Dieses Vorgehen entspricht grundsätzlich der inzwischen vorliegenden Richtlinie der EMA („Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy“), die einerseits empfiehlt, ein möglichst breites Spektrum an Patienten in eine confirmatorische Studie einzuschließen, andererseits angesichts der hohen Heterogenität in der Patientenpopulation Subgruppen-Analysen zu definieren. Es sollten verschiedene Cut-off-Werte auf angemessenen Skalen eingesetzt werden, um die Assay-Sensitivität zu gewährleisten. Die EMA schließt eine Extrapolation klinischer Daten von z.B. älteren auf jüngere Kinder oder *vice versa* unter bestimmten Bedingungen nicht grundsätzlich aus, insbesondere, wenn dies auf Basis des Wirkmechanismus begründbar ist (2). Eine Extrapolation der klinischen Ergebnisse von Ataluren auf Patientenpopulationen, die nicht in den klinischen Studien repräsentiert waren, erscheint zumindest daher prinzipiell gerechtfertigt. Auch die FDA weist in ihrem aktuellen Richtlinienentwurf darauf hin, dass Daten bei relativ geringen Patientenzahlen über verschiedene Krankheits-Subgruppen eine breitere Indikation stützen können (29).

Das Ausmaß des Zusatznutzens von Translarna wurde aus den beiden für die Zulassung relevanten Studien sowie den Meta-Analysen für die gemeinsamen patientenrelevanten Endpunkte nach Maßgabe der in Kapitel 5, § 5 Absatz 7 Satz 1 Nr. 1 bis 4 der Verfo (2015) festgelegten Kriterien hinsichtlich der therapeutischen Relevanz (qualitativ) bewertet (6).

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

Zur Bewertung des Ausmaßes des Zusatznutzens von Ataluren werden die Ergebnisse von Meta-Analysen unter Einschluss der beiden zulassungsrelevanten Studien (007 und 020) herangezogen. Das Design beider Studien war sehr ähnlich (s. Kapitel 4.4.2.5.1), und beide Studien hatten die identische Fragestellung. Es sollte die Wirksamkeit und Sicherheit von Ataluren bei einer Dosierung von 40 mg/kg/Tag in Patienten untersucht werden, die an einer durch nonsense-Mutationen verursachten Duchenne-Muskeldystrophie leiden. Die Studiendauer war mit 48 Wochen für beide Studien identisch, und beide Studien hatten denselben primären Endpunkt (Veränderung der 6MWD über 48 Wochen). Auch die sekundären Endpunkte (TFTs) waren weitgehend identisch. Die Auswertungsmethoden der Endpunkte waren entweder identisch (Hazard Ratio für den Endpunkt 10%ige Verschlechterung im 6MWT, relatives Risiko für unerwünschte Ereignisse) oder methodisch eng verwandt (MMRM-Model oder ANCOVA Model für den primären Endpunkt und andere sekundäre Endpunkte).

Bei den eingeschlossenen Studien handelt es sich um randomisierte, Placebo-kontrollierte Doppelblindstudien der Phase 2b bzw. der Phase 3 und der Evidenzstufe 1b. Das Verzerrungspotential wird für die eingeschlossenen Studien endpunktübergreifend als gering eingeschätzt. Das Verzerrungspotential wird für die primär relevanten Endpunkte Gehstrecke im 6MWT, TFTs und unerwünschte Ereignisse ebenfalls als gering eingestuft (siehe auch vorheriges Dossier zur Nutzenbewertung für die Studie 007). Die Meta-Analysen der Studien 007 und 020 sind nach der in der Verfo des G-BA dargestellten Evidenzhierarchie der Evidenzstufe 1a zuzuordnen.



Die Therapieeffekte waren für alle patientenrelevanten Endpunkte in der Gesamtpopulation zwischen den eingeschlossenen Studien homogen. Daher war es gerechtfertigt, die Ergebnisse beider Studien in Meta-Analysen zusammenzufassen, und die Gesamteffektschätzer für die Bewertung des Zusatznutzens heranzuziehen (s. Kapitel 4.3.1.3.1). Für fünf der sechs Morbiditätsendpunkte (Veränderung der Gehstrecke im 6MWT; Zeit bis zur 10%igen Verschlechterung der 6MWD; Zeit, um 4 Treppenstufen hinaufzusteigen; Zeit, um 4 Treppenstufen hinabzusteigen; und Zeit, um 10 m zu gehen oder zu laufen zeigten die Meta-Analysen signifikante Überlegenheit von Ataluren gegenüber Placebo. Da eine 10%ige Veränderung im 6MWT als patientenrelevante Veränderung bewertet wird, ergab sich für diesen Endpunkt in der Gesamtpopulation ein quantifizierbarer Zusatznutzen (Tabelle 4-2, Tabelle 4-83).

In der Studie 020 war im Erhebungsinstrument NSAA ein nicht signifikanter Trend zur Überlegenheit von Ataluren erkennbar, ebenso zeigte das Instrument PODCI für zwei Domänen einen nicht signifikanten Trend zur Überlegenheit (s. Tabelle 4-2). Eine Meta-analyse konnte für die beiden Endpunkte nicht durchgeführt werden, da sie nur in der Studie 020 erhoben wurden.

Die Verträglichkeit von Ataluren war insgesamt sehr gut, und die Inzidenz der unerwünschten Ereignisse (UEs) unterschied sich zwischen Ataluren und Placebo in den Meta-Analysen nicht signifikant. Hinsichtlich der Zahl der schweren (schwerwiegenden) UEs ergaben sich in einer gepoolten Analyse beider Studien keine Unterschiede zwischen Ataluren und Placebo. Keines der schwerwiegenden UEs stand mit der Behandlung mit Ataluren in Zusammenhang.

Tabelle 4-2: Übersicht der Ergebnisse für patientenrelevante Endpunkte aus PTC124-GD-007-DMD und PTC124-GD-020-DMD (gesamte Studienpopulationen, Meta-Analysen)

Endpunktkategorie	Endpunkt	Analysepopulation (s. Kapitel 4.3.1.2.1)	Effektschätzer Ataluren vs. Placebo (Woche 48)
Morbidity	Veränderung der Gehstrecke im 6MWT	Meta-Analyse der Gesamtpopulationen: kITT (007)/ITT (020)	$\Delta$ LSM: 18.21 m [95% KI: 1.55, 34.86] <b>p=0.032</b>
	Zeit bis zur 10%igen Verschlechterung der 6MWD	Meta-Analyse der Gesamtpopulationen: kITT (007)/ITT (020)	HR: 0.67 [95% KI: 0.48; 0.94] <b>p=0.021</b>
	Zeit, um 4 Treppenstufen hinaufzusteigen	Meta-Analyse der Gesamtpopulationen: kITT (007)/ITT (020)	$\Delta$ LSM: -1.74 sec [95%KI: -2.97, -0.51] <b>p=0.0056</b>
	Zeit, um 4 Treppenstufen hinabzusteigen	Meta-Analyse der Gesamtpopulationen: kITT (007)/ITT (020)	$\Delta$ LSM: -1.89 sec [95%KI: -3.20; -0.59] <b>p=0.004</b>
	Zeit, um 10 m zu gehen oder zu laufen	Meta-Analyse der Gesamtpopulationen: kITT (007)/ITT (020)	$\Delta$ LSM: -1.16 sec [95% KI: -2.28; -0.03] <b>p=0.043</b>
	NSAA Gesamtscore, transformiert	ITT; Studie 020	$\Delta$ LSM: 1.51 [95% KI: -1.16; 4.17] p=0.268
Gesundheitsbezogene Lebensqualität	PODCI Transfer/Basismobilität	ITT; Studie 020	$\Delta$ LSM: 1.64 [95% KI: -2.11; 5.39] p=0.391
	PODCI Sport/körperliche Funktion	ITT; Studie 020	$\Delta$ LSM: 2.15 [95% KI: -1.75; 6.05] p=0.281
Verträglichkeit	TEAEs inklusive schwerwiegender (bzw. schwerer) TEAEs	Meta-Analyse der Gesamtpopulationen: As treated (007) / as treated (020)	Keine statistisch signifikanten Unterschiede

Da eine mögliche Effektmodifikation durch die BL-Gehstrecke im 6MWT schon *a priori* bekannt war, wurde im SAP die Subgruppenanalyse für die Subgruppenmerkmale BL-Gehstrecke <300 m,  $\geq$ 300 m bis  $\leq$ 400 m und >400 m definiert (zur Rationale siehe Kapitel 4.4.2.4). Subgruppenanalysen für prädefinierte Subgruppen werden in diesem Dossier nur für diejenigen Endpunkte berichtet, für die statistisch signifikante Therapieeffekte in den gesamten Studienpopulationen nachweisbar waren, und für die eine signifikante Interaktion zwischen den Subgruppen in einer Meta-Analyse demonstriert werden konnte. Dabei handelt es sich um die Endpunkte Veränderung der Gehstrecke im 6MWT, Zeit um 4 Treppenstufen

hinaufzusteigen; und Zeit um 4 Treppenstufen hinabzusteigen (für den Endpunkt Zeit um 10 m zu gehen oder zu laufen ergaben sich ähnliche Subgruppeneffekte wie für die anderen Endpunkte, jedoch war der Interaktionstest nicht signifikant).

Die Resultate der Subgruppenanalysen zeigten hoch signifikante und patientenrelevante Therapieeffekte in der Subgruppe  $\geq 300$  m bis  $\leq 400$  m (Tabelle 4-3), während die Therapieeffekte in den beiden anderen Subgruppen nicht statistisch signifikant waren (s. Kapitel 4.3.1.3.2). Auffallend war die Konsistenz der Ergebnisse der Subgruppenanalysen sowohl zwischen den Studien als auch zwischen den unterschiedlichen Endpunkten. Die Ergebnisse der Subgruppenanalysen führten daher zu prädiktiven Arbeitshypothesen über die Bedeutung eines Therapiebeginns mit Ataluren in Abhängigkeit vom Krankheitsstadium, und die zu erwartenden Therapieeffekte in klinischen Studien:

In klinischen Studien mit einer Behandlungsdauer von 48 Wochen ist bei jungen DMD-Patienten in frühen Krankheitsstadien (BL-Gehstrecken von  $>400$  m) zu berücksichtigen, dass ein möglicher Behandlungseffekt durch Wachstum und Entwicklung und die dadurch bedingte zunehmende Gehstrecke überlagert wird (30). Zudem wird ein Behandlungseffekt in diesen Patienten erst deutlich werden, sobald der Schädigungsgrad des Muskelgewebes soweit fortgeschritten ist, dass er mit den verwendeten Messinstrumenten überhaupt erfassbar wird. Patienten in der mittleren Krankheitsphase mit schon erkennbar eingeschränkter Mobilität und einer progressiven Verschlechterung der Krankheitssymptomatik (BL-6MWD von  $\geq 300$  m bis  $<400$  m) nehmen eine distinkte Sonderstellung ein. Bei diesen Patienten ist die Sensitivität, einen Therapieeffekt nachweisen zu können, am höchsten, da das Krankheitsbild für diese Patienten nicht mehr stabil, sondern progredient ist. Demgegenüber verschlechtert sich bei Patienten in der späten ambulatorischen Phase (BL-6MWD unter 300 m) die Muskelfunktion über verschiedene Testsysteme so rasch, und der Anteil des intakten Muskelgewebes ist schon so gering, dass eine therapeutisch induzierte Re-Expression von Dystrophin nur noch limitierte Erfolge zeigen kann (s. Kapitel 4.4.2.4). Der Einschluss von Patienten mit BL Gehstrecken  $>400$  m oder  $<300$  m wird folglich zu einer reduzierten Assay-Sensitivität in klinischen Studien führen, in denen die hier eingesetzten etablierten Messinstrumente verwendet werden. Der Einsatz einer Dystrophin-restaurierenden Therapie sollte daher so bald wie möglich erfolgen, um eine irreversible Schädigung des Muskelgewebes schon möglichst frühzeitig zu verzögern. Die Entwicklung von prognostischen und prädiktiven Biomarkern, die eine Schädigung des gesunden Muskelgewebes in den frühen Krankheitsphasen erfassen können, in denen die hier eingesetzten klinischen Testsysteme noch keine ausreichende Sensitivität zeigen, wird zukünftig an Bedeutung gewinnen.

In diesem Zusammenhang ist wichtig zu betonen, dass auf Grund des Wirkmechanismus der eingesetzten Intervention – unabhängig von der Krankheitsphase - ein umfassender therapeutischer Effekt auf den Krankheitsverlauf zu erwarten ist. Positive Therapieeffekte sind nicht nur für die Skelettmuskulatur, sondern auch für die Atemmuskulatur, und den Herzmuskel zu erwarten, also auch für Rollstuhl-pflichtige Patienten, die in den vorgelegten Studien jedoch nicht eingeschlossen waren.

Tabelle 4-3: Übersicht der Ergebnisse für patientenrelevante Endpunkte aus PTC124-GD-007-DMD und PTC124-GD-020-DMD (Subgruppen, Meta-Analysen)

Endpunktkategorie	Endpunkt	Analysepopulation (s. Kapitel 4.3.1.2.1)	Effektschätzer Ataluren vs. Placebo (Woche 48)
Morbidität	Veränderung der Gehstrecke im 6MWT	Meta-Analyse der Subpopulationen $\geq 300$ m bis $< 400$ m BL-Gehstrecke kITT (007)/ITT (020)	$\Delta$ LSM: 45.33 m [95% KI: 20.16; 70.49] <b>p=0.0004</b> Hedges' g: 0.581 [95% KI: 0.246; 0.917]
	Zeit, um 4 Treppenstufen hinaufzusteigen	Meta-Analyse der Subpopulationen $\geq 300$ m bis $< 400$ m BL-Gehstrecke kITT (007)/ITT (020)	$\Delta$ LSM: -3.47 sec [95% KI: -5.35; -1.59] <b>p=0.0003</b> Hedges' g: -0.598 [95% KI: -0.934; -0.263]
	Zeit, um 4 Treppenstufen hinabzusteigen	Meta-Analyse der Subpopulationen $\geq 300$ m bis $< 400$ m BL-Gehstrecke kITT (007)/ITT (020)	$\Delta$ LSM: -4.31 sec [95% KI: -6.25; -2.37] <b>p=0.00001</b> Hedges' g: -0.720 [95% KI: -1.060; -0.381]

### 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 (31,32). Die nmDMD ist mit erheblichen Einschränkungen der Ausdauer und Mobilität bis hin zum Verlust der Gehfähigkeit assoziiert, sodass der Erhalt der Gehfähigkeit für jeden Patienten ein primäres Therapieziel darstellt (18). 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. Das Alter zum Zeitpunkt des Verlusts der Gehfähigkeit determiniert bei Jungen mit DMD den Beginn weiterer schwerer Folgekomplikationen wie z.B. Skoliose und respiratorische Insuffizienz (33). Bei dem progressiven und irreversiblen Verlust der Gehfähigkeit handelt es sich daher um ein sehr schwerwiegendes Krankheitssymptom. Die körperlichen Fähigkeiten gehen bei DMD in der Regel in einer zeitlich geordneten Reihenfolge verloren, sodass die Patienten mit Fortschreiten der Erkrankung im täglichen Leben mehr und mehr beeinträchtigt werden und zunehmend an Autonomie verlieren (25). Die Zeit, die benötigt wurde, um 10 m zu gehen/laufen, war mit dem Erhalt bzw. Verlust der Gehfähigkeit in den nächsten 12 Monaten assoziiert (19). Daher erfassen auch die TFTs die schwere Symptomatik der Erkrankung.

Bei einer sehr seltenen Erkrankung, bei der sehr große Studien nicht möglich sind, liefern Meta-Analysen aus Studien des höchsten Evidenzgrads mit niedrigem Verzerrungspotential

eine solide und belastbare Evidenz der Wirksamkeit und Verträglichkeit. Die hohe Konsistenz der vorgelegten Ergebnisse belegt die Robustheit des Behandlungseffekts von Ataluren. Aus den Meta-Analysen kann in Einklang mit den seitens des IQWiGs festgelegten Kriterien zur Ableitung von Aussagesicherheiten für Endpunkte mit demonstriertem Zusatznutzen die höchste Ergebnissicherheit abgeleitet werden (7).

Für den Endpunkt Zeit bis zur 10%igen Verschlechterung im 6MWT ergibt sich für die Gesamtpopulation in striktem Einklang mit den Kriterien des § 5 Absatz 7 des 5. Kapitels der Verfo (6) ein beträchtlicher Zusatznutzen für Ataluren. Der Endpunkt repräsentiert eine gegenüber der Vergleichstherapie bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens, insbesondere eine Abschwächung schwerwiegender Symptome (Tabelle 4-83).

Für den Endpunkt Veränderung der Gehstrecke im 6MWT ergibt sich aus dem Gesamteffektschätzer für die Subgruppe  $\geq 300$  m bis  $< 400$  m ein beträchtlicher Zusatznutzen. Auch dieser Endpunkt repräsentiert eine gegenüber der Vergleichstherapie bisher nicht erreichte deutliche Verbesserung des therapierelevanten Nutzens, insbesondere eine Abschwächung schwerwiegender Symptome.

Der Zusatznutzen für den Endpunkt Zeit um 10 m zu gehen oder zu laufen Endpunkt ergibt sich aus dem Gesamteffektschätzer der Meta-Analyse unter Einschluss der Gesamtpopulationen der eingeschlossenen Studien. Da für diesen Endpunkt kein formal validiertes Responsekriterium existiert, ist der Zusatznutzen als nicht quantifizierbar auszuweisen.

Für die Endpunkte Zeit, um 4 Treppenstufen hinaufzusteigen, und Zeit, um 4 Treppenstufen hinabzusteigen ergibt sich aus den Gesamteffektschätzern für die Subgruppe  $\geq 300$  m bis  $< 400$  m jeweils ein geringer Zusatznutzen. Die Endpunkte repräsentieren eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapierelevanten Nutzens im Sinne von § 3 Absatz 1 erreicht wird, insbesondere eine Verringerung von nicht schwerwiegenden Symptomen der Erkrankung (Tabelle 4-83).

Aus den Ergebnissen der Subpopulation mit einer BL-Gehstrecke von  $\geq 300$  bis  $< 400$  m ist es gerechtfertigt, Aussagen zum Ausmaß des Zusatznutzens von Ataluren für die Gesamtpopulation der Patienten in der zugelassenen Indikation abzuleiten (s. Kapitel 4.4.2.4). Dieses Vorgehen steht grundsätzlich im Einklang mit den aktuellen Empfehlungen der EMA und der FDA, die den Einschluss nicht zu restriktiver Patientenkollektive und prä-spezifizierte Subgruppen-Analysen empfehlen (2,30). Die EMA empfiehlt verschiedene Cut-Offs auf angemessenen Skalen, um die Assay-Sensitivität zu gewährleisten (2). In den beiden eingeschlossenen Studien 007 und 020 zeigte die Gruppe mit BL-Gehstrecken von  $\geq 300$  bis  $< 400$  m mit einem hohen Maß an Übereinstimmung die höchste Assay-Sensitivität für die eingesetzten Messinstrumente.

Die Meta-Analysen der UEs aus den Studien 007 und 020 liefern keine Hinweise auf ein relevantes Schadenspotential von Ataluren.

**Zusammenfassend wird das Ausmaß des Zusatznutzens von Translarna (Ataluren) in einer Dosierung von 10/10/20 mg/kg KG/Tag in der zugelassenen Indikation als beträchtlich eingestuft, da eine therapeutisch 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 Beutel) in der zugelassenen Indikation zur Behandlung der genetisch gesicherten durch eine Nonsense-Mutation verursachten Muskeldystrophie Typ Duchenne (nmDMD) bei gehfähigen Patienten im Alter von fünf Jahren und älter zu bewerten. Translarna erhielt im Jahr 2014 eine auflagengebundene europäische Zulassung, die den pharmazeutischen Unternehmer zur Durchführung einer konfirmatorischen Studie bei nmDMP verpflichtete (1). Der Gemeinsame Bundesausschuss (G-BA) hat das Ausmaß des Zusatznutzens von Translarna in der zugelassenen Indikation bereits in einem früheren Verfahren auf Basis einer Phase 2b-Studie (Studie 007) bewertet (3). Mit Beschluss vom 21.05.2015 stufte der G-BA das Ausmaß des Zusatznutzens von Translarna in der zugelassenen Indikation als gering ein. Die Geltungsdauer des Beschlusses war bis zum 1. Juni 2016 befristet (4). Die vorliegende Nutzenbewertung bezieht sich auf die konfirmatorische Phase 3-Studie (Studie 020) sowie die Meta-Analysen, welche die Phase 2b Studie und die Phase 3 Studie einschlossen. Hinsichtlich der Phase 2b-Studie wird auf das frühere Nutzendossier verwiesen.

Bei der DMD handelt es sich um eine seltene, lebensbedrohliche X-chromosomal vererbte Form der Muskeldystrophie, bei der kein funktionales Dystrophin exprimiert wird, welche zu

einem irreversibel fortschreitendem Muskelabbau führt (1,2). Bei einem Teil der DMD-Patienten wird die Erkrankung durch ein zusätzliches Stop-Codon im Dystrophin-Gen 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.

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. 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 G-BA) (6). 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.

### **Patientenpopulation**

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.

Zum Zeitpunkt der Planung der Phase 2b-Studie lagen nur begrenzte Daten zum Krankheitsverlauf der DMD vor, sodass in Übereinstimmung mit der später veröffentlichten Richtlinie der EMA (2) ein relativ breites und entsprechend heterogenes Patientenkollektiv in die Studie eingeschlossen wurde. Diesem Vorgehen lag selbstverständlich auch die Intention zugrunde, Patienten ohne Therapieoption eine Behandlung nicht vorzuenthalten, und die Rekrutierung nicht unnötig zu verzögern. Die Krankheitsprogression ist bei DMD hoch variabel und hängt neben dem Alter entscheidend von der Schwere der Erkrankung ab, z.B. der Baseline-6-Minuten-Gehstrecke (BL-6MWD). Grundsätzlich gilt, dass je schlechter die Ausgangssituation desto schlechter ist die Prognose und desto höher ist das Risiko einer raschen Progression bis hin zum vollständigen Gehverlust. Entsprechend bleiben Patienten mit guter initialer Gehfähigkeit in der Regel über ein Jahr symptomatisch stabil (19). Daher stellt die Studienplanung bei DMD eine Herausforderung dar, insbesondere die Selektion eines hinreichend homogenen aber nicht zu restriktiven Kollektivs, bei dem über eine bei dieser sehr schwerwiegenden Erkrankung ethisch vertretbare Studiendauer ein Therapieeffekt gegenüber „Best supportive care“ messbar ist. Bei der Planung der konfirmatorischen Phase 3-Studie wurden die Erkenntnisse der Phase 2b-Studie sowie die Daten anderer Arbeitsgruppen zum natürlichen Verlauf der DMD berücksichtigt, und im Vergleich mit der vorausgegangenen Phase 2b-Studie wurden etwas engere Einschlusskriterien definiert. In der konfirmatorischen Phase 3-Studie wurden Patienten mit einem hohen Risiko einer Krankheitsprogression (sogenannte „ambulatory decline phase“, ADP) rekrutiert, bei denen erwartungsgemäß innerhalb der Studiendauer von einem Jahr ein Behandlungseffekt einer



Muskel-stabilisierenden Therapie nachweisbar sein würde. Da auch bei Planung der Phase 3-Studie 020 noch Unsicherheiten hinsichtlich der optimalen Patientenselektion bestanden, und Hinweise auf Effektmodifikation durch die Baseline-Gehstrecke bestanden, wurden prospektiv Subgruppen-Analysen (z.B. BL-6MWD <300, ≥300 bis <400, ≥400 m) für den Fall definiert, dass sich die gewählten Einschlusskriterien als zu breit und das Patientenkollektiv als zu heterogen erweisen sollten.

### **Intervention**

In beiden zulassungsrelevanten Studien nahmen die Patienten dreimal täglich 10, 10, 20 mg/kg KG Ataluren (40 mg/kg KG pro Tag) oder Placebo über 48 Wochen oral ein. Die in der Studie 007 zusätzlich geprüfte nicht zugelassene Dosierung von Ataluren von 20, 20, 40 mg/kg KG (80 mg/kg KG pro Tag) findet in dem vorliegenden Dossier zur Nutzenbewertung keine Berücksichtigung.

### **Vergleichstherapie**

Nach § 35a SGB V Abs. 1 Satz 10 Halbsatz 1 entfällt bei Arzneimitteln, die für seltene Leiden zugelassen sind, die Verpflichtung zum Nachweis des medizinischen Zusatznutzens im Verhältnis zur zweckmäßigen Vergleichstherapie. Eine zweckmäßige Vergleichstherapie ist für Ataluren nicht zu benennen. Bei DMD ist keine andere kausal wirksame Therapie zugelassen, sodass die Wahl von Placebo plus Standardtherapie als Vergleichstherapie angemessen war. Nur Patienten, die mit Kortikosteroiden vorbehandelt waren, konnten an der Studie 020 teilnehmen. Die häufig eingesetzten Kortikosteroide sind in Europa für DMD formal nicht zugelassen (2).

### **Patientenrelevante Endpunkte**

Als Ergebnis eines Meetings von DMD-Experten, Patienten und Vertretern der Industrie (TREAT-NMD) können folgende Endpunkte je nach Fragestellung in Phase 2/3 Studien in Betracht gezogen werden (geeigneter Altersbereich der Patienten): „Griffiths Locomotor scale“ (0-8 Jahre), Bayley III Grobmotorik (1-42 Monate), „North Star Ambulatory Assessment“ (NSAA; 3.5 Jahre bis zum Verlust der Gehfähigkeit), „timed function tests“ (TFTs; 4 Jahre bis zum Verlust der Gehfähigkeit), „6-Minute Walk Test“ (6MWT; 5 Jahre bis zum Verlust der Gehfähigkeit), Muskelkraft anhand des manuellen Muskelkrafttests (4 Jahre bis Grad 2-), Muskelstärke quantitativ (untere Extremitäten: 5-12 Jahre, obere Extremitäten: 5 bis ≥20 Jahre), pulmonale Funktionstests (7 bis ≥20 Jahre), „Performance of upper limb“ (7 bis ≥20 Jahre), „Pediatric Outcomes Data Collection Instrument“ (PODCI, 3-21 Jahre), „PROM“ (7 bis ≥20 Jahre) und Muskelfunktion („myotools“, ≥5 Jahre). Lediglich das NSAA, die TFTs und der 6MWT erfüllten alle der Bewertung der Validität der Tests zugrundeliegenden Kriterien. Der 6MWT wurde als primärer Endpunkt bei DMD vorgeschlagen (24). Nach der Richtlinie der EMA sollten komplementäre klinisch relevante Endpunkte zur Muskelfunktion (z.B. „Motor Function Measure Scale“, NSAA, „Vignos lower limb score“, „Brooke upper limb score“, GSGC „[gait, stairs, Gowers, chair] assessment“, TFTs), motorische Funktionstests, Muskelstärke, Behinderung, Aktivitäten des täglichen Lebens

(z.B. „Functional Independence Measure“, Aktometer) oder der kardiopulmonalen Funktion begleitend erhoben werden, wenn der 6MWT als primärer Endpunkt gewählt wird. Zur Erfassung der Lebensqualität stehen das PedsQL („Pediatric quality of life inventory“) und bestimmte Domänen des PODCI („Pediatric Outcomes Data Collection Instrument“) bei gehfähigen DMD-Patienten zur Verfügung (2).

Die Bewertung des Ausmaßes des Zusatznutzens von Translarna basiert auf den patientenrelevanten Endpunkten und den relevanten Zulassungsstudien (8,34,35). Die vorliegende Bewertung erfolgte für die Intervention Translarna im Vergleich zu Placebo anhand der patientenrelevanten Endpunkte Veränderung der Gehstrecke im 6MWT (primärer Endpunkt), Zeit bis zu einer 10%igen Verschlechterung der Gehstrecke im 6MWT, TFTs und unerwünschter Ereignisse (UEs). In der Phase 3-Studie wurden zusätzlich die patientenrelevanten Endpunkte NSAA, Aktivitäten des täglichen Lebens mittels eines speziell entwickelten Tests und Gesundheitsbezogene Lebensqualität gemessen mittels PODCI-Domänen erhoben. Die Endpunkte sind gemäß Arzneimittel-Nutzenbewertungsverordnung (AM-NutzenV) und IQWiG Methoden – Version 4.2 den Zielgrößenkategorien Morbidität, Gesundheitsbezogene Lebensqualität und Nebenwirkungen zuzuordnen (7,36). Der patientenrelevante Endpunkt Mortalität konnte innerhalb der Dauer der eingeschlossenen Studien (48 Wochen) nicht evaluiert werden.

## Studientypen

Aufgrund der niedrigeren Ergebnissicherheit werden keine offenen, nicht kontrollierten Studien für die Bewertung des Ausmaßes des Zusatznutzens herangezogen. Da randomisierte, kontrollierte Studien den höchsten Grad der evidenzbasierten Medizin repräsentieren, werden in der folgenden Untersuchung nur randomisierte kontrollierte Studien berücksichtigt. Es liegen eine randomisierte, doppelblinde, Placebo-kontrollierte Parallelgruppenstudie der Phase 2b und eine konfirmatorische Phase 3-Studie vergleichbaren Designs jeweils über ein Jahr für Translarna vor, deren Ergebnisse in Meta-Analyse zusammengefasst wurden. Eine Meta-Analyse von zwei oder mehreren homogenen Studienergebnissen mit nicht notwendigerweise signifikanten Resultaten für die Einzelstudien kann bei Vorliegen eines statistisch signifikanten Effekts und hoher qualitativer Ergebnissicherheit einen Beleg für einen Zusatznutzen liefern (7,37). Die Ergebnisse der Meta-Analysen der Studien 007 und 020 sind nach der in der Verfo des G-BA dargestellten Evidenzhierarchie der Evidenzstufe 1a zuzuordnen (6).

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 (17). 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 (17). Entsprechend basiert die vorliegende Nutzenbewertung auf der der Zulassung zugrunde liegenden randomisierten, Placebo-kontrollierten Phase 3-Studie PTC124-GD-020-DMD (Studie 020) und der Phase 2b-Studie PTC124-GD-007-DMD (Studie 007; siehe frühere Nutzenbewertung; (3)). Aus diesen Studien leiten sich die zugelassene Indikation sowie die Anwendungsbedingungen von Ataluren ab. Eine detaillierte Betrachtung der Patientenpopulationen, der Intervention, der Endpunkte, und des Studiendesigns wird in Kapitel 4.4.2 durchgeführt. Die Einschlusskriterien der Studie 020 finden sich in Anhang 4-E.

### **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 (17). 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 nicht erforderlich (17).

#### **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 (17). 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 (6). Eine Bewertung der Aussagekraft der Nachweise ist für Arzneimittel zur Behandlung seltener Leiden daher nicht erforderlich (17). Dennoch wird zur Bewertung der dieser Nutzenbewertung zugrunde liegenden kontrollierten Studie (020) 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 (siehe vorheriges Dossier zur Nutzenbewertung für Studie 007). In Anhang 4-F werden lediglich die in der Meta-Analyse ausgewerteten Endpunkte bewertet, auf die sich die Ableitung des Ausmaßes des Zusatznutzens von Ataluren hauptsächlich stützt.

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 (6).

## 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 (17). Das Design und die Methodik der eingeschlossenen Zulassungsstudie PTC124-GD-020-DMD wurden in Kapitel 4.3.1.2, Anhang 4-E und Anhang 4-F entsprechend der Vorgaben für randomisierte Studien dargestellt. Insoweit erfüllt die Darstellung der zulassungsrelevanten, randomisierten kontrollierten Studie PTC124-GD-020-DMD mindestens die Anforderungen des CONSORT-Statements. Zu Design und die Methodik der

<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.

Studie PTC124-GD-007-DMD siehe Nutzenbewertungsverfahren zum Wirkstoff Ataluren, Beschlussfassung: 21.05.2015.

#### 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.*

Die Charakteristika der in die zulassungsrelevante Studie 020 eingeschlossenen Patienten sind in Kapitel 4.3.1.2 beschrieben. Angaben zur Operationalisierung, Validität und Patientenrelevanz der dieser Bewertung zugrundeliegenden patientenrelevanten Endpunkte sind in Kapitel 4.3.1.2 und 4.4.2 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

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



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 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.*

Es wurden Meta-Analysen auf der Basis von zwei eingeschlossenen Studien durchgeführt. Dazu wurden die prospektive, drei-armige, randomisierte, kontrollierte, doppelt verblindete, multizentrische Phase 2b Studie PTC124-GD-007-DMD sowie die prospektive, zwei-armige, randomisierte, kontrollierte, doppelt verblindete, multizentrische Phase 3 Studie PTC124-GD-020-DMD herangezogen. Aus der drei-armigen Studie PTC124-GD-007-DMD wurde in den Auswertungen des vorliegenden Dossiers jedoch nur die niedrigere, zugelassene Dosierung von 40 mg/kg/Tag berücksichtigt.

Zu den primären Analysepopulationen und Analysemethoden der in den Meta-Analysen eingesetzten Phase 2b Studie siehe Kapitel 4.3.1.2.1. Für die Effektgrößen des 6MWT und der TFTs wurden die aggregierten Daten (LSM Differenzen zwischen den beiden Behandlungsarmen) aus den Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD eingesetzt. Für den Endpunkt „Unerwünschte Ereignisse“ wurden relativen Risiken aus den Einzelstudien in den Meta-Analysen eingesetzt.

---

<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.

Auf Grund der geringen Patientenzahlen in den jeweiligen Studien ist es methodisch sinnvoll, die Ergebnisse in einer Meta-Analyse zusammenzufassen, um die Präzision der Resultate und die Ergebnissicherheit zu erhöhen. Dabei ist es formell unerheblich, ob die beiden Einzelstudien signifikante Ergebnisse liefern. Vielmehr bestehen die Anforderungen darin, dass die Zusammenfassung der Ergebnisse der Einzelstudien im Sinne einer Meta-Analyse zu einem statistisch signifikanten Unterschied führt (37).

Im vorliegenden Dossier wurden alle Meta-Analysen auf Basis von Modellen mit zufälligen Effekten durchgeführt und als Forest-Plots dargestellt (s. Kapitel 4.3.1.3.1). Der Grad der Heterogenität zwischen den Studienergebnissen und ein zusammenfassender Effektschätzer mit dazugehörigem Konfidenzintervall wurden berechnet. Die Untersuchung auf statistische Heterogenität der Effekte zwischen den Studien erfolgte mittels Cochran-Q-Statistik und dem  $I^2$ -Maß. Lag der p-Wert des Tests auf Heterogenität unterhalb des Signifikanzniveaus von 0.05, so lag eine signifikante Heterogenität vor. Die Meta-Analysen der patientenrelevanten Endpunkte wurden auf Basis aggregierter Daten (extrahiert aus den Outputs in den jeweiligen CSRs) der Einzelstudien berechnet.

Zur Durchführung der Meta-Analyse wurde die Software Review Manager 5.3 (Version 5.3.5; Build date 30.10.14, The Cochrane Collaboration) verwendet.

#### 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 Studie PTC124-GD-020-DMD *a priori* spezifiziert (34):

- MMRM-Analyse („mixed model repeated measures“ Analyse)
- Effekt der Zuteilung von Zwillingen/Drillingsen zu derselben Behandlungsgruppe
- Randomisierungs-basierte ANCOVA-Analyse (optional)

In den Meta-Analysen der beiden eingeschlossenen Studien wurde für die Studie PTC124-GD-007-DMD die kITT Population und die prädefinierte MMRM Analyse eingesetzt (9) (s. Kapitel 4.3.1.2.1).

In Sensitivitätsanalysen von Meta-Analysen wurden zusätzlich, wenn angegeben, folgende Analysepopulationen/Analysemethoden für die Studie PTC124-GD-007-DMD eingeschlossen:

- ITT Population/prädefiniertes MMRM Model
- kITT Population/augmentiertes MMRM Model (*post-hoc*) (38)
- ITT Population/ANCOVA Model (38)

#### 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 Effektmodifikation 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-Analysen und Subgruppenmerkmale waren im SAP der PTC124-GD-020-DMD *a priori* spezifiziert (34):

- BL-6MWD-Gruppen (<300 m,  $\geq$ 300 m bis <400 m,  $\geq$ 400 m; 6MWD-Stratifizierung nicht in das ANCOVA-Model eingeschlossen)
- Stratifizierungsfaktoren:
  - o BL-6MWD Stratifizierungsfaktoren ( $\geq$ 350 m vs. <350 m)
  - o Dauer der Kortikosteroid-Vorbehandlung zu BL (etwa  $\geq$ 6 bis <12 Monate vs.  $\geq$ 12 Monate)
  - o Altersgruppen zu BL (<9 Jahre vs.  $\geq$ 9 Jahre)

Folgende *post-hoc* Subgruppen-Analysen wurden im CSR berichtet (explorativ):

- BL-6MWD  $\geq$ 300 bis <350 m,  $\geq$ 350 bis <400 m,  $\geq$ 400 bis <450 m,  $\geq$ 450 m; deskriptiv)
- BL-6MWD (<250 m,  $\geq$ 250 bis <400 m)
- BL-6MWD (<200 m,  $\geq$ 200 bis <400 m)
- BL-6MWD  $\geq$ 300 bis <450 m,  $\geq$ 450 m)

Ein aus Interaktionstests resultierender p-Wert zwischen  $>0.05$  und  $<0.2$  wurde als Hinweis auf unterschiedliche Effekte zwischen den Subgruppen gewertet. Ein p-Wert von  $<0.05$  wurde als Beleg für unterschiedliche Effekte eingestuft. Im Falle von Belegen wurden die Einzelergebnisse separat für alle Subgruppen dargestellt und in die Bewertung des Zusatznutzens eingeschlossen. Die Subgruppen-Analysen nach BL-6MWD-Gruppen (<300 m,  $\geq$ 300 m bis <400 m,  $\geq$ 400 m) zeigten für die meisten bewertungsrelevanten Endpunkte Belege für unterschiedliche Effekte zwischen den Subpopulationen. Um die Konsistenz der Ergebnisse über verschiedene Testsysteme zu zeigen, werden diese Ergebnisse vollständig berichtet, auch wenn für einzelne Endpunkte keine Belege auf unterschiedliche Effekte zwischen den Subgruppen vorlagen. Aus Gründen der Vollständigkeit werden auch die Daten der prospektiv definierten 6MWD-Gruppen von <350 vs.  $\geq$ 350 m nahezu komplett berichtet. Post-hoc Subgruppen-Analysen werden nur berichtet, wenn sich Belege für zwischen den Gruppen unterschiedliche Effekte ergeben ( $p < 0.05$  im Interaktionstest).

In keiner der untersuchten Subgruppen traten *qualitative* Interaktionen auf. Die Subgruppenanalysen in diesem Dossier konzentrieren sich daher auf *quantitative* Interaktionen zwischen präspezifizierten Gruppen.

Die Auswahl der Subgruppen, die Entscheidung ob ein tatsächlicher Subgruppeneffekt mit hoher Wahrscheinlichkeit korrekt identifiziert wurde, sowie die Interpretation der Subgruppeneffekte wurde streng wissenschaftlich, primär nach den Kriterien von Sun et al. 2010 (39), und entsprechend der EMA „Guideline on the investigation of subgroups in confirmatory clinical trials“ (Draft) durchgeführt (40). Sun et al. 2010 schlagen 11 Kriterien vor, nach denen die Verlässlichkeit von Aussagen zu Subgruppeneffekten geprüft werden sollen; diese Kriterien sind in Tabelle 4-4 zusammengefasst.

Tabelle 4-4: Kriterien zur Prüfung der Aussagesicherheit zu Subgruppeneffekten (39)

Studienaspekte	Kriterien
Design	1. Ist die Subgruppenvariable ein Charakteristikum, das zu Baseline erhoben wurde (oder aber nach Randomisierung)?
	2. Wird der Subgruppeneffekt durch Vergleiche innerhalb von Studien beschrieben (oder durch Vergleiche zwischen Studien)?
	3. War die Hypothese <i>a priori</i> spezifiziert?
	4. War die Richtung des Subgruppeneffekts <i>a priori</i> definiert?
	5. Entspricht der Test auf den spezifischen Subgruppeneffekt einer von wenigen getesteten Hypothesen
Analyse	6. Ist der Interaktionstest mit einer geringen Wahrscheinlichkeit vereinbar, dass der Subgruppeneffekt zufällig auftritt?
	7. Ist der signifikante Subgruppeneffekt unabhängig (unterschiedliche Subgruppeneffekte sind voneinander statistisch unabhängig)?
Kontext	8. Ist der Subgruppeneffekt groß?
	9. Ist der Subgruppeneffekt über unabhängige Studien hinweg konsistent?
	10. Ist der Subgruppeneffekt über ähnliche Outcome-Parameter konsistent?
	11. Existiert indirekte Evidenz, welche die statistische Interaktion erklärt (biologische Rationale)?

Eine Überprüfung der präspezifizierten Subgruppenvariablen nach den in Tabelle 4-4 aufgeführten Kriterien ergab, dass alleinig das Subgruppenpaar BL-6MWD ( $\geq 300$  m bis  $< 400$  m versus  $< 300$  m plus  $\geq 400$  m) sämtliche Kriterien zur Prüfung der Aussagesicherheit von Subgruppeneffekten erfüllte. Insbesondere sind die Kriterien 9 bis 11 (Tabelle 4-4) nur durch das Subgruppenpaar BL-6MWD ( $\geq 300$  m bis  $< 400$  m versus  $< 300$  m plus  $\geq 400$  m) erfüllt. Die biologische Rationale für die statistische Interaktion (Kriterium 11) ist in Kapitel 4.4.2.4 detailliert beschrieben.

Wie dort dargelegt, stellt die Subpopulation der Patienten mit einer Gehstrecke von  $\geq 300$  m bis  $< 400$  m ein Kollektiv dar, welches sich nach neuesten Erkenntnissen (u.a. der Ataluren-Studien) über eine Behandlungsdauer von 48 Wochen in Bezug auf die gewählten Endpunktparameter klinisch kontinuierlich verschlechtert; die Krankheitsprogression ist daher in dieser Population im 6MWT auch messbar. Bei Patienten mit einer BL-Gehstrecke von  $\geq 400$  m ist dies nicht der Fall, da sie in Bezug auf die Endpunktparameter über die Behandlungsdauer von einem Jahr noch stabil bleiben. Die Assay-Sensitivität für den 6MWT ist somit in der Subpopulation  $\geq 300$  m bis  $< 400$  m gewährleistet, während sie in Patienten mit Gehstrecken von  $\geq 400$  m eingeschränkt ist.

Bei Patienten mit Baseline 6MWD von  $\geq 300$  m bis  $< 400$  m ist die Erkrankung auch noch nicht so weit fortgeschritten wie bei Patienten in der späten Krankheitsphase mit Gehstrecken  $< 300$  m, deren gesundes Muskelgewebe schon soweit reduziert ist, dass sich ihr Zustand auch unter einer Dystrophin-restaurierenden Therapie verschlechtert. Die Assay-Sensitivität des 6MWT ist daher auch für Patienten mit Gehstrecken  $< 300$  m deutlich eingeschränkt.

Die Subpopulation der Patienten mit einer Baseline-Gehstrecke von  $\geq 300$  m bis  $< 400$  m stellt somit diejenige Patientenpopulation dar, in welcher der 6-Minuten-Gehtest (6MWT) die Therapieeffekte von Ataluren am sensitivsten und mit hoher Reproduzierbarkeit erfassen kann. Die Ergebnisse der Subgruppenanalysen der Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD (siehe Kapitel 4.4.2) stützen dies eindrucksvoll.

Die fehlende Assay-Sensitivität für Patienten im frühen Krankheitsverlauf (BL-Gehstrecken von  $\geq 400$  m) in Studien mit zeitlich eingeschränkter Beobachtungsdauer (48 Wochen) bedeutet natürlich nicht, dass der eingesetzte Wirkstoff nicht wirksam wäre. Die Wirksamkeit einer Dystrophin-restaurierenden Therapie ist für diese Patienten erst nach längeren Beobachtungszeiten zu erwarten. Und zwar erst dann, wenn bei unbehandelten Patienten die Schädigung des Muskelgewebes soweit fortgeschritten wäre, dass sie mittels der Erhebungsinstrumente (6MWT) messbar wird.

Zur Bestimmung der Effektgrößen für die Wirksamkeitsparameter in den Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD ist es daher erforderlich, diejenige Studienpopulation heranzuziehen, in der die Assay-Sensitivität gewährleistet ist (Patienten mit einer Baseline-Gehstrecke von  $\geq 300$  m bis  $< 400$  m). Voraussetzung hierfür ist jedoch auch ein Beleg einer statistischen Interaktion zwischen Therapieeffekten, die in Subgruppen mit nachweislich unterschiedlicher Assay-Sensitivität erhoben wurden (s. Kapitel 4.3.1.3.2).

Da der Effektmodifikator (BL-6MWD) in der bewertungsrelevanten Subgruppenanalyse aus drei Kategorien besteht ( $\geq 300$  m bis  $< 400$  m;  $< 300$  m;  $\geq 400$  m) wurden paarweise Interaktionstest durchgeführt, um nicht-transitive Relationen zu adressieren (41)

#### 4.2.5.6 Indirekte Vergleiche

Zurzeit sind international 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>.

*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.*

---

<sup>7</sup> Glenny AM, Altman DG, Song F, Sakarovitch 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.

- *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 4-5: 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 <sup>1</sup>	ja (Phase 2b)	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)
PTC124-GD-020-DMD	ja (Phase 3)	ja	abgeschlossen	48 Wochen 26. März 2013 bis 20. August 2015	Placebo Ataluren 10/10/20 mg/kg KG (40 mg/kg KG/Tag)

<sup>1</sup> Die Studie PTC124-GD-007-DMD wurde in einem separaten Dossier zur Nutzenbewertung dargestellt (Nutzenbewertungsverfahren zum Wirkstoff Ataluren, Beschlussfassung: 21.05.2015).

*Geben Sie an, welchen Stand die Information in Tabelle 4-5 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.*

Nicht zutreffend, da es sich bei Ataluren um ein Arzneimittel zur Behandlung seltener Leiden handelt.

*Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-6 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.*

Tabelle 4-6: 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.*

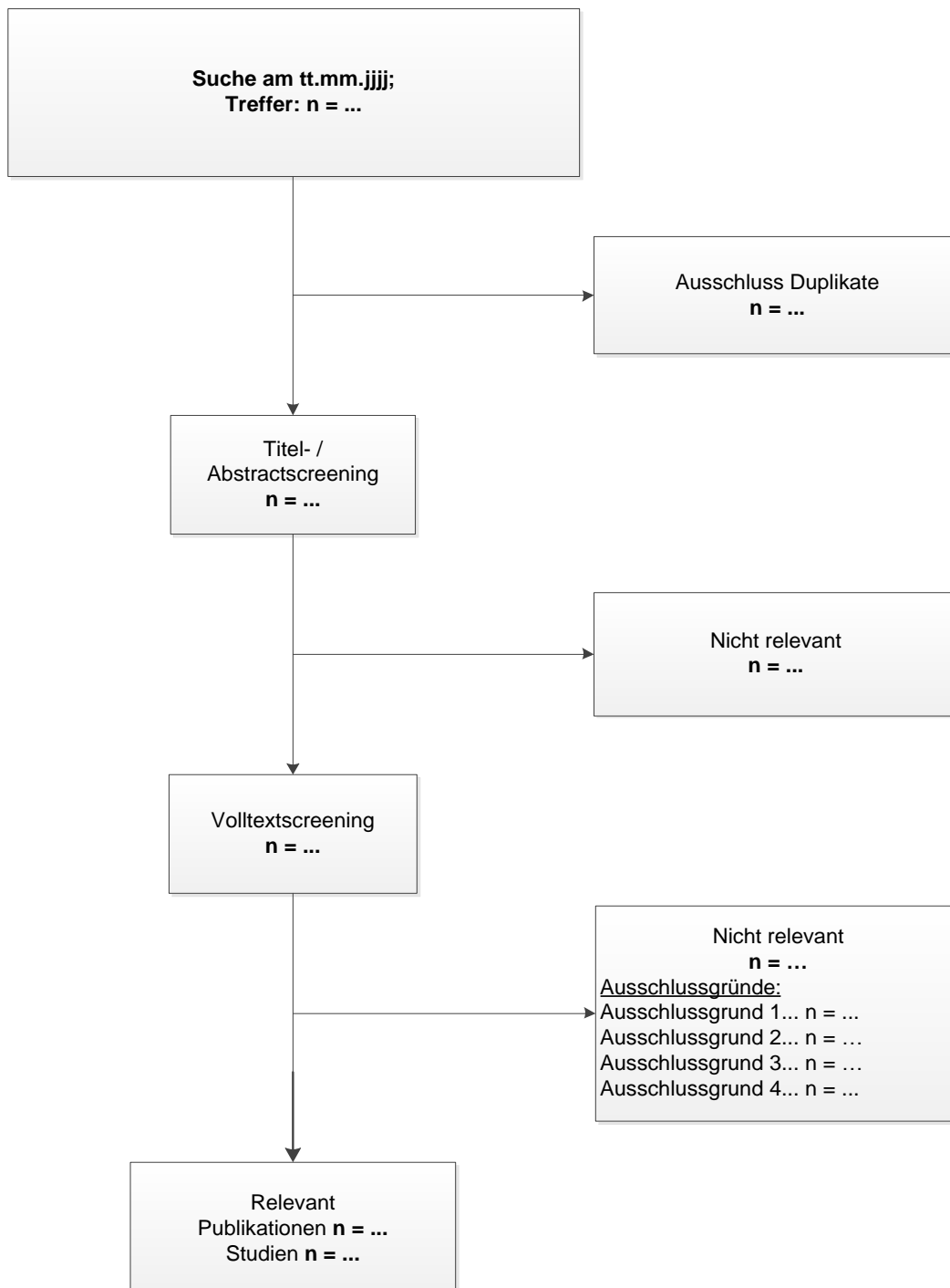


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

Da es sich bei Translarna® um ein Arzneimittel zur Behandlung seltener Leiden handelt, entfällt die Verpflichtung einer bibliographischen 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 4-5) 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 4-7: 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)
<Studie 1>	clinicaltrials.gov [6, 7]	ja	nein	abgeschlossen

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 4-7 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 Dossiers maßgeblichen Zeitpunkt liegen.*

Da es sich bei Translarna<sup>®</sup> 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 4-8: 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>ggf. Zwischenüberschrift zur Strukturierung des Studienpools</b>						
<b>placebokontrolliert</b>						
PTC124 -GD- 007- DMD	ja	nein	nein	ja (9)	ja Siehe Nutzenbewertungsverfahren zum Wirkstoff Ataluren, Beschlussfassung 21.05.2015	ja Siehe Nutzenbewertungsverfahren zum Wirkstoff Ataluren, Beschlussfassung 21.05.2015
PTC124 -GD- 020- DMD	ja	ja	nein	ja (8)	ja clinicaltrials.gov: NCT01826487(42) EudraCT: 2012-004527-20 (43)	nein
<b>aktivkontrolliert, zweckmäßige Vergleichstherapie(n)</b>						
	Nicht zutreffend					
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregister-einträge) 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>						

### 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 Phase 3 Studie PTC124-GD-020-DMD und die Charakteristika der Interventionen sind in Tabelle 4-10, Tabelle 4-11 und Anhang 4-E dargestellt. In Tabelle 4-12 sind die Charakteristika der Studienpopulation aufgeführt. Da im Folgenden Analysen von Subpopulationen bei Patienten mit einer BL-6MWD von  $\geq 300$  bis  $< 400$  m eine besondere Rolle spielen, sind die demographischen Daten dieses prä-definierten Kollektivs ebenfalls in der Tabelle 4-12 dargestellt. Wenn nicht anders angegeben beziehen sich die nachfolgend dargestellten Daten auf den Studienbericht der Studie PTC124-GD-020-DMD (8).

Die in Meta-Analysen für das vorliegende Dossier eingesetzten Analysepopulationen und Analysemethoden der Phase 2b Studie PTC124-GD-007-DMD sind in Tabelle 4-9 dargestellt (siehe auch Nutzenbewertung von Ataluren; Beschluss 21.05.2015). Wenn nicht anders angegeben wurden als primäre Analysepopulation die kITT Population und als primäre Analysemethode das *a priori* definierte MMRM Model in den vorliegenden Meta-Analysen eingesetzt.

Tabelle 4-9: Analysepopulationen und Analysemethoden der Phase 2b Studie PTC124-GD-007-DMD in Meta-Analysen

	Analysepopulation (9)	Analysemethode (9)
PrimäreAnalyse	kITT	Präspezifiziertes MMRM Model
Sensitivitätsanalysen	ITT	Präspezifiziertes MMRM Model
	kITT	Augmentiertes MMRM Model (post-hoc)
	ITT	ANCOVA

In der Phase 2b Studie PTC124-GD-007-DMD wurde für die kITT Population bei je einem Patienten in der Placebo- bzw. in der Ataluren-20/20/40mg-Gruppe der Baseline-Wert für den

6MWT durch den Screening-Wert ersetzt. Grund hierfür waren akute Verletzungen (Knie bzw. Sprunggelenk) wenige Tage vor der Erhebung der Baseline-Werte der unteren Extremität, die zu einem falsch-niedrigen, d. h. nicht repräsentativen Ergebnis zu Baseline führten (309 statt 395 m bei der Screening-Visite bzw. 125 statt 303 m bei der Screening-Visite). Die Werte wurden zunächst als gültig eingestuft und nachträglich protokollgemäß durch die Screening-Werte ersetzt, was zu einer *korrigierten ITT-Population* (kITT) führte. Dementsprechend wird die kITT-Population in der vorliegenden Bewertung als klinisch und wissenschaftlich begründete primäre Analysepopulation eingestuft. Diese Einschätzung wurde vom G-BA in den tragenden Gründen zum Beschluss (Ataluren) vom 21.05.2015 geteilt und als plausibel bewertet.



Tabelle 4-10: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
PCT124-GD-020-DMD	Multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte, 2-armige Parallelgruppen-Studie der Phase 3, GCP	Gefähige männliche Patienten im Alter zwischen $\geq 7$ und $\leq 16$ Jahren mit genetisch gesicherter nmDMD, die über mindestens sechs Monate systemische Kortikosteroide eingenommen hatten; valide Screening-6MWD von $\geq 150$ m und $\leq 80\%$ des nach Alter und Größe vorausgesagten Wertes	Placebo (n=115) Ataluren 10, 10, und 20 mg/kg KG (n=115) jeweils dreimal täglich p.o.	2-wöchiges Screening 48-wöchige doppelblinde Behandlungsphase 6-wöchiges ( $\pm 7$ d) Follow-up nach der letzten Dosis (Patienten, die nicht an einer offenen Folgestudie teilnahmen)	International: 54 Zentren in 18 Ländern (USA, Spanien, Frankreich, Türkei, Chile, Deutschland, UK, Italien, Australien, Polen, Kanada, Schweden, Belgien, Brasilien, Südkorea, Schweiz, Tschechische Republik, Israel) 26. März 2013 bis 20. August 2015	<u>Primärer Wirksamkeitsendpunkt:</u> Änderung der 6MWD <u>Sekundäre Wirksamkeitsendpunkte:</u> Zeit bis zu dauerhafter 10%iger Verschlechterung der 6MWD Änderung der TFTs (Zeit, um 10 m zu gehen/laufen und 4 Stufen hinauf- oder hinabzusteigen) <u>Explorative Wirksamkeitsendpunkte:</u> Änderung des NSAA (körperliche Funktion) inklusive Zeit, um aus dem Liegen aufzustehen Alltagsaktivitäten Änderung des PODCI (Transfer/Basismobilität und Sport/körperliche Funktion)

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

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
						<u>Arzneimittelsicherheit</u> PK siehe Tabelle 4-85 Compliance mit der Behandlung

Tabelle 4-11: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel (44)

Studie	Placebo	Ataluren	Komedikation
PTC124-GD-020-DMD	Morgens (etwa 7:00 h), Mittags (etwa 13:00 h) und Abends (etwa 19:00 h) p.o. jeweils innerhalb von etwa 30 min nach den Mahlzeiten, d.h. in Abständen von etwa 6, 6 bzw. 12 h (±1 h)	Ataluren 10, 10 und 20 mg/kg KG, analog Placebo  Gegebenenfalls Anpassung der Dosis an das KG in Woche 24 für den Rest der Studie	Nur Patienten, die folgenden Kriterien entsprachen durften in die Studie eingeschlossen werden:  Mindestens sechs Monate vor Beginn der Behandlung systemische Kortikosteroid-Therapie (Prednison, Prednisolon oder Deflacort); keine signifikante Änderung der Dosierung oder des Dosisregimes über mindestens drei Monate vor Beginn der Behandlung (ausgenommen Änderungen aufgrund des KGs); begründete Erwartung, dass sich die Dosierung oder das Dosisregime während der Studie nicht signifikant ändern.  Während der Studie sollte das Kortikosteroid-Regime möglichst stabil bleiben. Anpassungen der Dosis aufgrund von Änderungen des KGs waren zulässig. Das Protokoll enthielt spezifische Instruktionen bezüglich der Kortikosteroid-Therapie während der Studie.  Siehe Fußnote 1 für weitere Informationen.

<sup>1</sup> Bei Patienten, die vor Studieneinschluss kardiale Medikationen zur Prophylaxe/Therapie einer chronischen Herzinsuffizienz einnahmen, sollte die Dosis über mindestens drei Monate vor der Randomisierung stabil gewesen sein und diese sollte während der Studie möglichst stabil bleiben. Dosisanpassungen waren zulässig, um Symptome zu vermeiden (z.B. ACE-Inhibitor-assoziiertes Husten), die Medikationen sollten aber möglichst nicht abgesetzt werden. Wenn möglich sollten kardiale Medikationen nicht neu eingeleitet werden und falls dies unvermeidbar war, sollte möglichst auf ACE-Inhibitoren und ARBs verzichtet werden. ACE-Inhibitoren und ARBs können die Plasminogen-Werte erhöhen, was einen möglichen Ataluren-assoziierten Aldosteronmangel verschleiern könnte. Das Protokoll enthielt weitere Hinweise zur gleichzeitigen Gabe potentiell nephrotoxischer Antibiotika (insbesondere Aminoglycoside oder Vancomycin), der Wichtigkeit einer adäquaten Hydratation sowie Warnhinweise hinsichtlich weiterer möglicher Interaktionen (z.B. CYP2C8 und CYP2C9).

Tabelle 4-12: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel („as-treated“ Population)

<b>Studie Gruppe</b>	<b>N</b>	<b>Alter (Jahre) M (±SD) [Bereich]</b>	<b>Alter &lt;9 / ≥9 Jahre n (%)</b>	<b>Kaukasier n (%)</b>	<b>Körper- größe (cm) M (±SD) [Bereich]</b>	<b>Körper- gewicht (kg) M (±SD) [Bereich]</b>	<b>BMI (kg/m<sup>2</sup>) M (±SD) [Bereich]</b>	<b>Alter bei Diagnose- stellung M (±SD) [Bereich]</b>	<b>BL 6MWD (m) M (±SD) [Range]</b>
PTC124-GD-020-DMD									
Placebo	115	9.0 (1.65) [7, 14]	53 (46.1) / 62 (53.9)	86 (74.8%)	126.34 (10.420) [101.8, 151.0]	30.62 (10.437) [18.1, 68.0]	18.80 (4.099) [13.0, 36.0]	4.4 (2.47) (0, 11)	362.69 (81.424) [142.5, 526.0]
Ataluren	115	8.9 (1.79) [7, 14]	57 (49.6) / 58 (50.4)	89 (77.4%)	126.14 (10.733) [93.9, 163.5]	31.42 (10.762) [15.8, 63.0]	19.31 (4.355) [11.3, 36.2]	4.4 (2.08) (0, 9)	364.04 (73.342) [166.8, 511.0]

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

<b>Studie Gruppe</b>	<b>N</b>	<b>Alter (Jahre) M (±SD) [Bereich]</b>	<b>Alter &lt;9 / ≥9 Jahre n (%)</b>	<b>Kaukasier n (%)</b>	<b>Körper- größe (cm) M (±SD) [Bereich]</b>	<b>Körper- gewicht (kg) M (±SD) [Bereich]</b>	<b>BMI (kg/m<sup>2</sup>) M (±SD) [Bereich]</b>	<b>Alter bei Diagnose- stellung M (±SD) [Bereich]</b>	<b>BL 6MWD (m) M (±SD) [Range]</b>
PTC124-GD-020-DMD: Subgruppe BL-6MWD ≥300 bis <400 m									
Placebo	52	9.0 (1.49) [7, 12]	22 ( 42.3) / 30 ( 57.7)	38 (73.1)	126.73 (9.502) [108.0, 151.0]	30.71 (10.191) [18.1, 68.0]	18.78 (4.122) [13.0, 34.0]	4.5 (2.58) [0, 11]	354.52 (29.712) [300.7, 399.5]
Ataluren	47	8.9 (1.81) [7, 14]	23 ( 48.9) / 24 ( 51.1)	37 (78.7)	123.76 (10.750) [93.9, 147.8]	30.43 (11.174) [17.2, 63.0]	19.41 (4.844) [11.3, 36.2]	4.6 (2.10) [1, 9]	356.69 (26.672) [305.0, 398.0]

Studie Gruppe	N	Phänotyp 1 n (%)						Stop Codon-Typ n (%)			Lokalisation des Exons n (%)		Kortiko- steroid 6-<12 / ≥12 Monate n (%)
		A	B	C	D	E	F	UGA	UAG	UAA	1-39	40-80	
PTC124-GD-020-DMD													
Placebo	11 5	76 (66.1)	91 (79.1)	92 (80.0)	110 (95.7)	107 (93.0)	38 (33.0)	40 (34.8)	45 (39.1)	29 (25.2)	68 (59.1)	58 (50.4)	19 (16.5) / 96 (83.5)
Ataluren	11 5	83 (72.2)	83 (72.2)	91 (79.1)	105 (91.3)	111 (96.5)	36 (31.3)	57 (49.6)	31 (27.0)	26 (22.6)	58 (50.4)	56 (48.7)	19 (16.5) / 96 (83.5)
PTC124-GD-020-DMD: Subgruppe BL-6MWD ≥300 bis <400 m													
Placebo	52	34 (65.4)	43 (82.7)	38 (73.1)	49 (94.2)	49 (94.2)	15 (28.8)	18 (34.6)	20 (38.5)	14 (26.9)	28 (53.8)	24 (46.2)	7 (13.5) / 45 (86.5)
Ataluren	47	35 (74.5)	34 (72.3)	36 (76.6)	44 (93.6)	45 (95.7)	14 (29.8)	27 (57.4)	13 (27.7)	7 (14.9)	23 (48.9)	24 (51.1)	7 (14.9) / 40 (85.1)

<sup>1</sup> A: schwankender Gang, B: Gowers Manöver, C: Wadenhypertrophie, D: Genetisches Ergebnis, E: Erhöhte Kreatinkinase, F: Andere

Quelle (Studienbericht 020): Tabelle 9, Tabelle 14.1.3.2.1.2, Tabelle 14.1.3.1.3, Tabelle 14.1.3.2.1.3

*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.*

Die Studie PCT124-GD-007-DMD wurde in einem separaten Dossier zur Nutzenbewertung beschrieben und wird in diesem Dossier nicht erneut dargestellt. Auf Unterschiede zwischen den zulassungsrelevanten Studien 020 und 007 wird in Kapitel 4.4.2 eingegangen.

#### **Studie PCT124-GD-020-DMD (Studienbericht, -protokoll und SAP) (8,34,44)**

Die Wirksamkeit und Sicherheit von Ataluren bei Patienten mit nmDMD wurde in einer Phase 3, multizentrischen, randomisierten, doppelblinden, Placebo-kontrollierten, GCP-konformen Parallelgruppen-Studie überprüft. Die Studie bestand aus einer 2-wöchigen Screening-Phase, einer 48-wöchigen doppelblinden Behandlungsphase und einer 6-wöchigen Follow-up-Phase nach der letzten planmäßigen Dosis oder einem vorzeitigen Abbruch (Patienten, die nicht an der offenen Folgestudie teilnahmen).

**Primäres Ziel der Studie:** Wirksamkeit von Ataluren, die Krankheitsprogression gemessen als Verlust der Gehfähigkeit (Abnahme der 6MWD) aufzuhalten

**Studienpopulation:** Eingeschlossen wurden männliche Patienten im Alter zwischen  $\geq 7$  und  $\leq 16$  Jahren mit phänotypischem Nachweis einer Dystrophinopathie im Alter von sechs Jahren. Der Nachweis einer Non-Sense-Punktmutation im Dystrophin-Gen via Gensequenzierung durch ein zertifiziertes Labor war zum Einschluss erforderlich. Nur Patienten, die mit systemischen Kortikosteroiden vorbehandelt waren und deren Kortikosteroid-Therapie während der Studie erwartungsgemäß stabil bleiben würde (siehe Tabelle 4-11), konnten an der Studie teilnehmen. Nur Patienten mit einer validen Screening-6MWD von  $\geq 150$  m und  $\leq 80\%$  des entsprechend Alter und Größe vorhergesagten Wertes, welche während der BL-Visite reproduziert werden konnte ( $\leq 20\%$  Änderung verglichen mit dem Screening-Wert), konnten an der Studie teilnehmen.

Die Wahl der Haupteinschlusskriterien basierte im Wesentlichen auf einer Auswertung des Placeboarms der vorausgegangenen Ataluren-Studie 007. Bei Patienten bis zum Alter von 7 Jahren besserte sich die 6MWD tendenziell oder blieb stabil (15). Eine italienische Beobachtungsstudie bei Jungen mit DMD zeigte ebenfalls, dass die Gehfähigkeit ab einem Alter von 7 Jahren progressive abnimmt, während bei Jungen unter 7 Jahren die Gehfähigkeit innerhalb eines Jahres durch Wachstum kompensatorisch zunimmt. Die prognostische Relevanz eines kritischen Alters von 7 Jahren blieb über 36 Monate erhalten (45). Jungen mit einer 6MWD über 350 m bleiben über einen Zeitraum von 48 Wochen tendenziell stabil, während der Gehfähigkeit bei Jungen mit einer kürzeren initialen Gehstrecke abnimmt (46). Das CHMP bewertete die Auswahlkriterien der „ambulatory decline“ Phase-Gruppe, die den Einschlusskriterien der Phase 3-Studie im Wesentlichen entsprachen, als klinisch begründet

(1,46). Da Kortikosteroide Bestandteil des Behandlungsstandards sind, wurden nur mit Kortikosteroiden vorbehandelte Patienten in die Studie aufgenommen.

**Intervention:** In der vorausgegangenen Phase 2b Dosisfindungs-Studie (PTC124-GD-007-DMD) hatte sich die höhere Ataluren-Dosis von 20, 20, 40 mg/kg KG als unwirksam erwiesen, während eine Dosis von 10, 10 und 20 mg/kg KG wirksam war. Beide Dosen waren gut verträglich. Für Ataluren, wie auch andere Wirkstoffe mit ähnlichem Wirkmechanismus ist eine glockenförmige Konzentrations-Wirkungs-Kurve bekannt. Daher wurde die höhere Ataluren-Dosis in der Phase 3-Studie nicht geprüft. Die Patienten nahmen Ataluren (10, 10 und 20 mg/kg KG) oder ein Placebo jeweils dreimal täglich möglichst 30 min nach den Mahlzeiten oral ein (entsprechend Einnahme-Intervalle von 6, 6, bzw. 12 h [ $\pm 1$  h]). Die Prüfpräparate (Pulver) sollten vermischt mit etwas Flüssigkeit oder halbfester Nahrung eingenommen werden. Die Dosierung der Studienmedikationen wurde zentral mittels „Interactive Voice Response/Interactive Web Response“ (IVR/IWR) auf der Basis der Angaben der Patienten kalkuliert und vorgegeben.

Die doppelblinde Behandlungsphase betrug 48 Wochen und Visiten fanden während der Behandlungsphase alle 8 Wochen statt. Eine Behandlungsdauer von mehr als einem Jahr wurde in einer Placebo-kontrollierten Studie unter Berücksichtigung der Schwere der Erkrankung als ethisch nicht vertretbar angesehen.

Begleitende Kortikosteroide sowie kardiale Begleitmedikationen sollten soweit wie möglich standardisiert verabreicht werden (siehe Tabelle 4-11), um das Risiko möglicher Störeffekte zu minimieren. Das Studienprotokoll enthielt spezifische Instruktionen hinsichtlich der Kortikosteroid-Therapie (44). Es sollten, wenn möglich Gesamtdosen von 0.75 mg/kg Prednison oder 0.9 mg/kg Deflazacort nicht überschritten werden. Patienten mit einem Körpergewicht von 40 kg oder höher sollten maximal 30 mg Prednison oder 36 mg Delazacort erhalten, wenn die Patienten nicht bereits höhere Dosen zu Studienbeginn erhalten hatten. Dosiserhöhungen von mehr als 5 mg Prednison oder 6 mg Deflazacort sollten vermieden werden. Die Kortikosteroide sollten nur abgesetzt werden, wenn klinisch notwendig und wenn möglich nach einer Dosisreduktion bei Auftreten von UEs. Im Falle von Nebenwirkungen konnte die Gesamtdosis um  $\leq 5$  mg Prednison ( $\leq 6$  mg Deflacort) reduziert werden.

**Vergleichstherapie:** Ataluren ist das erste zur Behandlung der DMD in Europa zugelassene Arzneimittel („first-in-class“). 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.



***Primäre und sekundäre Endpunkte:***

Primärer Wirksamkeits-Endpunkt: Änderung der 6MWD von BL bis Woche 48

Sekundäre Wirksamkeits-Endpunkte:

- Zeit bis zu einer 10%igen Verschlechterung der 6MWD
- Änderung der proximalen Muskelfunktion gemessen in TFTs (Zeit, die benötigt wird, um vier Stufen hinauf- oder hinabzusteigen, 10 m zu gehen/laufen)

Explorative Wirksamkeits-Endpunkte:

- Änderung der körperlichen Funktion ermittelt anhand des NSAA-Gesamtscores
- Änderung der Aktivitäten des täglichen Lebens und des Krankheitsstatus, welche mittels eines standardisierten Fragebogens durch das Klinikpersonal erhoben wurden
- Änderung der PODCI-Scores (Transfer/Mobilität und Sport/körperliche Funktion)
- Ataluren-Plasmakonzentrationen (unmittelbar vor der morgendlichen Dosis)

Compliance der Patienten (Prozentsatz der tatsächlich eingenommenen Studienmedikation in Relation zur Menge, die hätte eingenommen werden müssen)

Die Sicherheitsendpunkte schlossen UEs, Laborabnormalitäten, Vitalzeichen, EKG-Änderungen, renale Ultraschall-Untersuchungen und körperliche Untersuchungen ein.

***Fallzahlschätzung:*** Der Studie lag die Hypothese zugrunde, dass die mittlere Änderung der 6MWD von BL bis Woche 48 in der Ataluren-Gruppe 30 m länger sein würde als in der Placebo-Gruppe. In der Phase 2b-Studie 007 betrug die SD der Änderung der 6MWD bis Woche 48 72.0 m in der Ataluren 10, 10, 20 mg/kg KG Gruppe. Bei einer 1:1-Randomisierung und einer Power von 85% ( $\alpha=0.05$ ) waren 210 Patienten (105 pro Gruppe) erforderlich, um einen Unterschied von 30 m detektieren zu können. Unter der Annahme, dass 5% der Patienten vorzeitig ausscheiden, sollten 220 Patienten (110 pro Gruppe) eingeschlossen werden.

***Randomisierung:*** Die Patienten wurden stratifiziert nach Alter (<9 vs.  $\geq 9$  Jahre), Dauer der Kortikosteroid-Vorbehandlung ( $\geq 12$  vs. <12-6 Monate vor BL) und der BL-6MWD ( $\geq 350$  vs. <350 m) im Verhältnis 1:1 randomisiert den Behandlungsgruppen zugeteilt. Die Patienten wurden den Gruppen mittels permutierter Block-Randomisierung zugeordnet, um eine ausgeglichene Verteilung hinsichtlich der Stratifizierungsfaktoren sowie der Zahl der Patienten pro Arm zu erreichen. Zur Randomisierung wurde das IVR/IWR-System eingesetzt.

***Interim-Analysen:*** Ein unabhängiges Experten-Komitee („data monitoring committee“, DMC) war für die Überwachung der Sicherheits-, Wirksamkeits- und anderer relevanter Daten verantwortlich und sollte Empfehlungen hinsichtlich der Fortführung der Studie und

gegebenenfalls notwendiger Design-Modifikationen aussprechen. Das DMC führte etwa alle sechs Monate verblindete Interim-Reviews durch und bewertete die Sicherheitsdaten.

### ***Studienpopulationen:***

ITT-Population: Alle randomisierten Patienten, die den Gruppen entsprechend der initialen Randomisierung zugeteilt worden waren, unabhängig davon ob die Patienten irgendein Prüfpräparat oder ein von der Randomisierung abweichendes Präparat erhalten hatten. Voraussetzung war das Vorliegen mindestens eines validen BL-6MWD- und eines validen 6MWD-Wert post-BL. Zwillinge wurden als unabhängige Personen angesehen. Alle Wirksamkeits-Endpunkte wurden in der ITT-Population analysiert.

Wie behandelt („as-treated“): Alle randomisierten Patienten, 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: Alle randomisierten Patienten, die die randomisiert zugeteilte Studienmedikation erhalten hatten, die Studie abschlossen, zumindest einen validen 6MWD-Wert zur BL-Visite und in Woche 48 hatten, eine Compliance zwischen  $\geq 75\%$  und  $\leq 125\%$  aufwiesen, keine der ausgeschlossenen Begleitmedikationen erhalten hatten und bei denen keine signifikanten Protokollverletzungen vorlagen. Signifikante Protokollverletzungen waren folgendermaßen definiert:  $< 5.5$  Monate Vortherapie mit Kortikosteroiden (vor BL), valide 6MWD von  $< 150$  m bei der Screening-Visite, 6MWD  $> 80.5\%$  des vorausgesagten Wertes zur Screening-Visite oder Visite in Woche 48 außerhalb des erlaubten Fensters von  $\geq 8$  d. Eine Liste dieser Patienten wurde vor der Entblindung zu Studienende erstellt. Diese Population konnte in unterstützenden Analysen des primären und der sekundären Wirksamkeits-Endpunkte ausgewertet werden.

Wie im SAP festgelegt wurde die PP-Population nicht analysiert. Die PP-Population wurde *a priori* definiert und dokumentiert für den Fall, dass entsprechende Analysen von Behörden-seite in Zukunft gefordert oder diese anderweitig benötigt werden sollten.

Auswertbare („evaluable“) Population: Alle „as-treated“ Patienten mit hinreichenden Daten, die eine Auswertung des jeweiligen Parameters von Interesse erlaubten (z.B. Ataluren-Plasmakonzentrationen, Erhebung der Alltagsaktivitäten und des Krankheitsstatus).

### ***Statistische Analysen:***

Der primäre Endpunkt (Änderung der 6MWD von BL bis Woche 48) wurde mittels ANCOVA nach multipler Imputation zum Ersatz fehlender Werte in der ITT-Population analysiert. Für fehlende Werte wurde das Markov-Chain-Monte-Carlo-Verfahren eingesetzt und 100 Imputationen wurden angesichts der erwarteten Zahl fehlender Werte als hinreichend erachtet. Im ANCOVA-Modell wurden die Stratifizierungs-Faktoren (Alter, Dauer der Kortikosteroid-Vorbehandlung und BL-6MWD) als Kovariaten berücksichtigt. Die MIANALYZE-Prozedur wurde eingesetzt, um die Ergebnisse der multiplen Imputation zu kombinieren. Die MIANALYZE-Prozedur lieferte Schätzungen der Behandlungsdifferenzen

und der korrespondierenden Standardfehler sowie Schätzungen der Freiheitsgrade, die in der t-Statistik zur Bestimmung der KIs und statistischer Tests eingesetzt wurden.

### Zusätzliche Analysen der 6MWD

Der primäre Endpunkt wurde für die Studie 020 zusätzlich in einem „mixed-model repeated-measures“ (MMRM)-Modell analysiert. Es wurde eine unstrukturierte Varianz-Kovarianz-Matrix angenommen und eine REML-Methode („restricted maximum likelihood“; default) eingesetzt. Neben den Stratifizierungsfaktoren und der BL-6MWD wurden Behandlung, Visite, die Interaktion zwischen Behandlung und Visite, zwischen Visite und BL-6MWD und zwischen jedem der Stratifizierungsfaktoren und Visite in das Model eingeschlossen. Die Visiten wurden als kategoriale Variable berücksichtigt. Der Behandlungseffekt in Woche 48 wurde in dieser Analyse ausgewertet.

Eine Randomisierungs-basierte ANCOVA-Methode war im SAP spezifiziert, die falls erforderlich für den primären Endpunkt eingesetzt werden konnte.

In eine weitere Analyse wurden nur die Daten des ersten Zwillings eingeschlossen, um zu evaluieren, ob die Zuteilung von Zwillingspaaren zu derselben Behandlungsgruppe einen Einfluss auf die primäre Wirksamkeitsanalyse gehabt haben könnte.

Der Einfluss der Stratifizierungsfaktoren (BL-6MWD, Dauer der Kortikosteroid-Vorbehandlung, Alter) auf die 6MWD wurde mittels ANCOVA unter Berücksichtigung zusätzlicher Interaktionsterms (Behandlung-Faktor) analysiert. Eine Analyse der BL-6MWD-Subgruppen (<300 m, ≥300 m bis <400 m, ≥400 m) war im SAP prospektiv definiert. Zur wissenschaftlichen Grundlage dieser Subgruppen-Analyse siehe Kapitel 4.4.2. Zur Berechnung der p-Werte für Interaktionen zwischen Subgruppe und Behandlung wurde F-Test-Statistik mit multipler Imputation eingesetzt.

### Sekundäre Wirksamkeits-Endpunkte

Die sekundären Endpunkte wurden für die ITT-Population analysiert. Die Analyse entsprach grundsätzlich derjenigen des primären Endpunkts, allerdings wurde der jeweilige BL-Wert als Faktor eingesetzt (siehe auch Operationalisierung der Endpunkte).

### ***Patientenfluss und Charakteristika der Studienpopulation***

Insgesamt 61 der 291 Patienten, die auf Eignung überprüft worden waren, entsprachen nicht den Einschlusskriterien. Zweihundertdreißig Patienten wurden randomisiert, wovon jeweils 115 Patienten der Placebo- bzw. Ataluren-Gruppe (10, 10, 20 mg/kg KG) zugeteilt wurden. Nur vier Patienten der Placebo-Gruppe (jeweils n=1 UE, Patient verloren, Patientenwunsch, Nicht-Compliance) und fünf Patienten der Ataluren-Gruppe (n=1 UE, n=3 Patientenwunsch, n=1 Nicht-Compliance) beendeten die Studie vorzeitig. Insgesamt 221 Patienten (96%) beendeten die Studie.

Die Patienten wurden in 54 Studienzentren in 18 Ländern rekrutiert. Die Mehrzahl der Patienten wurde in den USA, Kanada und europäischen Ländern in die Studie eingeschlossen, deren Versorgungskontext demjenigen in Deutschland ähnlich sein dürfte. Etwa 75% der eingeschlossenen Patienten beider Gruppen waren kaukasischer Herkunft (8.7% Placebo, 6.1% Ataluren keine Angaben), was etwas unterhalb des Prozentsatzes in Deutschland liegen dürfte.

Die Gruppen waren hinsichtlich der vordefinierten Stratifizierungsfaktoren (Alter <9 vs.  $\geq 9$  Jahre, Dauer der Kortikosteroid-Vorbehandlung  $\geq 12$  vs. <12-6 Monate, BL-6MWD  $\geq 350$  vs. <350 m) wie erwartet ausgeglichen. Etwas mehr als die Hälfte der Patienten war 9 Jahre oder älter (53.9% Placebo, 50.4% Ataluren) und 63.5% der Patienten beider Gruppen konnte bei der BL-Visite im 6MWT 350 m oder mehr ohne Hilfen gehen. Insgesamt 83.5% der Patienten beider Gruppen waren über 12 Monate oder mehr mit Kortikosteroiden vorbehandelt worden. Die Gruppen waren hinsichtlich Alter, Geschlecht, Größe, Gewicht und BMI vergleichbar (Tabelle 4-12). Die Krankheitscharakteristika wie Alter bei Diagnosestellung und Phänotyp waren zwischen den Gruppen vergleichbar. Alle drei Typen der Stop-Codons waren vertreten und UAG (Placebo) und UGA (Ataluren) kamen am häufigsten vor. Die Nonsense-Mutationen waren über die 79 Exons gleichmäßig verteilt und es wurde keine Häufung einer spezifischen Lokalisation beobachtet. Die mittlere BL-6MWD (Bereich) betrug 362.69 m (142.5, 526.0) in der Placebo-Gruppe und 364.04 m (166.8, 511.0) in der Ataluren-Gruppe und die Mittelwerte unterschieden sich um weniger als 2 m. Trotz der Intention, Patienten in der „ambulatory decline“ Phase zu rekrutieren d.h. ein etwas homogeneres Kollektiv als in der Studie 007, war die Patientenpopulation hinsichtlich der BL-6MWD heterogen (142.5-526.0 m).

Die demographischen und Krankheitscharakteristika der Patienten mit einer 6MWD von  $\geq 300$  bis <400 m zu Beginn der Studie waren zwischen den Behandlungsgruppen ausgeglichen.

In Woche 24 wurde bei mehr Patienten der Placebo- als der Ataluren-Gruppe die Dosis der Prüfpräparate adjustiert (59.6% vs. 34.2%; ITT). Per Einschlusskriterium waren alle Patienten mit Kortikosteroiden vorbehandelt und die Kortikosteroid-Therapie sollte bis zum Behandlungsende möglichst konstant bleiben. Hinsichtlich des Kortikosteroid-Typs ergaben sich zwischen den Gruppen keine nennenswerten Unterschiede (Placebo vs. Ataluren: Deflazacort 47.0% vs. 43.5%, Prednison: 32.2% vs. 33.0%, Prednisolon: 24.3% vs. 25.2%). Daten zur Kortikosteroid-Dosis wurden nicht analysiert. Nur 20.9% der Patienten erhielten zu irgendeinem Zeitpunkt während der Behandlung kardiale Medikamente (25.2% Placebo, 16.5% Ataluren). Der Prozentsatz der mit ACE-Inhibitoren behandelten Patienten war in der Placebo-Gruppe etwas höher (20.9% vs. 13.9%; 3.5% vs. 3.5% ARB, 3.5% vs. 1.7% Beta-Blocker). Keiner der Patienten wurde mit Aminoglykosiden behandelt.

Bei DMD übliche begleitende nicht-arzneiliche Therapie-Maßnahmen wurden bei der Mehrzahl der Patienten beider Gruppen eingesetzt (58.3% Placebo, 72.2% Ataluren). Es ergaben sich nur vergleichsweise geringe Gruppen-Unterschiede hinsichtlich einzelner physikalischer Therapiemaßnahmen. Der größte Unterschied bestand in der Zahl der Orthose-Anwender (34.8% Ataluren, 24.3% Placebo).

In beiden Gruppen lag die Compliance bei über 95%.

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 4-13: 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	Siehe vorheriges Nutzenbewertung (3)						
PTC124-GD-020-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 4-13 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Bei der Studie PTC124-GD-020-DMD handelt es sich um eine multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte internationale Parallelgruppenstudien der Phase 3. Eine adäquate Erzeugung der Randomisierungssequenz liegt vor. Eine permutierte Blockrandomisierung wurde eingesetzt, um eine ausgeglichene Zuteilung hinsichtlich der Stratifizierungsfaktoren sowie der Zahl der Patienten pro Arm zu erreichen. Die Gruppenzuteilung erfolgte verdeckt mittels eines durch eine unabhängige CRO entwickelten IVR/IWR-Systems. Die Dosierung der Studienmedikationen wurde zentral mittels IVR/IWR auf der Basis der Angaben der Patienten kalkuliert. Ataluren und Placebo waren äußerlich nicht unterscheidbar. Sowohl die Patienten als auch die Prüfarzte und das medizinische Personal waren bis zur Endauswertung der Studie verblindet. Das ITT-Prinzip wurde adäquat umgesetzt. Lediglich ein Patient pro Gruppen wurde protokollgemäß aus der ITT-Analyse ausgeschlossen. Bei beiden Patienten konnte via Gensequenzierung keine Non-Sense-Mutation im Dystrophin-Gens nachgewiesen werden und

sie wurden vorzeitig aus der Studie ausgeschlossen. Nach BL lag bei diesen Patienten kein valider 6MWD vor, was eine Voraussetzung für den Einschluss in die ITT-Population war. Nur vier Patienten der Placebo-Gruppe und fünf Patienten der Ataluren-Gruppe brachen die Studie vorzeitig ab. Fehlende Werte wurden mittels multipler Imputation ersetzt. Unter Berücksichtigung der Seltenheit der Erkrankung ist die Zahl der eingeschlossenen Patienten als vergleichsweise hoch einzuschätzen. Dies gilt umso mehr, als im SAP der Studie 020 eine Meta-Analyse beider zulassungsrelevanter Studien prospektiv geplant war. 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 4-14: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität	Gesundheits- bezogene Lebensqualität	Morbidität	Verträglichkeit
PTC124-GD-007-DMD <sup>1</sup>	Nein	Ja (PedsQL)	Ja (6MWD inklusive Progressor-Analyse, TFTs)	Ja
PTC124-GD-020-DMD	Nein	Ja (PODCI)	Ja (6MWD inklusive Progressor-Analyse, TFTs, NSAA, Alltags- aktivität/Symptome)	Ja

<sup>1</sup> Siehe vorherige Nutzenbewertung (3)

#### 4.3.1.3.1 Endpunkte – RCT

##### 4.3.1.3.1.1 Gehstrecke im 6MWT (primärer Endpunkt) und 10%ige Verschlechterung im 6MWD – 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 4-15: Operationalisierung der Endpunkte 6MWD und 10%ige Verschlechterung im 6MWD

Studie	Operationalisierung
PCT124-GD-020-DMD	<p>Das Gehvermögen der Patienten wurde mittels einer speziell für DBMD-Patienten modifizierten Version des 6MWT der ATS-Richtlinie erfasst (8,47), eines an der Universität Kalifornien-Davis entwickelten standardisierten Verfahrens, mit welchem die Gehstrecke innerhalb von 6 min ermittelt wird (34,44,48). Während des 6MWT durften die Patienten keine Gehhilfen benutzen und es waren keine Hilfsmittel außer Einlegesohlen erlaubt (nur bis zum Sprunggelenk). Den Prüfcentren 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. Unmittelbar vor dem Test sollten die Patienten auf Verletzungen der unteren Extremitäten untersucht werden und falls vorhanden sollte der Test ausfallen. Die Tests sollten in einem ausreichend breiten, ruhigen Flur ohne Teppichboden, der mit einem 25 m langen Massband markiert war, stattfinden. Der Start- bzw. Wendepunkt sollte jeweils mit Kegeln und halbkreisförmigen Pfeilen markiert werden. Die einzelnen Testschritte 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, sollte sich aber nicht hinsetzen und wurde aufgefordert, den Test so bald wie möglich fortzusetzen (8).</p> <p>Die 6MWD wurde während der Screening- und BL-Visite and alle 8 Wochen während der Behandlung mittels 6MWT erfasst. Während des 2-tägigen Klinikaufenthaltes zur BL-Visite und Behandlungsende wurde der Test jeweils an zwei aufeinander folgenden Tagen wiederholt. Es war geplant, die Tests per Videokamera zu überwachen, sofern eine Einverständniserklärung der Patienten/Betreuer vorlag, die keinen Einfluss auf die weitere Teilnahme an der Studie hatte.</p> <p>Wie im Manual und SAP festgelegt, wurde der 6MWT als valide eingestuft, wenn der Patient den Test wie geplant abschloss oder den Test ausschließlich aufgrund von Erschöpfung vorzeitig abbrach. Ein Pause - angelehnt an einer Wand - war erlaubt. Unter folgenden Bedingungen galt der 6MWT als nicht valide (8,34): Der Patient</p> <ul style="list-style-type: none"> <li>- folgte wiederholt nicht den Anweisungen,</li> <li>- behielt den Kurs wiederholt nicht bei,</li> <li>- ging wiederholt in eine entgegengesetzte Richtung,</li> <li>- zeigte Zeichen einer Verletzung der unteren Extremitäten,</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> <li>- Der Therapeut konnte den Test auch aus anderen Gründen als nicht valide einstufen. Diese Gründe sollten in einem 6MWT-Arbeitsblatt dokumentiert werden, ebenso wie Verletzungen, die einen Einfluss auf den 6MWT haben könnten (8).</li> </ul> <p>Grundsätzlich war die Validität der 6MWTs in einem Arbeitsblatt zu dokumentieren und falls der Test nicht valide war, war der Grund anzugeben. Im Falle der erste Test war nicht valide oder der Patient war zu erschöpft, um den ersten Test zu beenden, war der Prüfer via Manual angehalten, zu versuchen, den Test zu wiederholen (nach einer mindestens 10-minütigen Ruhepause an Tag 1 oder an Tag 2). Bei Testwiederholung war eine neues Arbeitsblatt anzufertigen, in dem notiert werden sollte, dass es sich um ein Wiederholung handelt. Falls mehr als ein valider Test für eine Visite vorlag, sollte die längste Gehstrecke analysiert werden (ausgenommen BL und Woche 48). Die BL-6MWD war als Mittelwert aus einem validen 6MWT an Tag 1 und 2 der BL-Visite definiert. Wenn nur ein valider 6MWD-Wert vorlag,</p>



	<p>sollte dieser Wert eingesetzt werden. Gleichmaßen sollte in Woche 48 respektive zu Behandlungsende verfahren werden (8,34).</p> <p>Patienten, die dauerhaft ihre Gehfähigkeit verloren, wurden auf Basis der UEs in den eCRFs identifiziert. Bei diesen Patienten wurde ab diesem Zeitpunkt ein Wert von 0 bei allen nachfolgenden Visiten eingegeben. Bei gehfähigen Patienten ohne valide 6MWD wurde der Wert als fehlend eingegeben (34).</p> <p>Der 6MWT ist ein mittlerweile etablierter und validierter primärer Endpunkt in Studien bei DMD-Patienten, der als Maß der Krankheitsprogression angesehen wird (18–20,46). Eine Veränderung des Populationsmittelwertes der 6MWD um 30 m wird allgemein als klinisch relevant akzeptiert, da diese ein Prognosefaktor für den zukünftigen Verlust der Gehfähigkeit ist (46).</p> <p>Die Daten wurden für die ITT-Population ausgewertet. Die statistische Auswertung ist in Kapitel 4.3.1.2.1 und Anhang 4-E beschrieben und weitere Daten zur Validität und Patientenrelevanz des Endpunkts und der Analysen sind in Kapitel 4.4.2 dargestellt.</p> <p><i>Sekundärer Endpunkt: Zeit bis zur 10%igen Verschlechterung der 6MWD</i></p> <p>Der Endpunkt wurde in der ITT-Population ausgewertet und fehlende Werte wurden wie in der Primäranalyse ersetzt. Bei Patienten, deren 6MWD sich nicht um 10% verschlechterte oder die aus der Studie ausschieden, wurde der letzte verfügbare 6MWD-Wert eingegeben. Ein Verlust der Gehfähigkeit wurde mit einer 10%igen Verschlechterung gleichgesetzt. Die Unterschiede zwischen den Behandlungsarmen wurden in einem Cox-proportionalen Hazard-Model analysiert, in welches die BL-6MWD, Behandlung sowie die drei Stratifizierungsfaktoren eingeschlossen wurden. Ein Kaplan-Meier-Plot war vorgesehen. Zur Analyse der 10%igen Verschlechterung wurden die imputierten Werte aus der Primäranalyse eingesetzt und der Anteil der Patienten mit einer andauernden 10%igen Verschlechterung in Woche 48 wurde kalkuliert (34).</p> <p>Die Zeit bis zu einer 10%igen Verschlechterung der 6MWD korreliert mit der Zeit bis zum Verlust der Gehfähigkeit (19).</p>
PTC124-GD-007-DMD	Die Studie PTC124-GD-007-DMD wurde in einem separaten Dossier zur Nutzenbewertung dargestellt, auf welches hier verwiesen wird (Nutzenbewertungsverfahren zum Wirkstoff Ataluren, Beschlussfassung: 21.05.2015) (3).

*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 4-16: Bewertung des Verzerrungspotenzials für Endpunkt 6MWD 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-020-DMD	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt 6MWD der Studie PTC124-GD-020-DMD wird unter Berücksichtigung der in Tabelle 4-16 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 Prüffärzte, das medizinische Personal und die Patienten waren verblindet. Es fand eine adäquate Operationalisierung des Endpunkts statt. Der 6MWT wurde unter standardisierten Bedingungen unter Einsatz eines Manuals, in dem die Testbedingungen inklusive Angaben zur Validität der Tests detailliert beschrieben waren, durchgeführt. Der 6MWT ist ein bei DMD-Patienten validierter und mittlerweile akzeptierter primärer Endpunkt in DMD-Studien (z.B. (18); siehe auch früheres Nutzendossier (38)). Auch die Wahl des Stratifizierungskriteriums (<350 m vs. ≥350 m) war klinisch begründet (siehe auch früheres Nutzendossier zur Studie PTC124-GD-007-DMD (38)).

In der vorausgegangenen Ataluren-Studie PTC124-GD-007-DMD fand eine adäquate Umsetzung des ITT Prinzips statt. Alle Patienten, die mindestens eine Dosis der Prüfpräparate erhalten hatten und die den Kriterien der ITT-Population entsprachen, wurden in die ITT-Population eingeschlossen. 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.

*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.*

Alle Patienten der Studie PTC124-GD-020-DMD beendeten den 6MWT zu BL/Woche 1 und 95.6% der Patienten beider Gruppen in Woche 48. Zu BL/Woche 1 lagen bei allen Patienten der Placebo-Gruppe und 98.2% der Patienten der Ataluren-Gruppe zwei valide 6MWTs vor (1.8%: ein valider Test). Zu Studienende waren bei jeweils fünf Patienten der Placebo- und Ataluren-Gruppe (4.4%) keine validen Testergebnisse verfügbar. Protokollgemäß wurde bei

Patienten, die während der Studie ihre Gehfähigkeit verloren, ein valider 6MWD-Wert von 0 vergeben.

### Primär- und Sensitivitätsanalysen der 6MWD (PTC124-GD-020-DMD, ITT)

Der Unterschied der Änderung der 6MWD von BL bis Woche 48 zwischen Ataluren (n=114) und Placebo (n=114) betrug 12.98 m (LSM; 95% KI -7.44, 33.39) zugunsten der Ataluren-Gruppe (ITT-Population). Der Unterschied zugunsten von Ataluren fiel geringer aus als erwartet und war in der Primäranalyse nicht signifikant (p=0.213; ANCOVA nach multipler Imputation; Tabelle 4-17). Tabelle 4-18 zeigt die zeitliche Veränderung der 6MWD im Studienverlauf. Über die Behandlungsdauer zeigte sich ein positiver Effekt von Ataluren verglichen mit Placebo im Sinne einer geringeren Abnahme der Gehstrecke, die Unterschiede waren im Gruppenvergleich allerdings zu keinem Zeitpunkt signifikant.

Tabelle 4-17: Änderung der 6MWD von BL bis Woche 48 für PTC124-GD-020-DMD (ITT, n=114 Placebo, n=114 Ataluren, Gesamtpopulation)

Gruppe	n <sup>2</sup>	BL-6MWD M (±SD) m	6MWD in Woche 48 M (±SD) m	Änderung von BL bis Woche 48 LSM (±SE) m	Unterschied vs. Placebo LSM (±SE) m <sup>1</sup>	95% KI (p-Wert) <sup>1</sup>
Placebo	114/ 109	363.54 (81.266)	306.50 (157.078)	-60.67 (9.323)	-	-
Ataluren	114/ 109	365.19 (72.610)	326.34 (134.079)	-47.69 (9.247)	12.98 (10.415)	-7.44, 33.39 (0.213)

Quelle (Studienbericht 020): Tabelle 14.2.1.2.1.1, Tabelle 14.2.1.3.1

<sup>1</sup> ANCOVA mit multipler Imputation. Faktoren: Stratifizierungsfaktoren (Alter <9 vs. ≥9 Jahre, BL 6MWD <350 vs. ≥350 m, Dauer der Kortikosteroid-Vorbehandlung ≥6 bis <12 vs. ≥12 Monate), Behandlung und BL-6MWD

<sup>2</sup> Anzahl Patienten zu BL / Woche 48

Tabelle 4-18: Änderung der mittleren 6MWD über den Studienverlauf für PTC124-GD-020-DMD (ITT)

Behandlungsgruppe	LSM ( $\pm$ SE) BL vs. Woche 24 (m)	LSM ( $\pm$ SE) Woche 24 vs. Placebo (m)	LSM ( $\pm$ SE) BL vs. Woche 32 (m)	LSM ( $\pm$ SE) Woche 32 vs. Placebo (m)	LSM ( $\pm$ SE) BL vs. Woche 40 (m)	LSM ( $\pm$ SE) Woche 40 vs. Placebo (m)	LSM ( $\pm$ SE) BL vs. Woche 48 (m)	LSM ( $\pm$ SE) Woche 48 vs. Placebo (m)
Placebo	-29.52 (6.191)	9.21 (6.879)	-48.37 (7.677)	14.54 (8.578)	-53.65 (8.189)	7.68 (9.113)	-60.67 (9.323)	12.98 (10.415)
Ataluren	-20.31 (6.123)	95% KI -4.27, 22.69 p=0.181	-33.84 (7.639)	95% KI -2.28, 31.35 p=0.090	-45.96 (8.113)	95% KI -10.18, 25.54 p=0.399	-47.69 (9.247)	95% KI -7.44, 33.39 p=0.213

ANCOVA mit multipler Imputation; Faktoren: Stratifizierungsfaktoren Alter, BL-6MWD-Gruppe, Dauer der Kortikosteroid-Vorbehandlung, Behandlung und BL-6MWD

p-Wert für den Vergleich der Behandlungsgruppen.

Quelle (Studienbericht 020): Tabelle 14.2.1.3.1

Die prä-definierten Sensitivitätsanalysen der 6MWD waren in ihrer Ausrichtung und Ausprägung der Primäranalyse gleichgerichtet und bestätigen diese.

Tabelle 4-19: Änderung der mittleren 6MWD über 48 Wochen für PTC124-GD-020-DMD in prä-spezifizierten Sensitivitätsanalysen (ITT)

Gruppe / Analyse	$\Delta$ LSM ( $\pm$ SE) in Woche 48 [95% KI] (m)	$\Delta$ LSM ( $\pm$ SE) vs. Placebo in Woche 48 (m)	95% KI (m)	p-Wert vs. Placebo
<i>MMRM-Modell (Studienreport 020 Tabelle 14.2.1.4)<sup>1</sup></i>				
Placebo (n=114)	-60.75 (9.339)		-	-
Ataluren (n=114)	-47.27 (9.251)	13.49 (10.374)	-6.96, 33.93	0.195
<i>Nur erster Zwilling eingeschlossen (Studienreport 020 Tabelle 14.2.1.3.2)<sup>2,3</sup></i>				
Placebo	-62.18 (9.446)	-	-	-
Ataluren	-50.11 (9.359)	12.07 (10.583)	-8.67, 32.81	0.254

<sup>1</sup> Faktoren: Stratifizierungsfaktoren Alter, BL-6MWD-Gruppe, Dauer der Kortikosteroid-Vorbehandlung, BL-6MWD, Behandlung, Visite, Interaktion zwischen Behandlung und Visite, zwischen Visite und BL-6MWD und Visite und jedem der Stratifizierungsfaktoren.

<sup>2</sup> ANCOVA mit multipler Imputation; Faktoren: Stratifizierungsfaktoren Alter, BL-6MWD-Gruppe, Dauer der Kortikosteroid-Vorbehandlung, Behandlung, und BL-6MWD

<sup>3</sup> Sieben Zwillingsskinder wurden aus der Analyse ausgeschlossen.

Die Zeit bis zu einer dauerhaften 10%igen Verschlechterung der 6MWD wurde als Sekundärendpunkt evaluiert. Die HR für Ataluren vs. Placebo betrug 0.75 (95% KI 0.51-1.12, p=0.1603), was einer 25%igen Reduktion des Hazards unter Behandlung mit Ataluren entsprach (ITT).

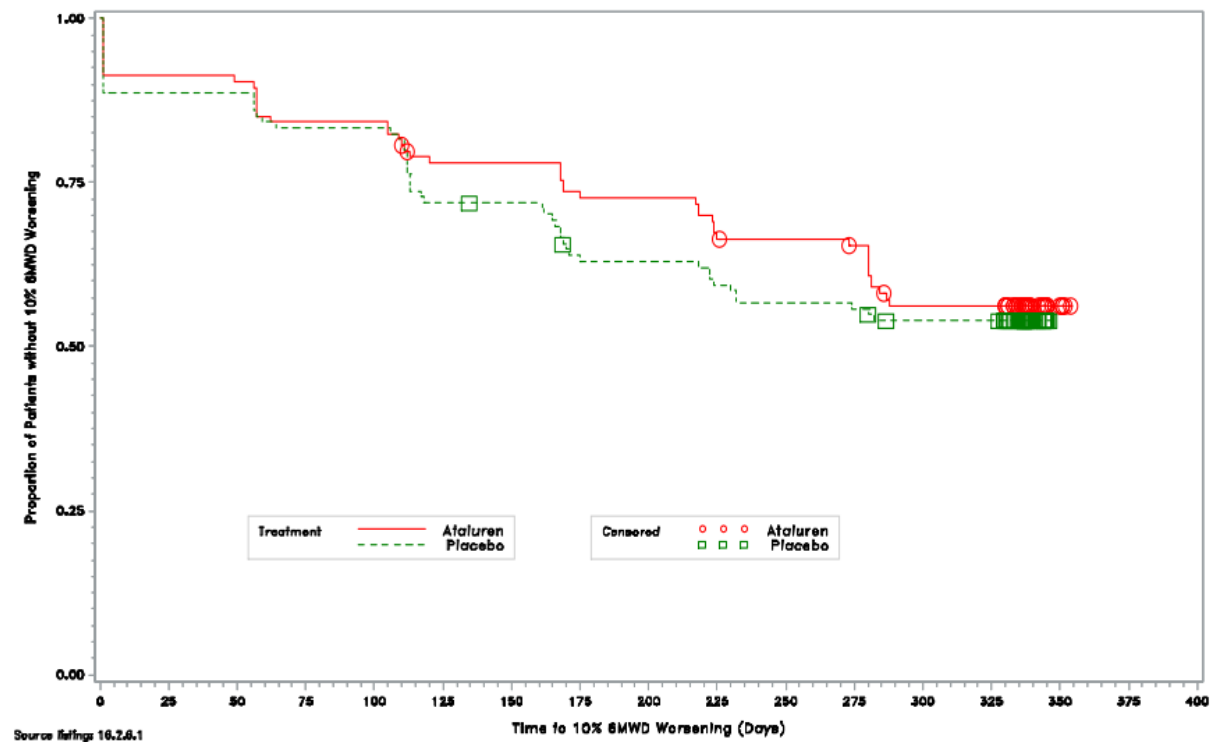


Abbildung 2: Kaplan-Meier-Plot der Zeit bis zu einer dauerhaften 10%igen Verschlechterung der 6MWD (ITT) für PTC124-GD-020-DMD; Ataluren versus Placebo

Quelle (Studienbericht 020): Abbildung 14.2.1.5.1.1

Neun Patienten der Ataluren-Gruppe (7.9%) und 14 Patienten der Placebo-Gruppe (12.3%) verloren während der Studie ihre Gehfähigkeit. Die HR der Zeit bis zum Verlust der Gehfähigkeit für den Vergleich von Ataluren vs. Placebo betrug 0.69 (95% KI 0.29, 1.64,  $p=0.404$ ; ITT), was einer 31%igen Risikoreduktion entspricht. Die Gruppenunterschiede wurden in einem Cox-proportionalen Modell analysiert.

## Fazit

Der primäre Endpunkt (Änderung der 6MWD von BL bis Woche 48) zeigte in der Studie PCT124-GD-020-DMD einen deutlichen Trend zur Überlegenheit von Ataluren gegenüber Placebo in der Gesamtpopulation (ITT). Statistische Signifikanz wurde jedoch nicht erreicht, was auf die große Heterogenität des Patientenkollektivs zurückzuführen war (siehe Kapitel 4.4.2).

Um die Ergebnisse der beiden eingeschlossenen Studien zusammenzufassen wurden Meta-Analysen unter Verwendung der Ergebnisse der Studien PCT124-GD-007-DMD und PCT124-GD-020-DMD durchgeführt.

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.

Da bei Studien für sehr seltene Erkrankungen immer geringe Patientenzahlen vorliegen, ist es methodisch sinnvoll, bei Existenz von mehr als einer Studie die Ergebnisse in einer Meta-Analyse zusammenzufassen (7). Dabei wird auch das in der Wissenschaft allgemein geforderte Prinzip der Replikation erfüllt. Aus medizinischer Sicht gibt es keinen offensichtlichen Grund, die beiden Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD nicht zusammenzufassen. Die beiden Studien sind in Bezug auf die medizinischen Fragestellungen, das Studiendesign, die Patientenpopulationen, die Endpunktparameter, und die Auswertungsmethodik nahezu identisch. Der einzige klar definierte Unterschied zwischen den beiden Studien ist der Unterschied bei den Einschlusskriterien in Bezug auf das Alter. Während in der Studie PTC124-GD-007-DMD Kinder ab 5 Jahren eingeschlossen wurden, war in der Studie PTC124-GD-020-DMD der Einschluss von Patienten erst ab einem Alter von 7 Jahren erlaubt. Da das Alter der Patienten einen Einfluss auf die Baseline-Gehstrecke nimmt, und die Baseline-Gehstrecke einen Effektmodifikator der 6MWD (und anderer Endpunkte) darstellt (38,49), ist ein Teil der Heterogenität zwischen den Studien durch diesen Unterschied erklärbar. Die Frage der Effektmodifikation durch die Baseline-Gehstrecke (synonym mit Krankheitsstadium), beziehungsweise die unterschiedliche Sensitivität der Erhebungsinstrumente in Abhängigkeit des Krankheitsstadiums (Alter, Baseline-Gehstrecke) wird nachfolgend in den Subgruppenanalysen in Kapitel 4.3.1.3.2 adressiert.

### Gehstrecke im 6MWT

Die Meta-Analyse für den Endpunkt Gehstrecke im 6MWT zeigte unter Einschluss der kITT Population für die Studie PTC124-GD-007-DMD eine statistisch signifikante Überlegenheit von Ataluren gegenüber Placebo (Abbildung 3). Für die Studie PTC124-GD-007-DMD wurde hierfür die primäre, im SAP definierte, MMRM-Analysemethode (RANCOVA) eingesetzt. Der Gesamteffektschätzer lag bei 18.21 m [95% KI: 1.55, 34.86]. Es existiert kein Hinweis auf Heterogenität zwischen den Studien ( $p=0.39$ ).

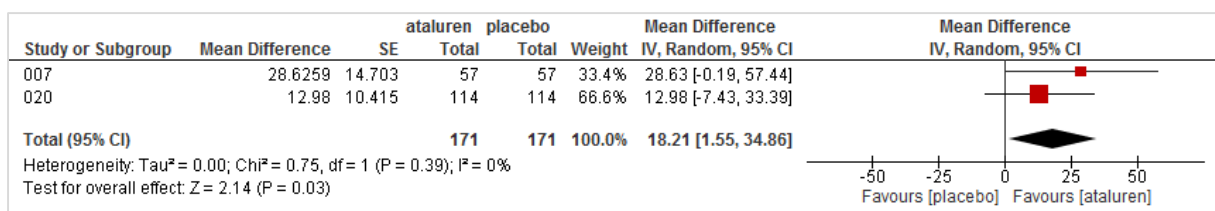


Abbildung 3: Meta-Analyse für 6MWT aus PTC124-GD-007-DMD (kITT, MMRM) und PTC124-GD-020-DMD (ITT, ANCOVA); Ataluren versus Placebo

Quellen: PTC124-GD-007-DMD; prädefiniertes MMRM Modell, SAS output 6MWT (12)  
PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.1.3.1 CSR 020 (8)

#### Sensitivitätsanalysen:

Die Meta-Analyse für den Endpunkt Gehstrecke im 6MWT zeigte auch unter Einschluss der ITT Populationen beider Studien eine signifikante Überlegenheit von Ataluren gegenüber Placebo (Abbildung 4). Für die Studie PTC124-GD-007-DMD wurden hierfür die nicht korrigierte ITT Population und die primäre, im SAP definierte, MMRM-Analysemethode (RANCOVA) eingesetzt. Der Gesamteffektschätzer lag bei 17.16 m [95% KI: 0.20; 34.11]. Es existiert kein Hinweis auf Heterogenität zwischen den Studien ( $p=0.47$ ); dies bestätigt, dass es methodisch sinnvoll ist, die beiden Studien zusammenzufassen.

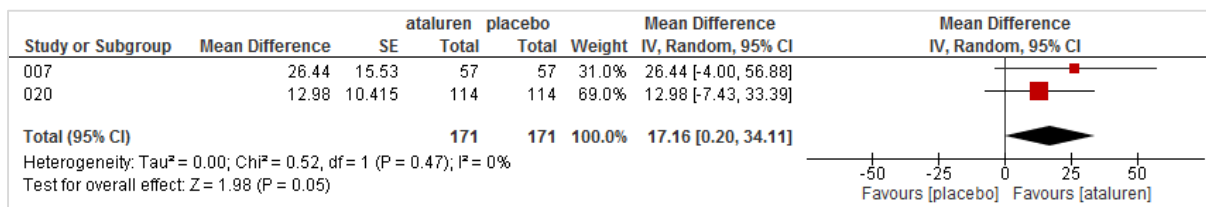


Abbildung 4: Meta-Analyse für 6MWT aus PTC124-GD-007-DMD (ITT, MMRM) und PTC124-GD-020-DMD (ITT, ANCOVA); Ataluren versus Placebo

Quellen: PTC124-GD-007-DMD; prädefiniertes MMRM Modell, Tabelle 14.2.1.8.5B CSR 007 (9)  
PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.1.3.1 CSR 020 (8)

Die Ergebnisse für die ITT Population der Studie PTC124-GD-007-DMD waren durch eine Knieverletzung eines Patienten (501-014) im Placeboarm vor der Erhebung der Baseline-Gehstrecke deutlich zuungunsten der Atalurenbehandlung verzerrt (38). Die Baselinewerte für diesen Patienten hätten nach den Vorgaben des SAP als invalide klassifiziert werden müssen. Aus diesem Grund wurde die Meta-Analyse unter Verwendung der korrigierten ITT Population (kITT) der Studie PTC124-GD-007-DMD wiederholt. Für die Auswertung wurde eine modifizierte MMRM Methode eingesetzt, in der ein zusätzlicher Interaktionsterm in das Modell aufgenommen wurde. Diese modifizierte MMRM Analyse war Grundlage der befristeten Zulassung auf Basis der Phase 2b Studie PTC124-GD-007-DMD (Abbildung 5). Unter Verwendung der modifizierten MMRM Analyse und lag der Gesamteffektschätzer für die kITT Population bei 20.40 m [95% KI: 2.45; 38.36]. Auch in dieser Analyse war die Heterogenität zwischen den Studien ausreichend gering ( $p=0.27$ ), um die Zusammenfassung der Ergebnisse in einer Meta-Analyse zu rechtfertigen. Der Gesamteffektschätzer der Meta-Analyse lag unter Einschluss der kITT Population der Studie PTC124-GD-007-DMD über dem Wert des Gesamteffektschätzers für die Meta-Analyse, in der die ITT Population in beiden Studien ausgewertet wurde.



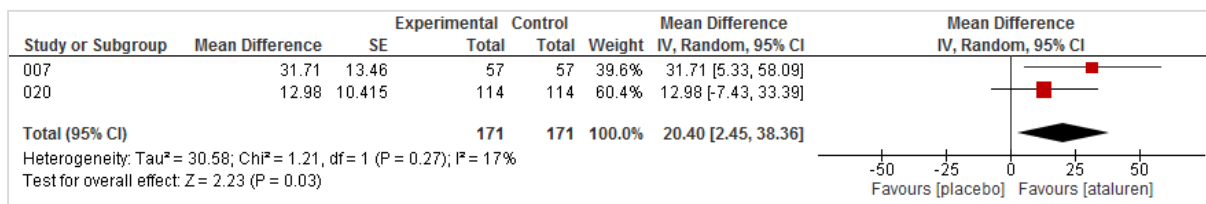


Abbildung 5: Meta-Analyse für 6MWT aus PTC124-GD-007-DMD (kITT, mMMRM) und PTC124-GD-020-DMD (ITT, ANCOVA); Ataluren versus Placebo

Quellen: PTC124-GD-007-DMD; augmentiertes MMRM Modell, Tabelle 14.2.1.24.2S CSR 007 (9)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.1.3.1 CSR 020 (8)

In den beiden eingeschlossenen Studien wurden ähnliche, jedoch nicht identische Auswertungsmethoden der Endpunktparameter verwendet. So wurde in der Studie PTC124-GD-007-DMD eine MMRM Methode eingesetzt, in der Studie PTC124-GD-020-DMD wurde eine ANCOVA-Analyse verwendet. Da in der vorangegangenen Phase 2b Studie PTC124-GD-007-DMD jedoch auch eine ANCOVA-Analyse als Sensitivitätsanalyse eingesetzt wurde, konnte eine Meta-Analyse durchgeführt werden, in der in beiden Studien eine ANCOVA-Analyse zur Auswertung des primären Endpunkts zum Einsatz kam (Abbildung 6). Der Gesamteffektschätzer lag bei 18.46 m [95% KI: 2.08; 34.84], es zeigte sich kein Hinweis auf eine Heterogenität zwischen den Studien. Die unterschiedlichen Analysemethoden führen somit zu keinen qualitativen Unterschieden in der Gesamtaussage.

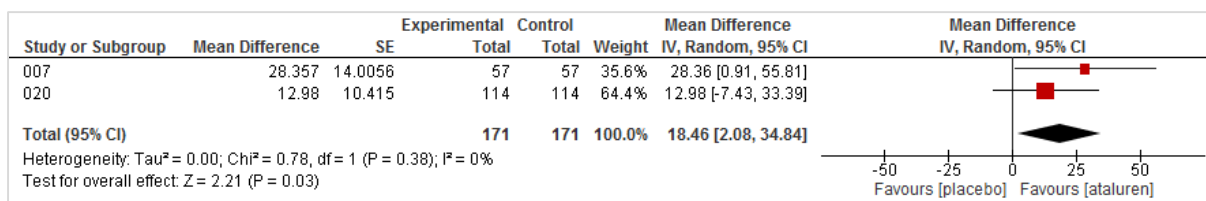


Abbildung 6: Meta-Analyse für 6MWT aus PTC124-GD-007-DMD (ITT, ANCOVA) und PTC124-GD-020-DMD (ITT, ANCOVA); Ataluren versus Placebo

Quellen: PTC124-GD-007-DMD; ANCOVA Modell, Tabelle 14.2.2.15 CSR 007 (9)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.1.3.1 CSR 020 (8)

In der Zusammenschau der Meta-Analysen ergeben sich mit allen unterschiedlichen Auswertungen der Studie PTC124-GD-007-DMD für die jeweiligen Gesamteffektschätzer signifikante Unterschiede der Gehstrecke im 6MWT (Ataluren im Vergleich zu Placebo) im in der Größenordnung von ca. 17-20 Metern über eine Studiendauer von 48 Wochen. In den nachfolgenden Subgruppenanalysen zeigt sich jedoch, dass der wahre Therapieeffekt von Ataluren in den eingeschlossenen heterogenen Gesamtpopulationen wahrscheinlich deutlich unterschätzt wird (s. Kapitel 4.3.1.3.2).

#### Zeit bis zur 10%igen Verschlechterung im 6MWT

Die Zeit bis zur 10%igen Verschlechterung korreliert mit der Zeit bis zum Verlust der Gehfähigkeit (19). Die 10%ige Verschlechterung kann somit als Surrogat für den irreversiblen Verlust der Gehfähigkeit angesehen werden. Unter Einschluss der ITT Populationen beider eingeschlossenen Studien zeigte sich in der Meta-Analyse eine

signifikante Reduktion des Hazards der 10%igen Verschlechterung für Patienten, die mit Ataluren behandelt wurden. Der Gesamteffektschätzer der Hazard Ratios lag bei 0.67 [95% KI: 0.48; 0.94] (Abbildung 7). Es ergab sich kein Hinweis für eine Heterogenität zwischen den Studien. Es ist bezeichnend, dass das Hazard Ratio der 10%igen Verschlechterung tatsächlich numerisch sehr ähnlich zum relativen Risiko des Verlustes der Gehfähigkeit ist (Abbildung 8).

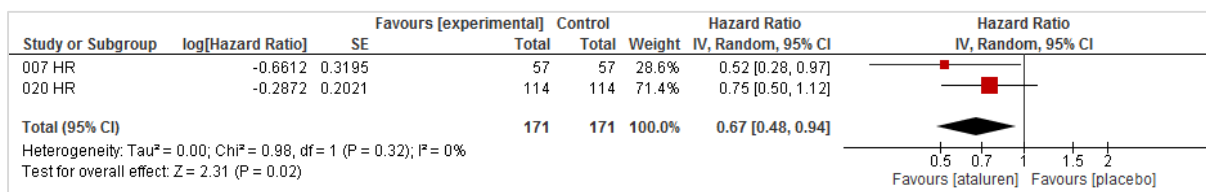


Abbildung 7: Meta-Analyse für Zeit bis zur 10%igen Verschlechterung aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quelle:

PTC124-GD-007-DMD; Cox proportionales Hazard Modell, Tabelle 14.2.2.11 CSR 007 (9)

PTC124-GD-020-DMD; Cox proportionales Hazard Modell, Tabelle 14.2.1.5.1 CSR 020 (8)

#### Risiko des Verlustes der Gehfähigkeit („loss of ambulation“)

Der Verlust der Gehfähigkeit ist eine der schwerwiegendsten Komplikationen der Duchenne Muskeldystrophie, mit extremen Folgen für Patienten und deren Familien. Obwohl die Einschlusskriterien der beiden Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD nicht darauf ausgelegt waren, den Verlust der Gehfähigkeit zu erfassen, traten Ereignisse mit geringer Inzidenz dennoch in beiden Studien auf. Das relative Risiko in den eingeschlossenen Patientenpopulationen innerhalb von 48 Wochen die Gehfähigkeit zu verlieren, lag in der Meta-Analyse der beiden Studien bei 0.65 [95% KI: 0.33; 1.26] für den Vergleich im Ataluren-Arm gegenüber dem Placebo-Arm (Abbildung 8). Es zeigte sich eine hohe Homogenität der Ergebnisse in beiden Studien (p-Wert im Heterogenitätstest: p=0.96). Aus den oben genannten Gründen wurde - wie zu erwarten – eine statistische Signifikanz für diesen Endpunkt über die limitierte Studiendauer nicht erreicht.

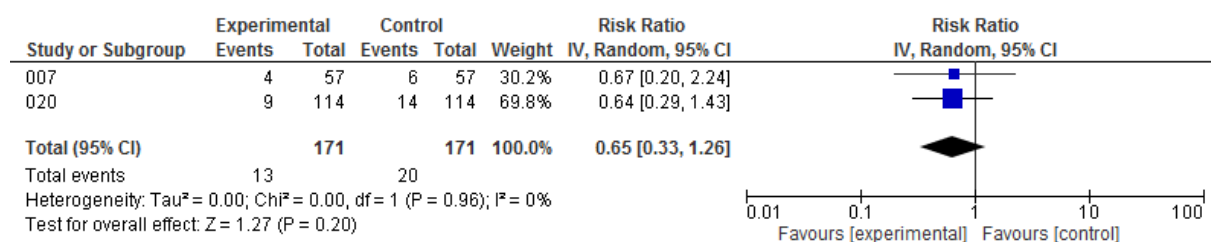


Abbildung 8: Meta-Analyse für Verlust der Gehfähigkeit aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quelle:

PTC124-GD-007-DMD; Tabelle 14.2.2.14 CSR 007 (9)

PTC124-GD-020-DMD; Tabelle 14.2.1.5.2 CSR 020 (8)

Die Ergebnisse der Subgruppenanalysen für diesen Endpunkt werden in Kapitel 4.3.1.3.2.1 dargestellt und in Kapitel 4.4.2.5.3 diskutiert.

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.

#### 4.3.1.3.1.2 Zeit-Funktionstests – 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 4-20: Operationalisierung von Endpunkt TFT („timed function tests“)

Studie	Operationalisierung
PTC124-GD-020-DMD	<p>Bei den TFTs wurde die Zeit erfasst, die die Patienten benötigten, um 10 m zu gehen oder zu laufen und vier Stufen von Standardgröße hinauf- bzw. hinabzusteigen. Die TFTs wurden detailliert in einem Manual beschrieben, das den Studienzentren zur Verfügung gestellt wurde (8,34).</p> <p>Die TFTs wurden zur Screening-Visite, BL, alle 8 Wochen während der Behandlung und Woche 48 bzw. zu Behandlungsende erhoben (34). Das Manual enthielt detaillierte Vorgaben hinsichtlich der verbalen Instruktionen während der Tests (8).</p> <p>Die Fähigkeit, 10 m zu gehen/laufen, wurde während des NSAA überprüft und ist in Tabelle 4-24 beschrieben. Die Zeit, die benötigt wurde, um aus dem Liegen aufzustehen, wurde ebenfalls während des NSAA erfasst.</p> <p>Der Test, vier Treppenstufen hinauf zu gehen, wurde wie folgt bewertet:</p> <ol style="list-style-type: none"> <li>1. Unfähig, 4 Treppenstufen zu ersteigen;</li> <li>2. Steigt 4 Treppenstufen mit Stillstand (d.h. mit jeweils einem Fuß und beiden Füßen auf einer Stufe vor dem Ersteigen der nächsten Stufe), beide Arme an einer oder beiden Handläufen;</li> <li>3. Steigt 4 Treppenstufen mit Stillstand (siehe oben), ein Arm an einem Handlauf;</li> <li>4. Steigt 4 Treppenstufen mit Stillstand (siehe oben), ohne den Handlauf zu Hilfe zu nehmen;</li> <li>5. Steigt 4 Treppenstufen alternierend mit beiden Füßen, Handlauf wird unterstützend benötigt;</li> <li>6. Steigt 4 Treppenstufen alternierend mit beiden Füßen, ohne den Handlauf zu Hilfe zu nehmen (8,34).</li> </ol> <p>Vier Treppenstufen hinabgehen: Wie 4 Treppenstufen hinaufgehen kategorisiert. Für den Fall, dass der Test wiederholt wurde, sollte das schnellste/beste Testergebnis dokumentiert werden (8).</p> <p>Die TFTs (Zeit, die benötigt wurde) wurden in der ITT-Population entsprechend des primären Endpunkts analysiert. Die folgenden Algorithmen wurden bei der Analyse eingesetzt: Wenn die für die Tests benötigte Zeit 30 sec überschritt (maximal für den Test zulässige Zeit) oder der Patient aufgrund einer Krankheitsprogression nicht in der Lage war, an dem Test teilzunehmen, wurde ein Wert von 30 sec eingesetzt. Bei Patienten, die aus anderen Gründen, an dem Test nicht teilnehmen konnten (z.B. Beinfaktur), wurde der Wert als fehlend eingegeben (34).</p> <p>In zwei Sensitivitätsanalysen wurde ein oberes Limit von 45 sec sowie der jeweils höchste Wert analysiert (34).</p>
PTC124-GD-007-DMD	<p>Die Studie PTC124-GD-007-DMD wurde in einem separaten Dossier zur Nutzenbewertung dargestellt, auf welches hier verwiesen wird (Nutzenbewertungsverfahren zum Wirkstoff Ataluren, Beschlussfassung: 21.05.2015).</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 4-21: 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-020-DMD	niedrig	ja	ja	ja	ja	niedrig <sup>1</sup>

<sup>1</sup> Hinsichtlich der Zeit, die benötigt wird, um aus dem Liegen aufzustehen, wird das Verzerrungspotential als hoch eingestuft.

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential für den Endpunkt TFTs (Zeit, die benötigt wurde) wird unter Berücksichtigung der in Tabelle 4-21 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F mit Ausnahme der Zeit, die benötigt wird, um aus dem Liegen aufzustehen, als niedrig eingeschätzt. Wie in Kapitel 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. Die Endpunkte wurden adäquat operationalisiert und den Endpunkterhebern wurde ein ausführliches Manual, in dem die Tests beschrieben waren, zur Verfügung gestellt. Der Endpunkt „Zeit, um aus dem Liegen aufzustehen“ wurde mit Amendement 1.2 des Studienprotokolls (datiert 13.06.2013) aufgenommen, d.h. nach Einschluss des ersten Patienten (26.03.2013). Daher wird das Verzerrungspotential für diesen Einzeltest als hoch eingestuft. TFTs sind ein in DMD-Studien häufig eingesetztes und etabliertes Instrument. In der früheren Nutzenbewertung der Ataluren-Studie 007 stuft der G-BA die Operationalisierung des Endpunkts TFTs und der Einzeltests, die derjenigen in Studie 020 weitgehend entsprach, als nachvollziehbar ein (3).

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.

In der Studie PTC124-GD-020-DMD nahm die Zeit, die die Patienten benötigen, um vier Treppenstufen hinaufzusteigen, bis Woche 48 in der Ataluren-Gruppe um 1.43 sec (LSM) weniger zu als in der Placebo-Gruppe (ITT  $p=0.058$ ; Tabelle 4-22). Der Unterschied zugunsten von Ataluren zeigte sich bereits in Woche 24 und blieb bis zum Behandlungsende erhalten.

Die Zeit, die benötigt wurde, um vier Treppenstufen hinabzusteigen, nahm bis Woche 48 in der Ataluren-Gruppe weniger stark zu als in der Placebo-Gruppe (LSM -1.97 sec, 95% KI -3.52, -0.43; ITT) und der Unterschied zugunsten von Ataluren war in der ANCOVA-Analyse mit multipler Imputation signifikant ( $p=0.012$ ; Tabelle 4-22).

In der ITT-Population betrug der Unterschied der Änderung der Zeit, um 10 m zu gehen oder zu laufen, in Woche 48 -1.07 sec (LSM, 95% KI -2.40, 0.27,  $p=0.117$ ; ITT) zugunsten von Ataluren (Tabelle 4-22). Der Unterschied zwischen der Ataluren- und Placebo-Gruppe war bereits nach 24 Wochen erkennbar und blieb bis zum Behandlungsende erhalten.

Tabelle 4-22: Änderung der Zeit für TFTs von BL bis Woche 48, PTC124-GD-020-DMD, (ITT; n=114 Placebo, n=114 Ataluren Gesamtpopulation) (8)

Gruppe	n BL/ Woche 48	M ( $\pm$ SD) BL (sec)	M ( $\pm$ SD) Woche 48 (sec)	$\Delta$ LSM ( $\pm$ SE) in Woche 48 (sec)	$\Delta$ LSM ( $\pm$ SE) vs. Placebo [95% KI] (sec)	p-Wert vs. Placebo
<i>Zeit, um 4 Treppenstufen hinaufzusteigen (Studienbericht 020: Tabelle 14.2.2.3.1, Tabelle 14.2.2.3.2<sup>1</sup>)</i>						
Placebo	111/ 106	6.37 (5.652)	10.61 (10.347)	5.31 (0.660)	-	-
Ataluren	111/ 107	5.92 (5.090)	8.81 (8.572)	3.88 (0.660)	-1.43 (0.754) [-2.90, 0.05]	0.058
<i>Zeit, um 4 Treppenstufen hinabzusteigen (Studienbericht 020: Tabelle 14.2.2.4.1, Tabelle 14.2.2.4.2)</i>						
Placebo	108/ 106	4.78 (4.516)	9.25 (10.194)	4.75 (0.694)	-	-
Ataluren	111/ 108	5.02 (4.968)	7.36 (8.200)	2.78 (0.685)	-1.97 (0.787) [-3.52, -0.43]	<b>0.012</b>
<i>Zeit, um 10 m zu gehen oder zu rennen (Studienbericht 020: Tabelle 14.2.2.2.1, Tabelle 14.2.2.2.2)</i>						
Placebo	114/ 110	6.80 (2.919)	10.27 (8.418)	3.43 (0.603)	-	-
Ataluren	114/ 109	6.63 (3.073)	8.80 (6.868)	2.36 (0.597)	-1.07 (0.680) [-2.40, 0.27]	0.117
<i>Zeit, um aus dem Liegen aufzustehen (Studienbericht 020: Tabelle 14.2.2.1.1, Tabelle 14.2.2.1.2)</i>						
Placebo	110/98	9.58 (7.785)	13.23 (10.954)	5.51 (0.728)	-	-
Ataluren	109/104	9.14 (7.565)	12.96 (9.851)	5.15 (0.718)	-0.36 (0.817) [-1.96, 1.24]	0.661

ANCOVA mit multipler Imputation; Faktoren: Stratifizierungsfaktoren (Alter, BL-6MD, Dauer der Kortikosteroid-Vorbehandlung), Behandlung und BL-TFT

### Sensitivitätsanalysen

Bei der primären Analyse wurde ein Wert von 30 sec eingegeben, wenn die für die Tests benötigte Zeit 30 sec überschritt oder der Patient aufgrund einer Progression der Erkrankung, nicht an dem Test teilnehmen konnte. Es wurden zwei Sensitivitätsanalysen ausgeführt, in denen ein oberes Limit von 45 sec bzw. der jeweils höchste Wert eingegeben wurde. Die Ergebnisse deuten an, dass der Behandlungseffekt von Ataluren vs. Placebo in der Primär-analyse unterschätzt wird (Tabelle 4-23).

Tabelle 4-23: Sensitivitätsanalysen der Änderung der TFTs bis Woche 48 für PTC124-GD-020-DMD; Ataluren versus Placebo (8)

Endpunkt (Zeit)	Maximum = 30 sec		Maximum = 45 sec		Maximum = Höchster Wert	
	Δ LSM, sec	p-Wert	Δ LSM, sec	p-Wert	Δ LSM, sec	p-Wert
10 m gehen/laufen	-1.07	0.117	-1.59	0.159	-2.26	0.196
4 Stufen hinauf	-1.43	0.058	-1.92	0.099	-2.78	0.326
4 Stufen hinab	-1.97	<b>0.012</b>	-2.44	<b>0.040</b>	-3.92	0.144

ANCOVA mit multipler Imputation; Faktoren: Stratifizierungsfaktoren (Alter, BL-6MWD-Gruppe, Dauer der Kortikosteroid-Vorbehandlung), Behandlung und BL-TFT.

In der Primäranalyse wurde ein Wert von 30 sec eingegeben, wenn die für Tests benötigte Zeit 30 sec überschritt oder der Patient aufgrund einer Progression der Erkrankung, nicht an dem Test teilnehmen konnte.

Quelle (Studienbericht 020): Tabelle 17, Tabelle 14.2.2.2.2, Tabelle 14.2.2.2.4, Tabelle 14.2.2.2.6, Tabelle 14.2.2.3.2, Tabelle 14.2.2.3.4, Tabelle 14.2.2.4.2, Tabelle 14.2.2.4.4, Tabelle 14.2.2.4.6

Wie für den primären Endpunkt wurden auch für die TFTs Meta-Analysen unter Einschluss der beiden Studien PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD durchgeführt, um die Präzision der Effektschätzung und die Ergebnissicherheit zu erhöhen.

*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.*

*Zeit, um 4 Treppenstufen hinaufzusteigen:*

Unter Einschluss der kITT Population und Verwendung der prädefinierten MMRM Analyse in der Studie PTC124-GD-007-DMD ergab sich in der Meta-Analyse mit der Studie PTC124-GD-020-DMD ein signifikanter Gesamteffektschätzer mit einer um -1.74 sec [95%KI: 2.97; -0.51] geringeren Zeit für Ataluren behandelte Patienten, um 4 Treppenstufen hinaufzusteigen (Abbildung 9). Es existierte kein Hinweis für eine Heterogenität zwischen den beiden Studien (p=0.45).

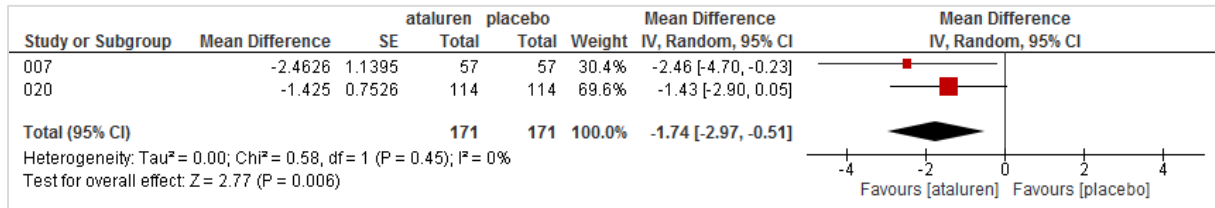


Abbildung 9: Meta-Analyse für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quelle:

PTC124-GD-007-DMD; prädefiniertes MMRM Modell, SAS output 4 stair climb test (10)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.3.2 CSR 020 (8)

Sensitivitätsanalysen:

In der Meta-Analyse ergab sich unter Einschluss der ITT Populationen der beiden Studien ein signifikanter Gesamteffektschätzer mit einer um -1.71 sec [95%KI: -2.95, -0.47] geringeren Zeit für Ataluren behandelte Patienten, um 4 Treppenstufen hinaufzusteigen (Abbildung 10). Es ergab sich kein Beleg für eine signifikante Heterogenität zwischen den beiden Studien (p=0.49).

Ein qualitativ und quantitativ vergleichbares Ergebnis wurde erzielt (-1.77 sec [95%KI: -3.00, -0.54]), wenn für die Meta-Analyse in der Studie PTC124-GD-007-DMD die kITT Population/das augmentierte MMRM Modell eingeschlossen wurde (Abbildung 11).

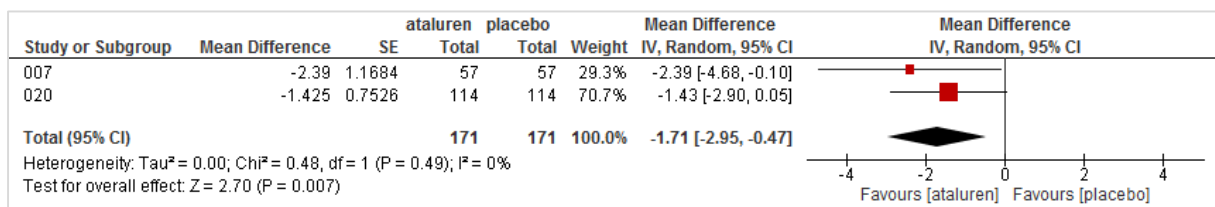


Abbildung 10: Meta-Analyse für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quelle:

PTC124-GD-007-DMD; prädefiniertes MMRM Modell, Tabelle 14.2.3.3.2B CSR 007 (9)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.3.2 CSR 020 (8)



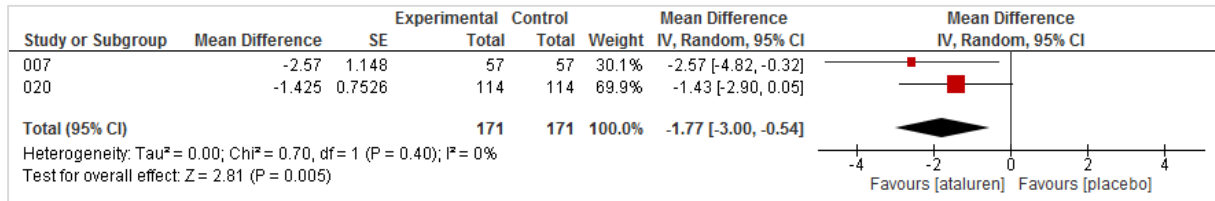


Abbildung 11: Meta-Analyse für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quelle:

PTC124-GD-007-DMD; augmentiertes MMRM Modell, Tabelle 14.2.3.3.2.2S CSR 007 (9)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.3.2 CSR 020 (8)

*Zeit, um 4 Treppenstufen hinabzusteigen:*

Unter Einschluss der kITT Population und Verwendung der prädefinierten MMRM Analyse in der Studie PTC124-GD-007-DMD ergab sich in der Meta-Analyse mit der Studie PTC124-GD-020-DMD ein signifikanter Gesamteffektschätzer mit einer um -1.89 sec [95%KI: -3.20, -0.59] geringeren Zeit für Ataluren behandelte Patienten, um 4 Treppenstufen hinabzusteigen (Abbildung 12). Es existierte kein Hinweis auf eine Heterogenität zwischen den beiden Studien ( $p=0.85$ ).

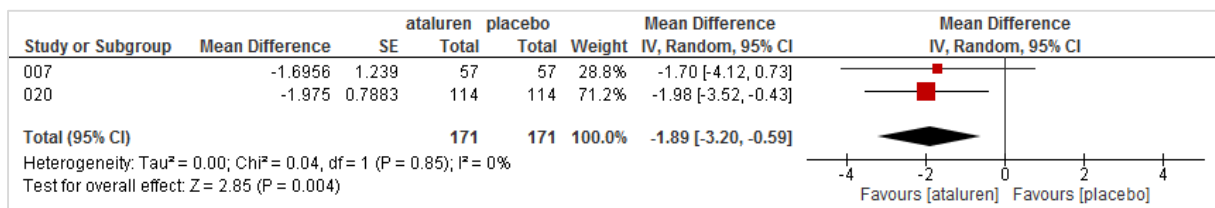


Abbildung 12: Meta-Analyse für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quelle:

PTC124-GD-007-DMD; prädefiniertes MMRM Modell, SAS output 4 stair descend test (11)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.4.2 CSR 020 (8)

Sensitivitätsanalysen:

In der Meta-Analyse der Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD ergab sich unter Einschluss der beiden ITT Populationen ein signifikanter Gesamteffektschätzer mit einer um -1.88 sec [95%KI: -3.19, -0.57] geringeren Zeit für Ataluren behandelte Patienten, um 4 Treppenstufen hinabzusteigen (Abbildung 13). Es ergab sich kein Hinweis für eine Heterogenität zwischen den beiden Studien ( $p=0.81$ ).

Ein qualitativ und quantitativ vergleichbares Ergebnis wurde erzielt (-1.91 sec [95%KI: -3.22, -0.60]), wenn für die Meta-Analyse in der Studie PTC124-GD-007-DMD die kITT Population/das augmentierte MMRM Modell eingeschlossen wurde (Abbildung 14).

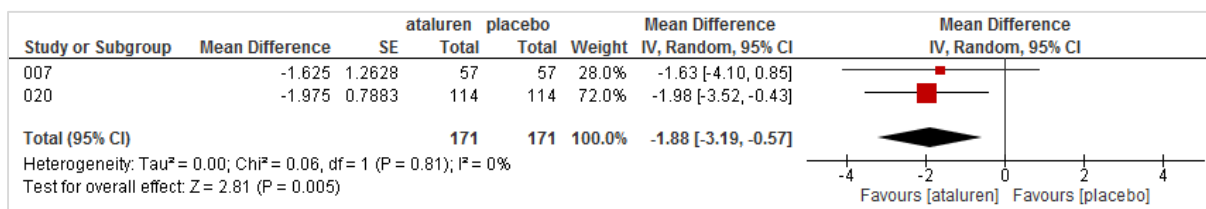


Abbildung 13: Meta-Analyse für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quelle:

PTC124-GD-007-DMD; prädefiniertes MMRM Modell, Tabelle 14.2.3.4.2B CSR 007 (9)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.4.2 CSR 020 (8)

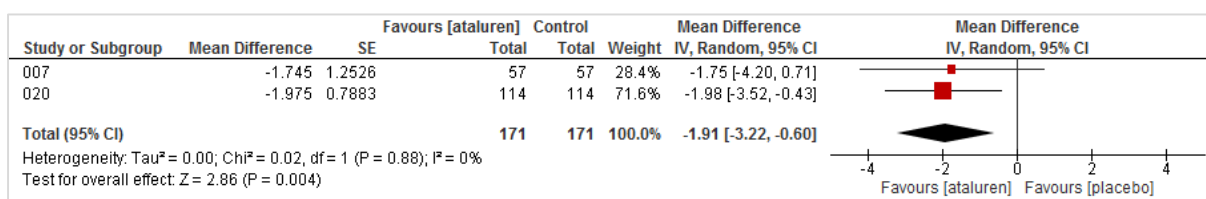


Abbildung 14: Meta-Analyse für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quelle:

PTC124-GD-007-DMD; augmentiertes MMRM Modell, Tabelle 14.2.3.4.2.2S CSR 007 (9)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.4.2 CSR 020 (8)

*Zeit, um 10 m zu gehen oder zu laufen:*

Unter Einschluss der kITT Population und Verwendung der prädefinierten MMRM Analyse in der Studie PTC124-GD-007-DMD ergab sich in der Meta-Analyse mit der Studie PTC124-GD-020-DMD ein signifikanter Gesamteffektschätzer mit einer um -1.16 sec [95%KI: -2.28; -0.03] geringeren Zeit für Ataluren behandelte Patienten, um 10 m zu gehen oder zu laufen (Abbildung 15). Es existierte kein Hinweis auf eine Heterogenität zwischen den beiden Studien (p=0.85).

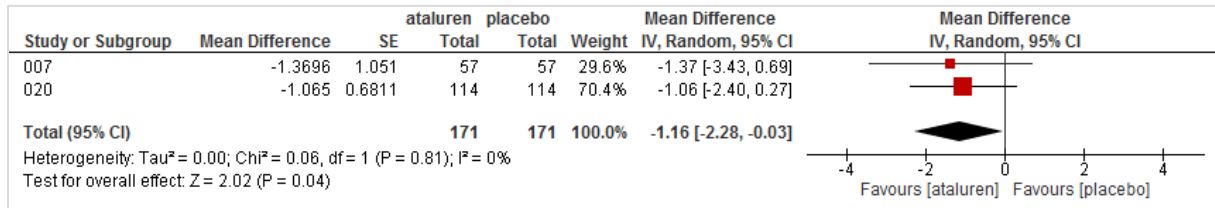


Abbildung 15: Meta-Analyse für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quellen: PTC124-GD-007-DMD; prädefiniertes MMRM Modell, SAS output 10 m run/walk test (13)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.2.2 CSR 020 (8)

### Sensitivitätsanalysen:

In der Meta-Analyse der Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD ergab sich unter Einschluss der ITT Populationen kein signifikanter Effekt, jedoch ein eindeutiger Trend ( $p=0.06$ ) mit einem Gesamteffektschätzer von einer um  $-1.08$  sec [95%KI:  $-2.22$ ;  $0.06$ ] geringeren Zeit für Ataluren behandelte Patienten, um 10 m zu gehen oder zu laufen (Abbildung 16). Es zeigte sich kein Hinweis auf Heterogenität zwischen den beiden Studien ( $p=0.97$ ).

Ein signifikantes Ergebnis wurde erzielt ( $-1.15$  sec [95%KI:  $-2.28$ ;  $-0.02$ ]), wenn für die Meta-Analyse in der Studie PTC124-GD-007-DMD die kITT Population/das augmentierte MMRM Modell eingeschlossen wurde (Abbildung 17).

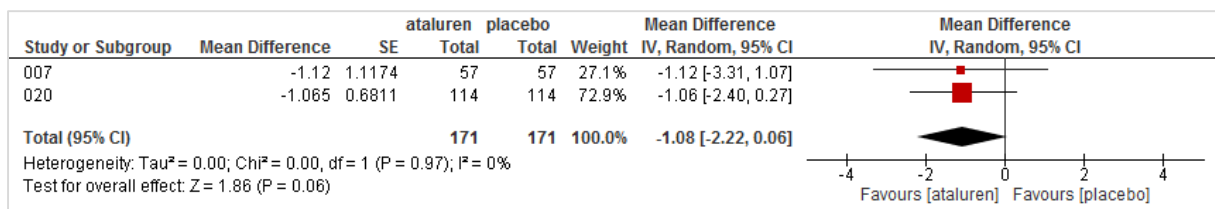


Abbildung 16: Meta-Analyse für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD (ITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quellen: PTC124-GD-007-DMD; prädefiniertes MMRM Modell, Tabelle 14.2.3.2.2B CSR 007 (9)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.2.2 CSR 020 (8)

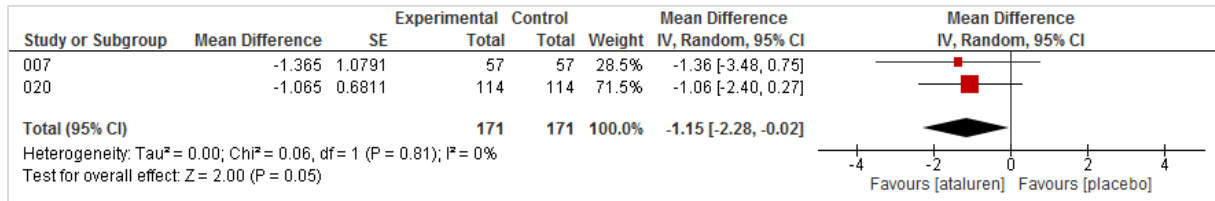


Abbildung 17: Meta-Analyse für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD (kITT) und PTC124-GD-020-DMD (ITT); Ataluren versus Placebo

Quellen: PTC124-GD-007-DMD; augmentiertes MMRM Modell, Tabelle 14.2.3.2.2.2S CSR 007 (9)

PTC124-GD-020-DMD; prädefiniertes ANCOVA Modell, Tabelle 14.2.2.2.2 CSR 020 (8)

In der Gesamtschau ergeben sich für die Meta-Analysen aller TFTs unter Einschluss der primären Analysepopulation (kITT) und der primären Analyseverfahren (prädefinierte MMRM Analyse) für die Studie PTC124-GD-007-DMD signifikante Gesamteffektschätzer. In keiner der Meta-Analysen zeigte sich ein Hinweis auf Heterogenität zwischen den Studien.

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.

#### 4.3.1.3.1.3 North Star Ambulatory Assessment (NSAA) – 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 4-24: Operationalisierung von Endpunkt NSAA („North Star Ambulatory Assessment“)

Studie	Operationalisierung
PTC124-GD-020-DMD	<p>Das NSAA ist ein Instrument zur Messung der körperlichen Funktion anhand von 17 Aktivitäten, die jeweils numerisch bewertet werden (0, 1, 2). Die Validität des NSAA wurde für Jungen mit DMD etabliert (50). Das NSAA wurde durch das „North Star Clinical Network“, welches durch die „Muscular Dystrophy Campaign“ unterstützt wird, und Elaine Scott entwickelt. Das den Prüfbüros zur Verfügung gestellte Manual wurde durch zusätzliche Photographien ergänzt und durch Elaine Scott mit Michelle Eagle und Anna Mayhew weiterentwickelt (8).</p> <p>Die Tests zu den Fähigkeiten, aus dem Liegen aufzustehen und 10 m zu gehen/laufen, waren nur einmal während des NSAA vorgesehen, sie wurden allerdings mittels zweier Skalen kategorisiert (0-2 für das NSAA, 1-6 für die TFTs). Die Patienten sollten während der Tests keine Schuhe oder Socken tragen und wenn möglich sollte keine Matte verwendet werden (8).</p> <p>Die Tests schlossen folgende Aktivitäten ein: 1. Stehen, 2. Gehen, 3. vom Stuhl aufstehen, 4. und 5. auf dem rechten oder linken Fuß stehen, 6. und 7. mit dem rechten oder linken Fuß zuerst auf eine Box steigen, 8. und 9. mit dem rechten oder linken Fuß von einer Box heruntersteigen, 10. aus Rückenlage in sitzende Position, 11. Aufstehen aus dem Liegen, 12. im Liegen Kopf heben, 13. auf den Fersen stehen, 14. Hüpfen, 15. und 16. mit dem rechten und linken Bein hüpfen und 17. gehen/laufen (10 m). Jede Aktivität wurde auf einer Skala von 0-2 bewertet (0 = unfähig, die Aktivität selbständig auszuführen, 1 = modifizierte Methode, aber fähig die Aktivität, selbständig ohne Hilfe eines Anderen auszuführen, 2 = normal, fähig, die Aktivität ohne jede Hilfe auszuführen). Das Manual enthielt detaillierte Vorgaben hinsichtlich der verbalen Instruktionen während der Tests (8).</p> <p>Die Tests zur Fähigkeit, aus der Rückenlage aufzustehen und 10 m zu gehen/laufen, konnten wiederholt werden. Die jeweils schnellste Zeit und der jeweils höchste Score sollten dokumentiert werden. Die Fähigkeit, aus dem Liegen aufzustehen, wurde wie folgt kategorisiert (TFT):</p> <p>1 = Unfähig, aus dem Liegen aufzustehen selbst bei Benutzung eines Stuhls 2 = Benötigt Möbel, um aus der Rückenlage in eine aufrechte Position aufzustehen („assisted</p>

Gowers“)

3 = Sich drehen, Aufrichten mittels Abstützen auf den Beinen, um aufrechte Position zu erreichen („full Gowers“)

4 = Sich drehen, Aufrichten mittels Abstützen mit einer Hand auf den Beinen („half Gowers“)

5 = Auf die Seite drehen und Aufstehen nach Abstützen mit einer oder beiden Händen auf dem Boden, aber nicht auf den Beinen

6 = Aufstehen ohne Drehen oder Abstützen auf den Beinen (8).

Während des Tests, um 10 m zu gehen oder zu laufen, wurde die Methode, die die Patienten einsetzten, dokumentiert:

1. Unfähig, selbständig zu gehen;
2. Unfähig, selbständig zu gehen, kann aber mit Unterstützung durch eine Person oder Hilfen gehen (Orthese oder Gehhilfe);
3. Ausgeprägt adaptierter Gang mit starker Lordose, Geschwindigkeit kann nicht erhöht werden;
4. Mäßig adaptierter Gang, Geschwindigkeit kann erhöht werden, laufen nicht möglich;
5. Geschwindigkeit kann erhöht werden, laufen nur mit Standphase möglich (d. h. nie beide Füße vom Boden abgehoben);
6. Laufen möglich, beide Füße vom Boden abgehoben (8).

Aus den 17 Einzelscores wurde ein Summenscore gebildet. Das NSAA wurde zur Screening-Visite, BL, alle 8 Wochen während der Behandlung und Woche 48 bzw. bei Behandlungsende erhoben (34).

Es wurde ein Wert von 0 eingegeben, wenn die Patienten nicht in der Lage waren, die Tests aufgrund von Krankheitsprogression/Verlust der Gehfähigkeit auszuführen (34).

Wenn weniger als 13 der 17 Datensätze vorlagen wurde der Summenscore als fehlend eingegeben. Falls zwischen 13 und 16 Datensätze vorlagen wurde der Summenscore aus diesen berechnet (34).

Der Gesamtscore wurde in der ITT-Population analog des primären Endpunkts ausgewertet. Eine Analyse nach linearer Transformation der NSAA-Scores auf einer Skala von 0-100 (0 = schlechtester bis 100 = bester Wert) war im SAP spezifiziert (34). Eine Transformation der Rohdaten des NSAA auf einer Skala von 0-100 wurde von Mathew et al. (2011, 2013) beschrieben und die transformierten Daten werden im Rahmen einer klinischen Studie als angemessener angesehen als die Rohdaten (50). Ein Test des NSAA (Kopf heben) erwies sich bei DMD-Patienten als weniger vertrauenswürdig und wurde in einer Studien bei DMD-Patienten aus dem Gesamtscore der transformierten NSAA-Daten ausgeschlossen (50,51). Entsprechend wurde bei der Analyse der Studie 020 verfahren.

*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 4-25: Bewertung des Verzerrungspotenzials für Endpunkt NSAA 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
PCT124-GD-020-DMD	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt NSAA wird unter Berücksichtigung der in Tabelle 4-25 aufgeführten Faktoren als niedrig eingeschätzt. Wie in Kapitel 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. Die Endpunkte wurden adäquat operationalisiert und den Endpunkterhebern wurde ein ausführliches Manual, in dem die Tests beschrieben waren, zur Verfügung gestellt. Das NSAA ist ein in DMD-Studien häufig eingesetztes, etabliertes und validiertes Instrument (24).

*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.*

Der NSAA wurde in einen linearen Gesamtscore von 0 (schlechtester Wert) bis 100 (bester Wert) transformiert. In der Gesamtpopulation (ITT) zeigte die Analyse der Änderung des Gesamtscores bis Woche 48 einen Vorteil für Ataluren, der allerdings weder in der Analyse der nicht transformierten noch der transformierten Daten signifikant war (Tabelle 4-26).

Tabelle 4-26: Änderung des NSAA von BL bis Woche 48 (ITT; n=114 Placebo, n=114) (8)

Gruppe	n BL/ Woche 48	M (±SD) BL (Score)	M (±SD) Woche 48 (Score)	Δ LSM (± SE) in Woche 48 (Score)	Δ LSM (±SE) vs. Placebo [95% KI] (Score)	p-Wert vs. Placebo
<i>Änderung des NSAA-Gesamtscores, nicht transformiert (Studienbericht 020: Tabelle 14.2.3.1, Tabelle 14.2.3.2)</i>						
Placebo	114/ 108	21.9 (8.03)	18.4 (9.95)	-3.76 (0.470)	-	-
Ataluren	114/ 105	22.2 (7.80)	20.1 (8.83)	-2.79 (0.468)	0.80 (0.523) [-0.23, 1.82]	0.128

Gruppe	n BL/ Woche 48	M ( $\pm$ SD) BL (Score)	M ( $\pm$ SD) Woche 48 (Score)	$\Delta$ LSM ( $\pm$ SE) in Woche 48 (Score)	$\Delta$ LSM ( $\pm$ SE) vs. Placebo [95% KI] (Score)	p-Wert vs. Placebo
<i>Änderung des transformierten (linearen) NSAA-Gesamtscores (Studienbericht 020: Tabelle 14.2.3.3, Tabelle 14.2.3.4)</i>						
Placebo	114/ 108	60.2 (18.37)	52.2 (23.76)	-8.49 (1.214)	-	-
Ataluren	114/ 106	60.9 (17.97)	55.7 (20.63)	-6.98 (1.212)	1.51 (1.357) [-1.16, 4.17]	0.268

ANCOVA mit multipler Imputation; Faktoren: Stratifizierungsfaktoren Alter, BL-6MWD-Gruppe, Kortikosteroid-Vorbehandlung, Behandlung, BL-NSAA Gesamtscore (nicht transformiert oder transformiert). Die Aktivität, den Kopf im Liegen zu heben, wurde in der Analyse der transformierten Werte nicht berücksichtigt. Bei Patienten, die ihre Gehfähigkeit verloren hatten, wurde ein Wert von 0 eingegeben.

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

Das NSAA wurde in der Studie 007 nicht erhoben und konnte daher nicht meta-analytisch ausgewertet werden.

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.

#### 4.3.1.3.1.4 Gesundheitsbezogene Lebensqualität (PRO) – 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 4-27: Operationalisierung von Endpunkt PODCI

Studie	Operationalisierung
PTC124-GD-020-DMD	<p>Die Gesundheitsbezogene Lebensqualität wurde mittels PODCI für die Domänen Transfer/Basismobilität und Sport/körperliche Funktion ermittelt (8,26).</p> <p>Die PODCI Transfer/Basismobilität-Domäne erfasst Routineaktivitäten des täglichen Lebens, während die PODCI Domäne Sport/körperliche Funktion etwas aktivere Freizeitaktivitäten abfragt. Jede Domäne wird auf einer Skala von 0-100 bewertet, wobei 100 der höchsten Aktivität und den geringsten Schmerzen entspricht (8).</p> <p>Das PODCI wurde zur Screening- und BL-Visite und alle 8 Wochen während der Behandlung jeweils vor den anderen Studienbedingten Prozeduren erhoben. Das PODCI war in allen für die Studie relevanten Sprachen verfügbar. Bei Patienten im Alter zwischen <math>\geq 7</math> und <math>&lt; 11</math> Jahren wurde der Fragebogen durch die Eltern oder Betreuer beantwortet, während bei älteren Patienten bis 16 Jahre der Patient und die Eltern/Betreuer die Fragen beantworteten. Wenn möglich sollte dasselbe Elternteil oder derselbe Betreuer die Fragen zu jeder Visite beantworten oder zumindest zur BL- und Woche 48-Visite (8).</p> <p>Das PODCI zeigte bei Patienten mit DMD Korrelationen mit der Krankheitsprogression und klinischen Parametern (26).</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 4-28: Bewertung des Verzerrungspotenzials für Endpunkt PODCI in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunktherheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
PTC124-GD-020-DMD	niedrig	ja	ja	ja	nein	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt PODCI wird unter Berücksichtigung der in Tabelle 4-28 aufgeführten Faktoren als niedrig eingeschätzt. Wie in Kapitel 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. Der Endpunkt wurde adäquat operationalisiert. Der PODCI ist ein bei DMD-Patienten validiertes Instrument und

korreliert mit der Krankheitsprogression, es ist allerdings unklar, ob der PODCI auf eine Behandlung anspricht (24). Es existiert kein formell validiertes Responsekriterium für das Instrument in der untersuchten Patientenpopulation.

*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.*

Die Behandlung mit Ataluren zeigt einen Trend zur Überlegenheit gegenüber Placebo auf die Einschätzung der Patienten/Betreuer hinsichtlich der Fähigkeiten, an üblichen Alltagsaktivitäten, sowie an Freizeitaktivitäten teilzunehmen, die Effekte waren aber weder in der Gesamtpopulation noch den untersuchten Subpopulationen signifikant.

Tabelle 4-29: Änderung des PODCI von BL bis Woche 48 (ITT; n=114 Placebo, n=114) (8)

Gruppe	n BL/ Woche 48	M (±SD) BL (Score)	M (±SD) Woche 48 (Score)	Δ LSM (± SE) in Woche 48 (Score)	Δ LSM (±SE) vs. Placebo, [95% KI] (Score)	p-Wert vs. Placebo
<i>Änderung des PODCI Transfer/Basismobilität (Studienbericht 020: Tabelle 14.2.4.1.1, Tabelle 14.2.4.1.2)</i>						
Placebo	114/110	81.4 (15.79)	72.4 (24.78)	-10.05 (1.681)	-	-
Ataluren	114/109	83.9 (13.10)	77.2 (22.16)	-8.41 (1.663)	1.64 (1.914) [-2.11, 5.39]	0.391
<i>Änderung des PODCI Sport/körperliche Funktion (Studienbericht 020: Tabelle 14.2.4.2.1, Tabelle 14.2.4.2.2)</i>						
Placebo	114/110	56.0 (20.94)	48.3 (26.04)	-8.66 (1.733)	-	-
Ataluren	114/109	56.2 (18.94)	51.1 (22.76)	-6.51 (1.736)	2.15 (1.990) [-1.75, 6.05]	0.281

ANCOVA mit multipler Imputation; unabhängige Variablen: Stratifizierungsfaktoren Alter, BL-6MWD-Gruppe, Kortikosteroid-Vorbehandlung, Behandlung, BL-PODCI.

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

Da das PODCI in der Studie 007 nicht erhoben wurde, konnte der Endpunkt nicht in der Meta-Analyse ausgewertet werden.

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.

#### **4.3.1.3.1.5 Aktivitäten des täglichen Lebens und Krankheitssymptome (PRO) – 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 4-30: Operationalisierung von Endpunkt Aktivitäten des täglichen Lebens und Krankheitssymptome (PRO)

Studie	Operationalisierung
PTC124-GD-020-DMD	<p>Zu jeder Visite wurden prospektiv Informationen zu Aktivitäten des täglichen Lebens und Krankheitssymptomen aus Sicht der Patienten/Betreuer nach einem von PTC Therapeutics spezifisch für DMD-Patienten entwickelten Testsystem erhoben (34). Das Instrument wurde ausgehend von Berichten von Teilnehmern früherer Ataluren-Studien entwickelt. Zur Screening- oder BL-Visite wurden die Patienten und/oder deren Betreuer gebeten, Alltagsaktivitäten (z.B. Gehen, Gleichgewicht, persönliche Hygiene/Körperpflege, An- und Ausziehen, selbständiges Essen, Badezimmernutzung, mit der Hand schreiben, Leistungen in der Schule oder Energieniveau) oder Symptome zu benennen, die durch die DMD eingeschränkt sind. Zu jeder der nachfolgenden Visiten wurde derselbe Patient und/oder dasselbe/derselbe Elternteil/Betreuer gebeten, Änderungen dieser Aktivitäten in den folgenden Kategorien zu benennen: Körperliche Funktion; allgemeines Energieniveau; Gedächtnis/Leistungen in der Schule; emotionale Ebene/soziale Funktion; und Schlaf. Die Änderungen wurden auf einer 5-Punkt Likert-Skala dokumentiert: 1 = viel schlechter, 2 = etwas schlechter, 3 = unverändert, 4 = etwas besser oder 5 = viel besser. Der Erhebungsbogen war in allen für die Studie relevanten Sprachen verfügbar und ein Mitarbeiter des Studienzentrums führte die Befragungen aus (8).</p> <p>Der Endpunkt konnte für die ITT- und andere relevante Populationen ausgewertet werden. Die statistische Auswertung war deskriptiv (34).</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 4-31: Bewertung des Verzerrungspotenzials für Endpunkt Aktivitäten des täglichen Lebens und Krankheitssymptome (PRO) 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-020-DMD	niedrig	ja	unklar	ja	nein	hoch

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt Alltagsaktivitäten und Krankheitssysteme wird unter Berücksichtigung der in Tabelle 4-31 aufgeführten Faktoren als hoch eingeschätzt. Wie in Kapitel 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. In den Analysen dieses Endpunkts blieb bei einigen der Subkategorien ein z.T. nicht unerheblicher Anteil der randomisierten Patienten unberücksichtigt und die Daten wurden rein deskriptiv ohne Imputation für fehlende Werte ausgewertet. Das ITT-Prinzip wurde dementsprechend wahrscheinlich nicht adäquat umgesetzt. Daher wird das Verzerrungspotential für diesen Endpunkt als hoch eingestuft. Die Patientenrelevanz dieses Endpunkts dürfte bei DMD-Patienten außer Frage stehen, die Sensitivität und Validität des Testsystems sind allerdings unbekannt.

*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.*

Hinsichtlich der Subkategorien Gehen (11% Differenz), Treppen steigen (5%), Aktivitäten der Selbstpflege (obere Extremitäten; 4%) und andere körperliche Funktionen (8%) ergaben sich in Woche 48 Vorteile zugunsten von Ataluren verglichen mit Placebo (Prozentwerte beziehen sich auf die Kategorien „Unverändert“, „Etwas besser“ und „Viel besser“). Die Ergebnisse der nicht-körperlichen Subkategorien waren variabel, allerdings berichteten mehr Patienten/Betreuer der Ataluren- als der Placebo-Gruppe über einen stabilen oder besseren Energiestatus. Die 6MWD-Subpopulationen wurden nicht analysiert.

Tabelle 4-32: Änderung der Alltagsaktivitäten und Krankheitssymptome in Woche 48 aus Sicht der Patienten/Betreuer (ITT) (8)

<b>Endpunkt / Änderung</b>	<b>Placebo (n=114)</b>	<b>Ataluren (n=114)</b>
<i>Aktivitäten der Selbstpflege (obere Extremitäten)</i>		
n	82	88
Viel besser	1 (1.2)	4 (4.5)
Etwas besser	7 (8.5)	7 (8.0)
Unverändert	67 (81.7)	73 (83.0)
Etwas schlechter	6 (7.3)	2 (2.3)
Viel Schlechter	1 (1.2)	2 (2.3)
<i>Gehen</i>		
n	112	108
Viel besser	5 (4.5)	8 (7.4)
Etwas besser	13 (11.6)	16 (14.8)
Unverändert	57 (50.9)	60 (55.6)
Etwas schlechter	19 (17.0)	21 (19.4)
Viel Schlechter	18 (16.1)	3 (2.8)

<b>Endpunkt / Änderung</b>	<b>Placebo (n=114)</b>	<b>Ataluren (n=114)</b>
<i>Treppen hinaufgehen</i>		
n	105	110
Viel besser	4 (3.8)	4 (3.6)
Etwas besser	8 (7.6)	13 (11.8)
Unverändert	61 (58.1)	65 (59.1)
Etwas schlechter	17 (16.2)	15 (13.6)
Viel Schlechter	15 (14.3)	13 (11.8)
<i>Andere körperliche Funktionen</i>		
n	65	68
Viel besser	4 (6.2)	5 (7.4)
Etwas besser	9 (13.8)	10 (14.7)
Unverändert	38 (58.5)	44 (64.7)
Etwas schlechter	7 (10.8)	8 (11.8)
Viel Schlechter	7 (10.8)	1 (1.5)
<i>Gedächtnis / soziale Funktion</i>		
n	92	104
Viel besser	6 (6.5)	5 (4.8)
Etwas besser	12 (13.0)	20 (19.2)
Unverändert	71 (77.2)	74 (71.2)
Etwas schlechter	3 (3.3)	3 (2.9)
Viel Schlechter	0 (0.0)	2 (1.9)
<i>Emotional / soziale Funktion</i>		
n	94	105
Viel besser	3 (3.2)	8 (7.6)
Etwas besser	9 (9.6)	21 (20.0)
Unverändert	75 (79.8)	68 (64.8)
Etwas schlechter	5 (5.3)	6 (5.7)
Viel Schlechter	2 (2.1)	2 (1.9)
<i>Energieniveau</i>		
n	83	89
Viel besser	3 (3.6)	8 (9.0)
Etwas besser	12 (14.5)	12 (13.5)
Unverändert	54 (65.1)	63 (70.8)
Etwas schlechter	11 (13.3)	5 (5.6)
Viel Schlechter	3 (3.6)	1 (1.1)
<i>Schlaf</i>		
n	85	91
Viel besser	3 (3.5)	3 (3.3)
Etwas besser	4 (4.7)	8 (8.8)
Unverändert	73 (85.9)	76 (83.5)
Etwas schlechter	5 (5.9)	2 (2.2)
Viel Schlechter	0 (0.0)	2 (2.2)

Endpunkt / Änderung	Placebo (n=114)	Ataluren (n=114)
<i>Andere nicht-körperliche Funktionen</i>		
n	23	23
Viel besser	0 (0.0)	1 (4.3)
Etwas besser	5 (21.7)	6 (26.1)
Unverändert	15 (65.2)	13 (56.5)
Etwas schlechter	2 (8.7)	3 (13.0)
Viel Schlechter	1 (4.3)	0 (0.0)

Studienbericht 020: Tabelle 14.2.5.1, 14.2.5.2

### Fazit

Für den Endpunkt liegen protokollgemäß keine Tests auf Signifikanz vor und der Endpunkt hat keinen Einfluss auf die Bewertung des Ausmaßes des Zusatznutzens von Ataluren.

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

Da der Endpunkt in der Studie 007 nicht erhoben wurde, konnte dieser nicht meta-analytisch ausgewertet werden.

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.

#### **4.3.1.3.1.6 Unerwünschte Ereignisse – 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 4-33: Operationalisierung des Endpunkts „Unerwünschte Ereignisse“ (UEs)

Studie	Operationalisierung
PTC124-GD-020-DMD	<p>Für die Auswertung der Verträglichkeit von Ataluren wurde die Safety Population („as treated“) herangezogen. In die Safety Population wurden alle randomisierten Patienten einbezogen, die Studienmedikation erhielten, die Auswertung erfolgte nach der tatsächlichen Zuteilung der Studienmedikation. Ein UE war als jedes unerwünschte medizinische Ereignis unabhängig vom Kausalzusammenhang definiert, welches bei einem Patienten auftrat, der die Prüfmedikation erhalten hatte (8). Alle UEs („treatment-emergent adverse events“, TEAE), die zwischen dem Tag der ersten Dosis und innerhalb von sechs Wochen (<math>\pm 7</math> d; Patienten, die nicht an der offenen Folgestudie teilnahmen) nach der letzten Dosis der Prüfpräparate auftraten oder sich verschlimmerten waren zu dokumentieren (8,34).</p> <p>Ein schwerwiegendes UE war definiert als jedes unerwünschte medizinische Ereignis unabhängig davon ob ein Kausalzusammenhang bestand oder nicht, welches fatal oder lebensbedrohlich war, eine stationäre Behandlung erforderte oder verlängerte oder zu dauerhafter oder signifikanter Behinderung oder substantiellen Einschränkungen, am normalen Leben teilzunehmen, die nicht mit der DMD assoziiert waren, führte. Wichtige medizinische Ereignisse, die nicht unmittelbar lebensbedrohlich oder fatal waren oder eine Hospitalisierung erforderten, den Patienten aber gefährdeten oder Interventionen erforderten, um eines der vorgenannten Ereignisse zu verhindern (z.B. Intensivbetreuung zuhause), sollten ebenfalls als schwerwiegend klassifiziert werden (8).</p> <p>UEs wurden nach MedDRA Version 15.1 kodiert und soweit möglich wurde deren Schwere entsprechend CTCAE (Version 3.0) klassifiziert (8,34). Falls keine CTCAE-Kriterien vorlagen, sollte der Prüfarzt protokollspezifische Kriterien zur Beurteilung des Schweregrads verwenden (Grad 1 = mild, Grad 2 = moderat, Grad 3 = schwer, Grad 4 = lebensbedrohlich, Grad 5 = fatal) (8). Die Inzidenz spezifischer UEs wurde entsprechend der Behandlungsgruppen, Systemorganklassen (SOCs) und „preferred terms“ (PTs) ausgewertet. Ein UE, welches bei einem Patienten mehrmals auftrat, wurde jeweils nur einmal gezählt und es wurde der jeweils schwerste CTCAE-Grad berücksichtigt (8,34).</p> <p>Während der Screening-Visite, vor der ersten Dosis, alle 8 Wochen während der Behandlung und zur Visite nach der letzten Behandlung sollten Blut- und Urinproben entnommen werden. Alle klinischen Laboruntersuchungen wurden in einem Zentrallabor ausgeführt. Diese umschlossen hämatologische und biochemische Parameter und Routine-Parameter im Urin. Hepatische, renale und adrenale UEs wurden besonders überwacht. Es wurden Maßnahmen spezifiziert, die bei spezifischen Laborabnormalitäten, zu ergreifen waren. Auch wiederkehrende oder andauernde Laborabnormalitäten des Grades 2 konnten unter bestimmten Bedingungen als klinisch signifikant eingestuft werden. Abweichungen sollten bis zum Abklingen oder zur Stabilisierung überwacht werden. Laborabweichungen, die keine klinische Intervention oder weitere Untersuchungen (abgesehen von Wiederholungs- oder Bestätigungs-Tests) erforderten, sollten nicht als UE gemeldet werden (8). Klinisch relevante renale und hepatische UEs wurden vor der Entblindung der Daten durch den medizinischen Monitor identifiziert (34). EKGs wurden zur Screening-Visite, in Woche 48 und nach Behandlungsende abgeleitet oder falls aus Sicht des Prüfarztes klinisch indiziert. Zur Screening-Visite und in Woche 48 wurden die Nieren und das ableitende System mittels Ultraschall untersucht. Zur Screening-Visite, in Woche 24 und 48 und nach Behandlungsende fand eine körperliche Untersuchung statt (8).</p> <p>Der Vergleich der Inzidenzraten der TEAEs in beiden Studienarmen wurde unter Verwendung der relativen Risiken (RR) und den zugehörigen 95% KIs dargestellt.</p>
PTC124-GD-007-DMD	Die Studie PTC124-GD-007-DMD wurde in einem separaten Dossier zur Nutzenbewertung dargestellt, auf welches hier verwiesen wird (Nutzenbewertungsverfahren zum Wirkstoff Ataluren, Beschlussfassung: 21.05.2015).

*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 4-34: Bewertung des Verzerrungspotenzials für Endpunkt UEs 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-020-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 4-34 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.*

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.

Alle randomisierten Patienten wurden in die Sicherheitsanalyse aufgenommen (n=115 pro Gruppe; „as-treated“). Die Mehrzahl der Patienten beider Gruppen entwickelte mindestens ein TEAE (87.7% Placebo, 89.6% Ataluren; Tabelle 4-35). Die TEAEs waren mehrheitlich milden oder mäßigen Schweregrads. Ein lebensbedrohliches oder fatales TEAE wurde nicht berichtet. Mehr Patienten der Ataluren- als der Placebo-Gruppe entwickelten ein TEAE mit

möglichem oder wahrscheinlichem Zusammenhang mit der Behandlung, die Unterschiede waren allerdings nicht signifikant. Ein Patient der Ataluren-Gruppe brach die Behandlung aufgrund einer Grad 2 Obstipation mit möglichem Zusammenhang mit der Behandlung vorzeitig ab und ein Patient der Placebo-Gruppe aufgrund einer Progression der Erkrankung ohne Kausalzusammenhang. Jeweils vier Patienten pro Gruppe entwickelten ein schwerwiegendes TEAE. Die Ergebnisse inklusive statistischer Auswertung sind in Tabelle 4-35 dargestellt.

Tabelle 4-35: Ergebnisse für TEAEs aus PTC124-GD-020-DMD („as-treated“); Ataluren versus Placebo

<b>TEAE</b>	<b>Placebo (n=115) n (%)</b>	<b>Ataluren (n=115) n (%)</b>	<b>RR Pla- cebo vs. Ataluren</b>	<b>Unteres 95% KI</b>	<b>Oberes 95% KI</b>
Mindestens 1 TEAE	101 (87.8)	103 (89.6)	0.9806	0.8941	1.0754
Schweregrad <sup>1,2</sup>					
Grad 1: Leicht	54 (47.0)	61 (53.0)	0.8852	0.6830	1.1475
Grad 2: Moderat	37 (32.2)	35 (30.4)	1.0571	0.7207	1.5507
Grad 3: Schwer	9 (7.8)	7 (6.1)	1.2857	0.4956	3.3356
Grad 4: Lebens- behrohlich	0	0	-	-	-
Zusammenhang mit Behandlung					
Kein	47 (40.9)	44 (38.3)	Nicht		
Unwahrscheinlich	30 (26.1)	30 (17.4)	analysiert		
Möglich	18 (15.7)	27 (23.5)	0.6667	0.3895	1.1411
Wahrscheinlich	6 (5.2)	12 (10.4)	0.5000	0.1943	1.2868
Abbruch aufgrund von TEAE	1 (0.9)	1 (0.9)	1.0000	0.0633	15.7964
Schwerwiegende TEAEs	4 (3.5)	4 (3.5)	1.0000	0.2562	3.9025
Todesfälle	0	0	-	-	-

Patienten, die dasselbe TEAE mehr als einmal entwickelten, wurden nur einmal gezählt.

<sup>1</sup> Die Angabe bezieht sich auf die schwerste Kategorie, die der Prüfarzt berichtet hatte.

<sup>2</sup> Ein TEAE (Skoliose) in der Placebo-Gruppe wurde inkorrekt als Grad 4 kodiert. Entsprechend des „Trial Master File“ war Grad 1 korrekt.

Quelle (Studienbericht 020): Tabelle 20

Das Nebenwirkungsprofil war in beiden Gruppen vergleichbar und die quantitative Analyse zeigte keine signifikanten Unterschiede hinsichtlich des Auftretens der häufigsten TEAEs (>5% der Patienten; Tabelle 4-35). Die häufigsten TEAEs waren in beiden Gruppen Erbrechen (20.4% gesamt) und Nasopharyngitis (20.0% gesamt), welches bei Kindern typische Erkrankungen sind. Diarrhoe trat in der Ataluren-Gruppe doppelt so häufig wie in der Placebo-Gruppe auf und Obstipation dreimal so häufig in der Placebo- wie in der Ataluren-Gruppe, die Unterschiede waren jedoch nicht signifikant. Die meisten TEAEs traten innerhalb der SOCs „Infektionen und parasitäre Erkrankungen“ (49.1% gesamt),

„Erkrankungen des Gastrointestinaltrakts“ (43.5%), „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ (30.0%), „Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen“ (27.8%), „Erkrankungen der Atemwege, des Brustraums und Mediastinums“ (27.8%) und „Erkrankungen des Nervensystems“ (22.2%) auf. Die Verteilung zwischen den SOC's war zwischen den Gruppen ausgeglichen. Eine Analyse der Inzidenzen der TEAEs pro 8-Wochen-Intervall ergab keine Hinweise auf eine mit der Zeit zunehmende Toxizität. In der Ataluren-Gruppe traten einige der gastrointestinalen TEAEs (Erbrechen, Diarrhoe und Abdominalschmerz) während der ersten 8 Wochen etwas häufiger auf.

Tabelle 4-36: Ergebnisse für TEAEs stratifiziert nach Organsystem, die bei >5% der Patienten der Gesamtpopulation über 48 Wochen auftraten aus PTC124-GD-020-DMD („as-treated“)

<b>MedDRA SOC / PT</b>	<b>Placebo (n=115) n (%)</b>	<b>Ataluren (n=115) n (%)</b>	<b>RR Pla- cebo vs. Ataluren</b>	<b>Unteres 95% KI</b>	<b>Oberes 95% KI</b>
<i>Erkrankungen des Gastrointestinaltrakts</i>	48 (41.7)	52 (45.2)	0.9231	0.6872	1.2400
Erbrechen	21 (18.3)	26 (22.6)	0.8077	0.4832	1.3500
Diarrhoe	10 (8.7)	20 (17.4)	0.5000	0.2449	1.0208
Schmerzen Oberbauch	13 (11.3)	9 (7.8)	1.4444	0.6428	3.2459
Übelkeit	7 (6.1)	7 (6.1)	1.0000	0.3623	2.7601
Obstipation	10 (8.7)	3 (2.6)	3.3333	0.9417	11.7994
Abdominalschmerz	5 (4.3)	7 (6.1)	0.7143	0.2335	2.1852
<i>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</i>	32 (27.8)	29 (25.2)	1.1034	0.7171	1.6979
Fieber	12 (10.4)	16 (13.9)	0.7500	0.3715	1.5140
Progression der Erkrankung <sup>1</sup>	14 (12.2)	9 (7.8)	1.5556	0.7014	3.4499
<i>Infektionen und parasitäre Erkrankungen</i>	50 (43.5)	63 (54.8)	0.7937	0.6080	1.0360
Nasopharyngitis	22 (19.1)	24 (20.9)	0.9167	0.5463	1.5381
Infektion der oberen Atemwege	6 (5.2)	11 (9.6)	0.5455	0.2087	1.4254
Rhinitis	4 (3.5)	8 (7.0)	0.5000	0.1549	1.6144
<i>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</i>	34 (29.6)	35 (30.4)	0.9714	0.6545	1.4418
Sturz	20 (17.4)	21 (18.3)	0.9524	0.5466	1.6593

<b>MedDRA SOC / PT</b>	<b>Placebo (n=115) n (%)</b>	<b>Ataluren (n=115) n (%)</b>	<b>RR Pla- cebo vs. Ataluren</b>	<b>Unteres 95% KI</b>	<b>Oberes 95% KI</b>
<i>Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen</i>	32 (27.8)	32 (27.8)	1.0000	0.6595	1.5163
Schmerz in einer Extremität	14 (12.2)	10 (8.7)	1.4000	0.6487	3.0214
Rückenschmerzen	8 (7.0)	11 (9.6)	0.7273	0.3037	1.7416
<i>Erkrankungen des Nervensystems</i>	23 (20.0)	28 (24.3)	0.8214	0.5046	1.3372
Kopfschmerz	21 (18.3)	21 (18.3)	1.0000	0.5788	1.7278
<i>Erkrankungen der Atemwege, des Brust- raums und Media- stinums</i>	30 (26.1)	34 (29.6)	0.8824	0.5813	1.3394
Husten	13 (11.3)	19 (16.5)	0.6842	0.3549	1.3191
Schmerzen im Oropharynx	6 (5.2)	7 (6.1)	0.8571	0.2972	2.4725

<sup>1</sup> Die Prüfarzte waren gebeten worden, bei Patienten, die während der Studie ihre Gehfähigkeit verloren, eine Progression der Erkrankung zu berichten.

Quelle (Studienbericht 020): Tabelle 22

Renale und hepatische TEAEs traten in der Ataluren-Gruppe etwas häufiger als in der Placebo-Gruppe auf, was hauptsächlich auf einer höheren Inzidenz des TEAEs „Hämaturie“ in der Ataluren-Gruppe beruhte (6.1% vs. 0.9%). Bei nur zwei Patienten (1.7%) der Ataluren-Gruppe wurde die Hämaturie (Grad 1 und 2) von den Prüfarzten als möglicherweise oder wahrscheinlich mit der Behandlung in Zusammenhang stehend klassifiziert. Alle renalen und hepatischen TEAEs waren mild mit Ausnahme eines schweren peripheren Ödems, einer mittelschweren Hämaturie und einer mittelschweren Harninkontinenz jeweils in der Ataluren-Gruppe. Zwei Patienten der Placebo-Gruppe entwickelten Nephrolithiasis. Die statistische Analyse erbrachte keine signifikanten Unterschiede zwischen den Gruppen (Tabelle 4-37). Die Urinanalyse und die Labordaten lieferten keine Hinweise auf eine renale Toxizität.

Tabelle 4-37: Ergebnisse für renale und hepatische TEAEs aus PTC124-GD-020-DMD („as-treated“); Ataluren versus Placebo

<b>MedDRA SOC / PT</b>	<b>Placebo (n=115) n (%)</b>	<b>Ataluren (n=115) n (%)</b>	<b>RR Pla- cebo vs. Ataluren</b>	<b>Unteres 95% KI</b>	<b>Oberes 95% KI</b>
Patienten mit $\geq 1$ TEAE	9 (7.9)	15 (13.0)	0.6000	0.2737	1.3154
<i>Allgemeine Erkrankun- gen und Beschwerden am Verabrei- chungsort</i>	0	3 (2.6)	0.0000	-	-
Ödem peripher	0	3 (2.6)	0.0000	-	-
<i>Leber- und Gallen- erkrankungen</i>	1 (0.9)	0	-	-	-
Steatosis hepatis	1 (0.9)	0	-	-	-
<i>Erkrankungen der Nieren und Harnwege</i>	8 (7.0)	12 (10.4)	0.6667	0.2831	1.5699
Hämaturie	1 (0.9)	7 (6.1)	0.1429	0.0179	1.1427
Anomalie des Urins	2 (1.7)	1 (0.9)	2.0000	0.1839	21.7517
Harninkontinenz	1 (0.9)	2 (1.7)	0.5000	0.0460	5.4379
Nephrolithiasis	2 (1.7)	0	-	-	-
Myoglobinuria	1 (0.9)	0	-	-	-
Pollakisurie	1 (0.9)	0	-	-	-
Polyurie	0	1 (0.9)	0.0000	-	-
Nierenzyste	0	1 (0.9)	0.0000	-	-

Quelle (Studienbericht 020): Tabelle 24

Mehr Patienten der Ataluren- als der Placebo-Gruppe entwickelten mindestens ein aus Sicht der Prüfarzte möglicherweise oder wahrscheinlich mit der Behandlung in Zusammenhang stehendes TEAE. Erbrechen war das einzige mit der Behandlung möglicherweise in Zusammenhang stehende TEAE, welches in der Ataluren-Gruppe signifikant häufiger auftrat als in der Placebo-Gruppe (dies bestätigte sich jedoch nicht in der nachfolgenden Meta-Analyse für die beiden eingeschlossenen Studien). Nur ein Patient (Ataluren) beendete die Behandlung aufgrund eines mit der Behandlung in Zusammenhang stehenden TEAEs (Grad 2 Obstipation). Erbrechen, Kopfschmerz, Diarrhoe und Übelkeit traten in der Ataluren-Gruppe mit einer >2.5-fach höheren Inzidenz auf als in der Placebo-Gruppe. Tabelle 4-38 zeigt die Inzidenz sowie eine statistische Auswertung der mit der Behandlung in Zusammenhang stehenden TEAEs.

Tabelle 4-38: Ergebnisse für die mit der Behandlung in Zusammenhang stehende TEAEs, die bei mehr als einem Patienten auftraten, aus PTC124-GD-020-DMD („as-treated“), Ataluren versus Placebo

<b>MedDRA SOC / PT</b>	<b>Placebo (n=115) n (%)</b>	<b>Ataluren (n=115) n (%)</b>	<b>RR Pla- cebo vs. Ataluren</b>	<b>Unteres 95% KI</b>	<b>Oberes 95% KI</b>
Patienten mit $\geq 1$ TEAE	24 (20.9)	39 (33.9)	0.6154	0.3972	0.9535
<i>Erkrankungen des Gastrointestinaltrakts</i>	7 (6.1)	19 (16.5)	0.3684	0.1611	0.8425
Erbrechen	1 (0.9)	9 (7.8)	0.1111	0.0143	0.8629
Abdominalschmerz	3 (2.6)	5 (4.3)	0.6000	0.1468	2.4523
Diarrhoe	1 (0.9)	5 (4.3)	0.2000	0.0237	1.6854
Übelkeit	0	4 (3.5)	0.0000	-	-
Flatulenz	1 (0.9)	2 (1.7)	0.5000	0.0460	5.4379
Abdominale Beschwerden	0	2 (1.7)	0.0000	-	-
<i>Allgemeine Erkran- kungen und Beschwer- den am Verabrei- chungsort</i>	2 (1.7)	1 (0.9)	2.0000	0.1839	21.7517
Fieber	1 (0.9)	1 (0.9)	1.0000	0.0633	15.7964
<i>Leber- und Gallen- erkrankungen</i>	2 (1.7)	0	-	-	-
Leberfunktion anomal	2 (1.7)	0	-	-	-
<i>Untersuchungen</i>	2 (1.7)	0	-	-	-
Renin erhöht	2 (1.7)	0	-	-	-
<i>Stoffwechsel- und Ernährungsstörungen</i>	6 (5.2)	8 (7.0)	0.7500	0.2687	2.0934
Hypertriglyzeridämie	2 (1.7)	3 (2.6)	0.6667	0.1135	3.9157
Appetit vermindert	0	2 (1.7)	0.0000	-	-
Hyperlipidämie	1 (0.9)	1 (0.9)	1.0000	0.0633	15.7964
<i>Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen</i>	2 (1.7)	4 (3.5)	0.5000	0.0934	2.6764
Schmerz in einer Extremität	1 (0.9)	2 (1.7)	0.5000	0.0460	5.4379
<i>Erkrankungen des Nervensystems</i>	3 (2.6)	7 (6.1)	0.4286	0.1136	1.6166
Kopfschmerz	3 (2.6)	6 (5.2)	0.5000	0.1281	1.9512
<i>Leber- und Gallen- erkrankungen</i>	2 (1.7)	5 (4.3)	0.4000	0.0792	2.0199
Hämaturie	0	2 (1.7)	0.0000	-	-



<b>MedDRA SOC / PT</b>	<b>Placebo (n=115) n (%)</b>	<b>Ataluren (n=115) n (%)</b>	<b>RR Pla- cebo vs. Ataluren</b>	<b>Unteres 95% KI</b>	<b>Oberes 95% KI</b>
<i>Erkrankungen der Atemwege, des Brust- raums und Media- stinums</i>	1 (0.9)	1 (0.9)	1.0000	0.0633	15.7964
Epistaxis	1 (0.9)	1 (0.9)	1.0000	0.0633	15.7964
<i>Erkrankungen der Haut und des Unter- hautzellgewebes</i>	2 (1.7)	2 (1.7)	1.0000	0.1433	6.9788
Erythematöser Hautausschlag	0	2 (1.7)	0.0000	-	-

Quelle (Studienbericht 020): Tabelle 26

Die TEAEs waren mehrheitlich mild (Grad 1) oder mittelschwer (Grad 2). Sieben Prozent aller Patienten entwickelten ein schweres (Grad 3) TEAE (7.8% Placebo, 6.1% Ataluren). Hinsichtlich der schweren (Grad 3) TEAEs ergab sich kein klares Muster. Die Prüfarztze waren angehalten, bei Patienten, die während der Studie ihre Gehfähigkeit verloren, eine Grad 3 Progression der Erkrankung anzugeben. Eine Progression der Erkrankung war das häufigste Grad 3 TEAE und wurde in der Placebo-Gruppe etwas häufiger berichtet als in der Ataluren-Gruppe (5.2% vs. 1.7%). Keines der anderen schweren TEAEs trat bei mehr als einem Patienten auf. In Tabelle 4-39 ist die Inzidenz und eine statistische Analyse der schweren TEAEs dargestellt.

Tabelle 4-39: Inzidenz schwerer (Grad 3) TEAEs aus PTC124-GD-020-DMD („as-treated“), Ataluren versus Placebo

<b>MedDRA SOK / PT</b>	<b>Placebo (n=115) n (%)</b>	<b>Ataluren (n=115) n (%)</b>	<b>RR Pla- cebo vs. Ataluren</b>	<b>Unteres 95% KI</b>	<b>Oberes 95% KI</b>
<i>Erkrankungen des Blutes und des Lymphsystems</i>	0	1 (0.9)	0.0000	-	-
Neutropenie	0	1 (0.9)	0.0000	-	-
<i>Erkrankungen des Gastrointestinaltrakts</i>	1 (0.9)	2 (1.7)	0.5000	0.0460	5.4379
Obstipation	0	1 (0.9)	0.0000	-	-
Diarrhoe	0	1 (0.9)	0.0000	-	-
Erbrechen	1 (0.9)	0	-	-	-

<b>MedDRA SOK / PT</b>	<b>Placebo (n=115) n (%)</b>	<b>Ataluren (n=115) n (%)</b>	<b>RR Pla- cebo vs. Ataluren</b>	<b>Unteres 95% KI</b>	<b>Oberes 95% KI</b>
<i>Allgemeine Erkrankun- gen und Beschwer- den am Verabrei- chungsort</i>	7 (6.1)	4 (3.5)	1.7500	0.5266	5.8160
Progression der Erkrankung	6 (5.2)	2 (1.7)	3.0000	0.6183	14.5550
Abasie	0	1 (0.9)	0.0000	-	-
Ödem peripher	0	1 (0.9)	0.0000	-	-
Fieber	0	1 (0.9)	0.0000	-	-
Gangstörung	1 (0.9)	0	-	-	-
<i>Infektionen und parasitäre Erkran- kungen</i>	0	1 (0.9)	0.0000	-	-
Virale Gastroenteritis	0	1 (0.9)	0.0000	-	-
<i>Verletzung, Vergiftung und durch Eingriffe bedingte Komplika- tionen</i>	0	2 (1.7)	0.0000	-	-
Sturz	0	1 (0.9)	0.0000	-	-
Posttraumatische Schmerzen	0	1 (0.9)	0.0000	-	-
<i>Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen</i>	3 (2.6)	1 (0.9)	3.0000	0.3167	28.4178
Schmerz in einer Extremität	0	1 (0.9)	0.0000	-	-
Rückenschmerzen	1 (0.9)	0	-	-	-
Muskelschwäche	1 (0.9)	0	-	-	-
Skoliose	0	0	-	-	-

Die Schweregrade wurden durch die Prüfarzte mittels CTCAE oder der protokollgemäßen Definition bestimmt.  
Quelle (Studienbericht 020): Tabelle 23

Jeweils vier Patienten der Ataluren- und Placebo-Gruppe entwickelten ein schwerwiegendes TEAE und bei vier Patienten trat mehr als ein schwerwiegendes TEAE auf. Alle schwerwiegenden TEAEs bei den mit Ataluren behandelten Patienten wurden als nicht mit der Behandlung in Zusammenhang stehend klassifiziert. Ein Fall einer anomalen Leberfunktion bei einem Patienten der Placebo-Gruppe stand nach Auffassung des Prüfarztes möglicherweise oder wahrscheinlich mit der Behandlung in Zusammenhang (Tabelle 4-40).

Tabelle 4-40: Inzidenz schwerwiegender TEAEs aus PTC124-GD-020-DMD („as-treated“)

MedDRA PT	Zusammenhang mit Behandlung	Gruppe
Sehnenerkrankung	Kein Zusammenhang	Ataluren
Posttraumatische Schmerzen	Kein Zusammenhang	Ataluren <sup>1</sup>
Pneumonie	Kein Zusammenhang	
Bronchiolitis	Kein Zusammenhang	Ataluren <sup>1</sup>
Pneumonie	Kein Zusammenhang	
Hypertrophie der Rachen- mandeln	Kein Zusammenhang	Ataluren <sup>1</sup>
Hypertrophie der Nasen- muscheln	Kein Zusammenhang	
Leberfunktion anomal	Möglich	Placebo
Myokarditis	Kein Zusammenhang	Placebo
Gastroenteritis	Kein Zusammenhang	Placebo
Oberschenkelfraktur	Kein Zusammenhang	Placebo <sup>1</sup>
Fraktur einer unteren Extremität	Kein Zusammenhang	

<sup>1</sup> Zwei Patienten entwickelten zwei schwerwiegende TEAEs, die als separate Fälle berichtet wurden und zwei Patienten zwei schwerwiegende TEAEs, die als ein Fall berichtet wurden.

Quelle (Studienreport 020): Tabelle 27

### Laboruntersuchungen

Ataluren hatte keinen klinisch relevanten Einfluss auf die hämatologischen Parameter. Änderungen von normalen BL-Werten zu Werten außerhalb des Normbereichs bei der Follow-up-Visite traten nur gelegentlich auf und waren zwischen den Gruppen gleich verteilt.

Erhöhte Serumspiegel der Kreatinkinase, Asparat- und Alanin-Aminotransferase und Laktat-Dehydrogenase wurden vor und während der Behandlung beobachtet. Diese sind durch den Austritt der Enzyme aus dem geschädigten Muskel erklärbar und sind bei DMD zu erwarten. Klinisch relevante Änderungen der Gamma-Glutamyl-Transferase, alkalischen Phosphatase und des Gesamtbilirubins wurden nicht beobachtet. Bei zwei Patienten der Placebo-Gruppe wurde aus Sicherheitsgründen im Protokoll definierten Schwellenwert für die Gamma-Glutamyl-Transferase überschritten.

Als Folge der Muskelschädigung waren die Kreatininspiegel niedrig und die Konzentrationen von Kreatinin, Blut-Harnstoff-Stickstoff und Cystatin C nahmen unter Ataluren geringfügig zu. Bei keinem Patienten wurden die aus Sicherheitsgründen im Protokoll definierten Schwellenwerte für Serum-Kreatinin, -Cystatin C und Blut-Harnstoff-Stickstoff überschritten.

Die mittleren Konzentrationen von Gesamtcholesterin und Triglyceriden waren bereits zu Beginn der Studie im oberen Normbereich, was bei DMD-Patienten zu erwarten ist. Während der Studie nahmen LDL (0.285 vs. 0.002 mmol/l Placebo) und Cholesterin (0.340 vs. 0.008 mmol/l Placebo) in der Ataluren-Gruppe geringfügig allerdings nicht in einem klinisch

relevanten Ausmaß zu. Das mittlere HDL änderte sich in der Placebo-(0.015 mmol/l) und Ataluren-Gruppe (0.340 mmol/l) nur minimal. Die mittleren Serum-Triglycerid-Konzentrationen nahmen bei den mit Ataluren behandelten Patienten im Studienverlauf geringfügig zu (0.114 vs. 0.047 mmol/l Placebo). Bei mehr Patienten der Ataluren- als der Placebo-Gruppe stiegen das Gesamtcholesterin (17.4% vs. 6.5%), LDL (17.4% vs. 13.3%), und die Triglyceride (18.3% vs. 13.9%) von initial normalen zu hohen Werten in Woche 48 an. HDL sank bei mehr Patienten der Ataluren- als der Placebo-Gruppe von initial normalen Werten zu niedrigen Werten in Woche 48 ab (4.6% vs. 1.9%). Keiner der mit Ataluren behandelten Patienten entwickelte eine Grad 3 oder höher Hypercholesterolämie oder Hypertriglyceridämie.

Die mittleren Änderungen anderer biochemischer Parameter waren klein und klinisch unbedeutend. Bei nur wenigen Patienten wurden Elektrolytstörungen beobachtet und diese waren nicht klinisch relevant. Klinisch signifikante Behandlungsspezifischen Laborabweichungen wurden bei keinem der hämatologischen und biochemischen Tests beobachtet.

Die Urinanalysen lieferten keine Hinweise auf Sicherheitsrisiken. Die im Protokoll aus Sicherheitsgründen definierten Schwellenwerte wurden hinsichtlich des Protein-Kreatinin-Verhältnisses ( $>0.4$  mg:mg) und des Verhältnisses von Protein zur Osmolalität ( $>0.3$  mg/l:mOsm/kg) im Harn bei 3.6% vs. 1.8% und 7.1% vs. 8.0% der Patienten der Placebo- und Ataluren-Gruppe überschritten. Bei zwei Patienten (1.8%) der Placebo-Gruppe wurden geringe Mengen Blut im Urin detektiert. Bei zwei (1.8%) bzw. 10 Patienten (8.8%) der Ataluren-Gruppe wurden moderate bzw. große Mengen Blut im Urin nachgewiesen. Mit Ausnahme eines Patienten, der zu Beginn und während der Studie eine Hämaturie hatte, wurden diese Anomalitäten nur einmal beobachtet und waren bei Wiederholung negativ.

#### *Vitalzeichen, körperliche Befunde und andere Sicherheitsbefunde*

Bei etwa 34% der Patienten war der Blutdruck bereits zu Studienbeginn erhöht. Der systolische und diastolische Blutdruck stieg in der Placebo- (0.6/1.5 mmHg) und Ataluren-Gruppe (0.4/2.0 mmHg) während der Studie nur minimal und nicht klinisch relevant an. Der Anteil der Patienten, die an Prä- oder Hypertension litten, war zu Beginn und in Woche 48 ähnlich. Zwei Patienten der Placebo- und ein Patient der Ataluren-Gruppe entwickelten einen leichten (Grad 1) Bluthochdruck.

Das mittlere Körpergewicht nahm in der Placebo-Gruppe um 4.3 kg und in der Ataluren-Gruppe um 3.0 kg zu.

Bei den körperlichen Untersuchungen wurden keine auffälligen Befunde erhoben.

Bei zwei der mit Placebo behandelten Patienten (1.8%) waren die EKGs zu Studienbeginn klinisch auffällig und bei einem Patienten der Placebo- (1.8%) und zwei Patienten der Ataluren-Gruppe (1.8%) in Woche 48. Bei beiden Patienten der Ataluren-Gruppe schrieb der behandelnde Arzt die klinisch signifikante EKG-Abweichung einer Rechtsherzhypertrophie zu.

Die renalen Ultraschalluntersuchungen zeigten bei wenigen Patienten zu Studienbeginn und/oder in Woche 48 Auffälligkeiten. Bis auf einen Patienten der Ataluren-Gruppe, dessen Milz in Woche 48 vergrößert war, wurde keine der Abweichungen von den Prüfarzten als klinisch signifikant eingeschätzt. Die vergrößerte Milz war Grad 1 (mild) und stand nicht mit der Behandlung in Zusammenhang. Die renalen Laborparameter waren bei diesem Patienten unauffällig.

### Fazit

Ataluren wurde grundsätzlich in der Studie PTC124-GD-020-DMD gut vertragen. Insgesamt war das Nebenwirkungsprofil von Ataluren hinsichtlich Inzidenz, Art, Schwere und möglichem Kausalzusammenhang demjenigen unter Placebo vergleichbar, allerdings trat unter den TEAEs mit zumindest möglichem Kausalzusammenhang Erbrechen unter Ataluren häufiger auf als unter Placebo. Diese Beobachtung wurde jedoch in der nachfolgenden Meta-Analyse der beiden eingeschlossenen Studien nicht bestätigt. Die häufigsten TEAEs waren in beiden Gruppen Erbrechen (20.4% gesamt) und Nasopharyngitis (20.0% gesamt). Es ergaben sich keine Hinweise auf eine renale oder hepatische Toxizität, einschließlich der im Protokoll spezifizierten Überwachung renaler und hepatischer Laborparameter. Hinsichtlich der Inzidenz schwerer (Grad 3) TEAEs ergaben sich keine signifikanten Unterschiede, allerdings wurde eine Progression der Erkrankung in der Placebo-Gruppe numerisch häufiger berichtet als in der Ataluren-Gruppe. Es traten nur sehr wenige schwerwiegende TEAEs auf und für keines der schwerwiegenden TEAEs wurde ein Kausalzusammenhang mit Ataluren vermutet.

Da bislang in Deutschland keine Referenztherapie bei DMD zugelassen ist, ist es nicht möglich, einen Zusatznutzen von Ataluren im Sinne einer Vermeidung von (schwerwiegenden) Nebenwirkungen oder deren Ausmaß zu belegen. Es ergeben sich jedoch keine Hinweise auf ein relevantes Schadenspotential von Ataluren, insbesondere in Anbetracht der besonderen Schwere der Erkrankung. Vor diesem Hintergrund ist das aussergewöhnlich gute Verträglichkeitsprofil von Ataluren besonders hervorzuheben.

*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.*

### Ergebnisse der Meta-Analysen für Inzidenzraten der UE aus PTC124-GD-007-DMD und PTC124-GD-020-DMD

Da die beiden Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD vergleichbare Patientenpopulationen eingeschlossen hatten, den beiden Studien dasselbe Studiendesign zugrund lag, und die Dosierung von Ataluren in beiden Studien identisch war, konnten die Ergebnisse der Einzelstudien sinnvoll in einer Meta-Analyse zusammengefasst werden. Meta-Analysen wurden für die Inzidenzraten der UEs (PT) bzw. UEs in den jeweiligen SOC's durchgeführt, die bei der Entscheidungsfindung für oder gegen den Einsatz der zu evaluierenden Intervention einen wesentlichen Stellenwert haben, und für die nicht unmittelbar erkennbar war, ob zwischen den beiden Studienarmen ein signifikanter Unterschied besteht (Abbildung 21 bis Abbildung 47). Die Inzidenzraten stammten für beide eingeschlossenen Studien aus der Safety Population („as treated“).

In den Meta-Analysen der einzelnen UEs, bzw. der zusammengefassten UEs in den entsprechenden SOC's zeigte sich bis auf zwei Ausnahmen zwischen den Studien keine Heterogenität. In keiner der Meta-Analysen konnte ein signifikanter Unterschied der Inzidenz von unerwünschten Ereignissen in den beiden Studienarmen nachgewiesen werden.

In der Meta-Analyse der UE „Abdominalbeschwerden“ (Abbildung 29) und der SOC „Infektionen und parasitäre Erkrankungen“ (Abbildung 38) war eine signifikante Heterogenität der Ergebnisse zwischen den Studien erkennbar. In beiden Fällen waren die Effektschätzer nicht gleichgerichtet, in keiner der beiden eingeschlossenen Studien waren die Effekte in den Einzelstudien signifikant, noch waren die Gesamteffektschätzer signifikant.

Auf Basis der Meta-Analysen der beiden eingeschlossenen Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD ist für die Intervention mit Ataluren somit kein Schaden gegenüber einer Placebobehandlung erkennbar. Die Ergebnisse aller Meta-Analysen für diesen Endpunkt sind tabellarisch in Kapitel 4.4.2.5.3 zusammengefasst.

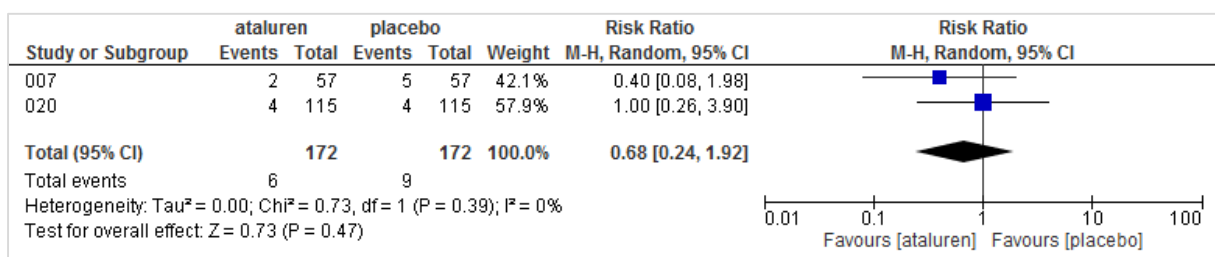


Abbildung 21: Meta-Analyse für schwerwiegende UE aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

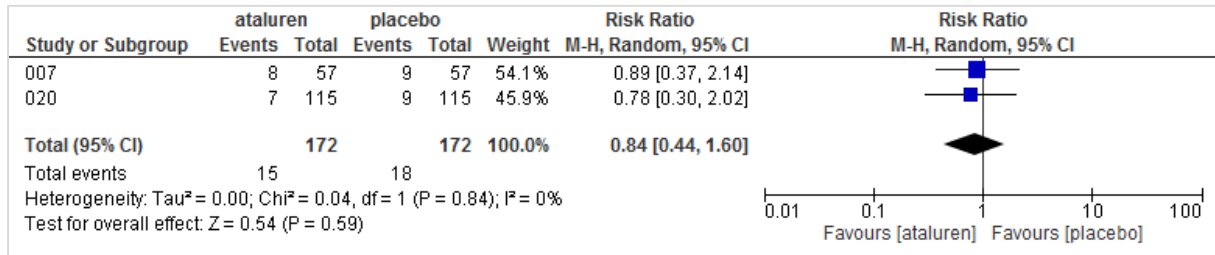


Abbildung 22: Meta-Analyse für schwere UE (Grad ≥ 3) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

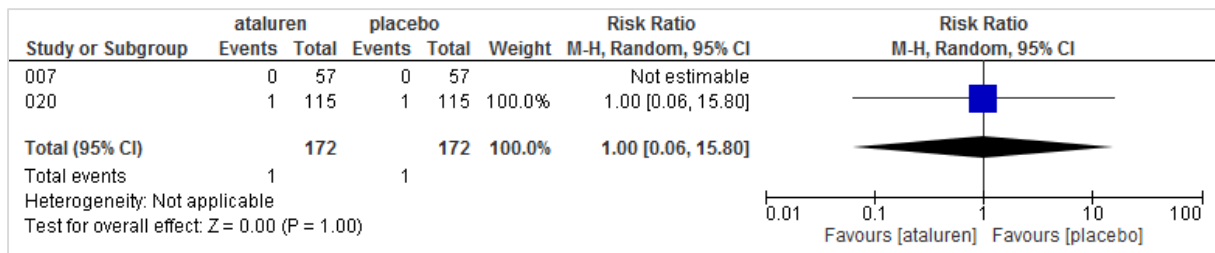


Abbildung 23: Meta-Analyse für UE, die zum Studienabbruch führten aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

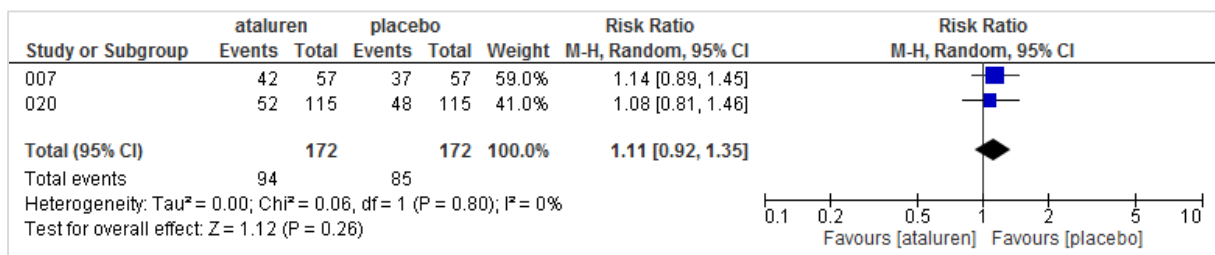


Abbildung 24: Meta-Analyse für UE der SOC „Erkrankungen des Gastrointestinaltrakts“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

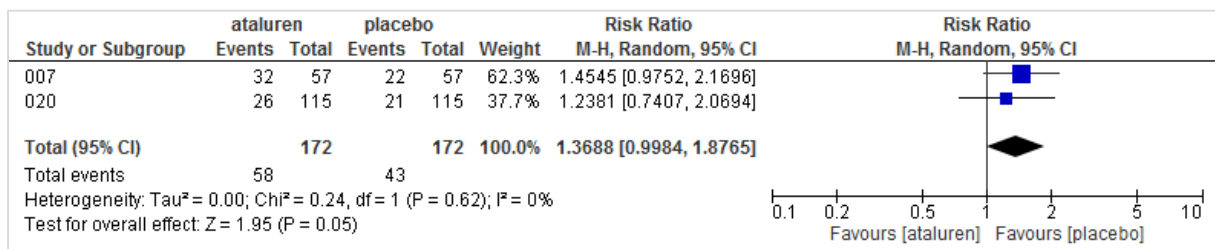


Abbildung 25: Meta-Analyse für UE „Erbrechen“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

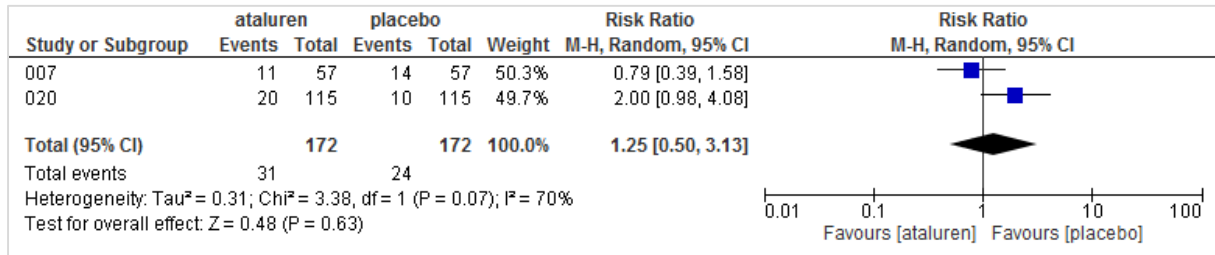


Abbildung 26: Meta-Analyse für UE „Diarrhöe“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

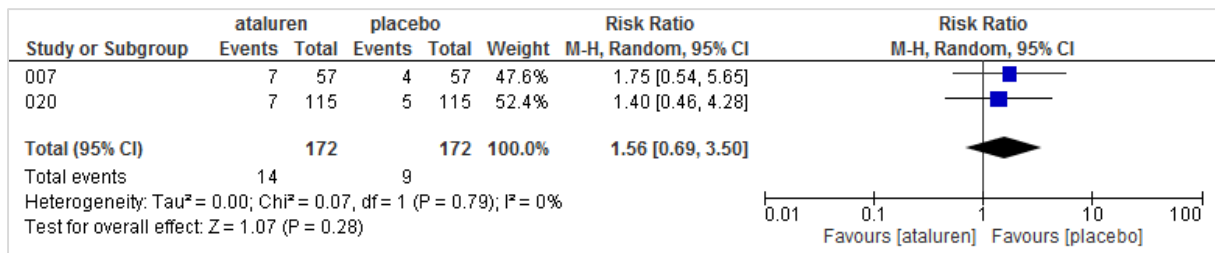


Abbildung 27: Meta-Analyse für UE „Abdominalschmerz“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

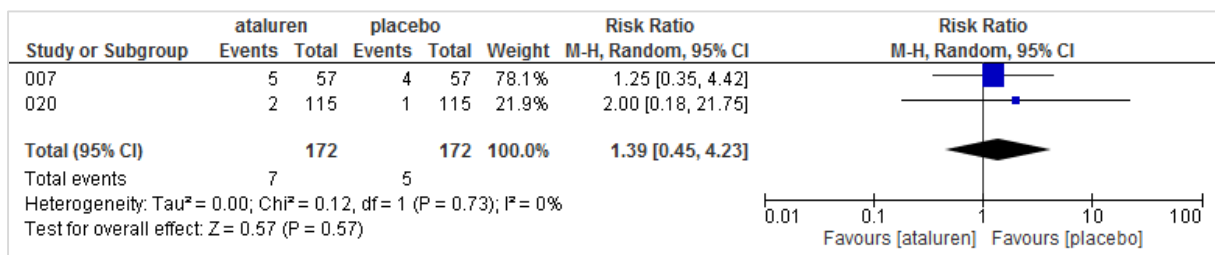


Abbildung 28: Meta-Analyse für UE „Flatulenz“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

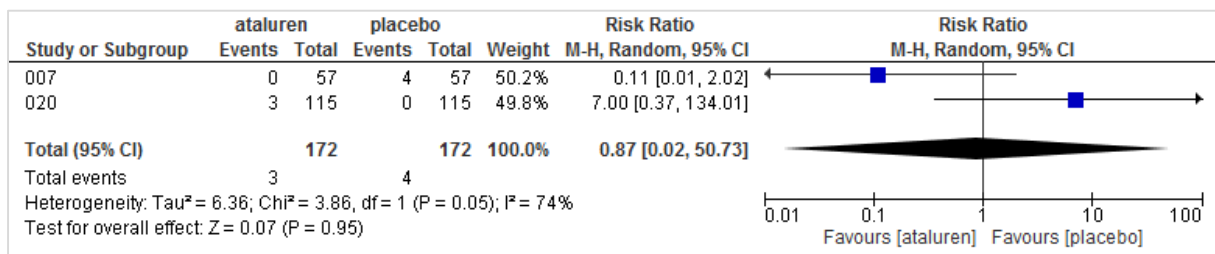


Abbildung 29: Meta-Analyse für UE „Abdominalbeschwerden“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo



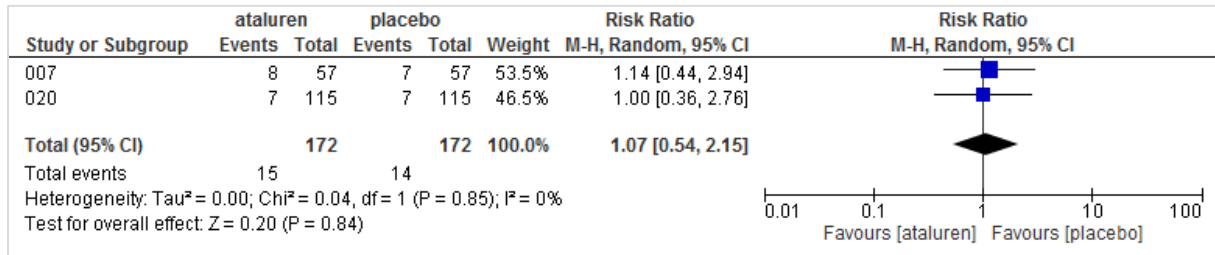


Abbildung 30: Meta-Analyse für UE „Nausea“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

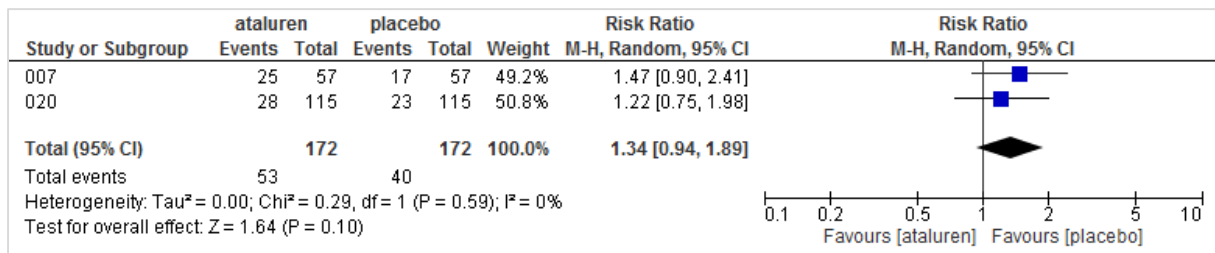


Abbildung 31: Meta-Analyse für UE der SOC „Erkrankungen des Nervensystems“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

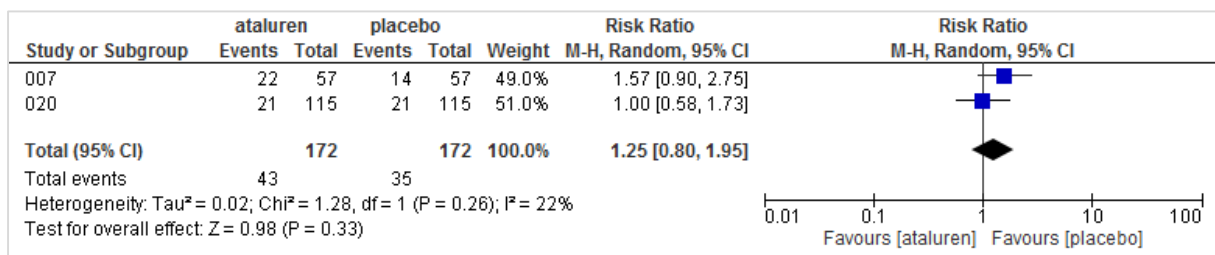


Abbildung 32: Meta-Analyse für UE „Kopfschmerzen“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

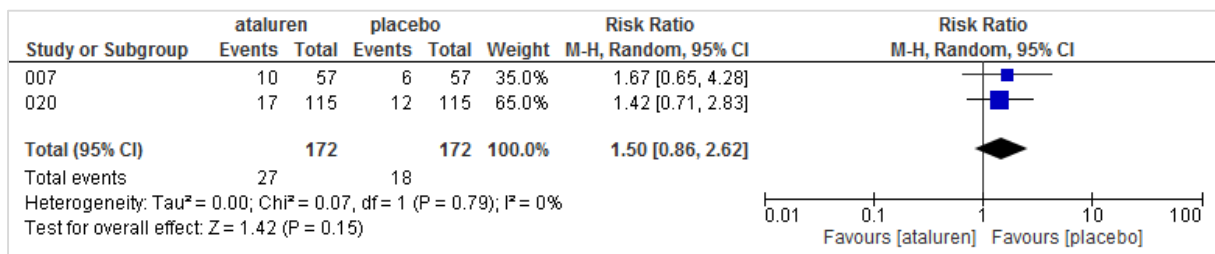


Abbildung 33: Meta-Analyse für UE der SOC „Erkrankungen der Niere und der Harnwege“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

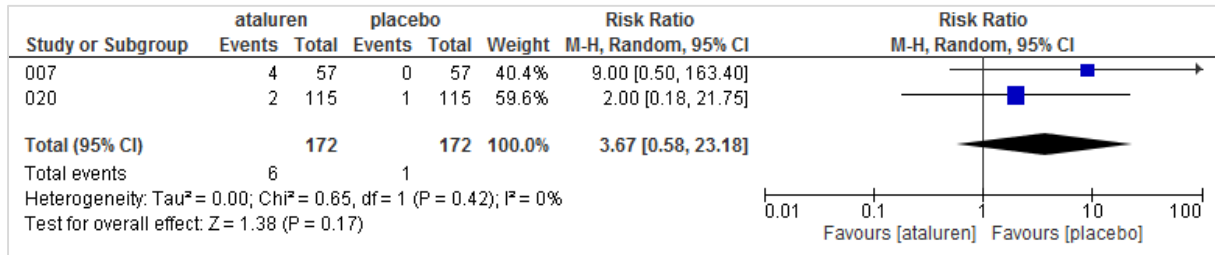


Abbildung 34: Meta-Analyse für UE „Enuresis“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

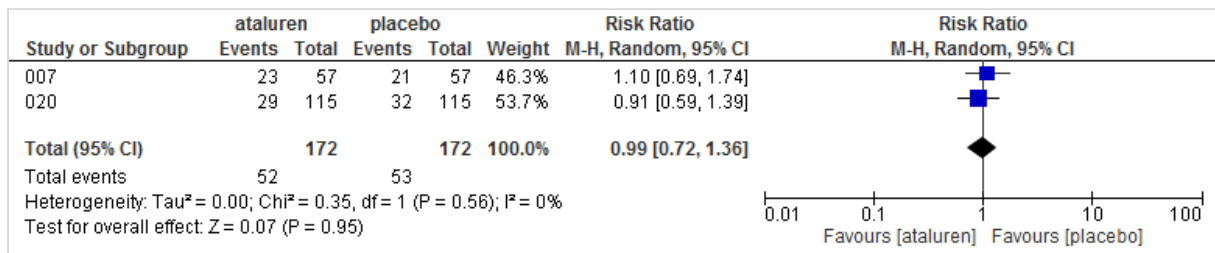


Abbildung 35: Meta-Analyse für UE der SOC „Allgemeine Erkrankungen und Beschwerden am Verabreichungsort“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

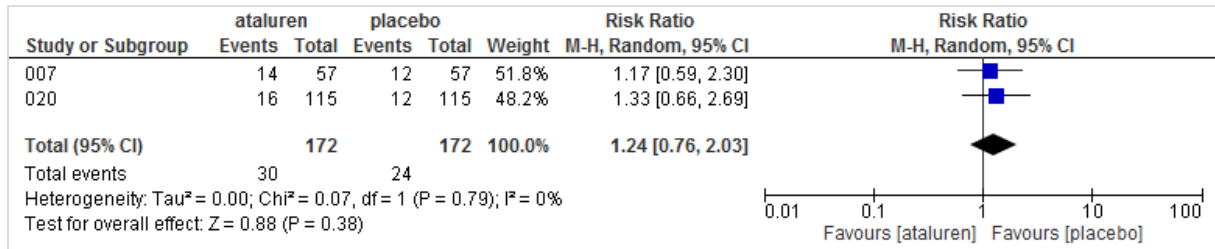


Abbildung 36: Meta-Analyse für UE „Fieber“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

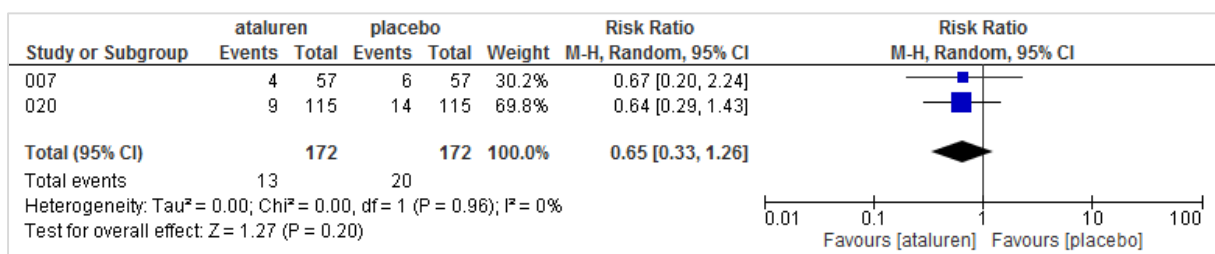


Abbildung 37: Meta-Analyse für UE „Progression einer Erkrankung“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

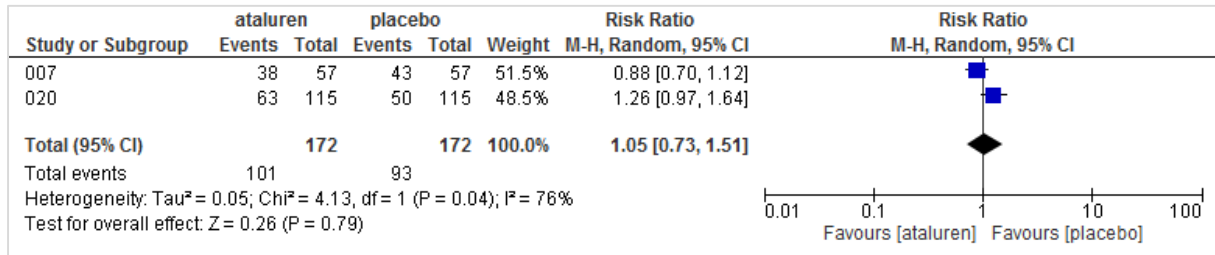


Abbildung 38: Meta-Analyse für UE der SOC „Infektionen und parasitäre Erkrankungen“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

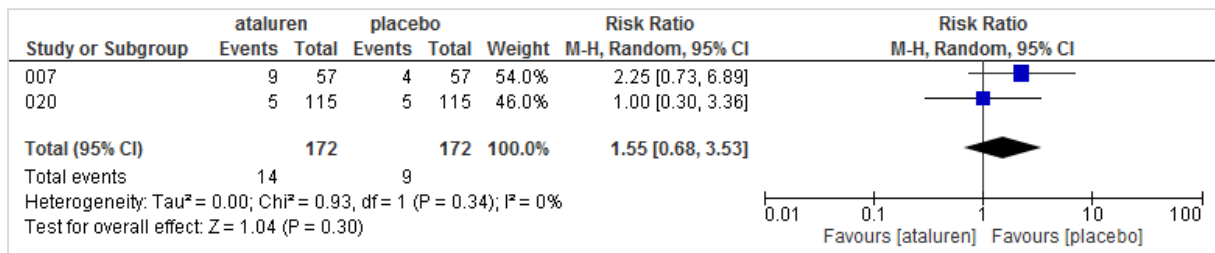


Abbildung 39: Meta-Analyse für UE „Gastroenteritis“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

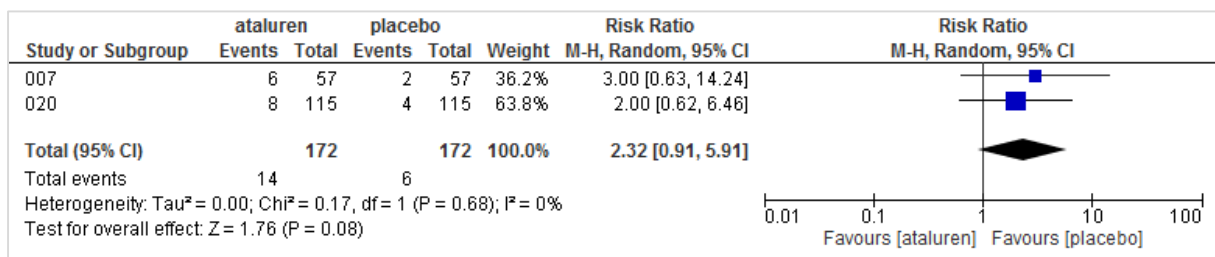


Abbildung 40: Meta-Analyse für UE „Rhinitis“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

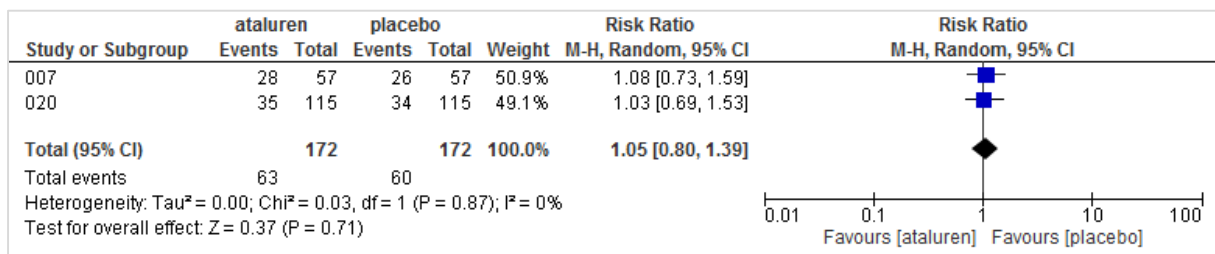


Abbildung 41: Meta-Analyse für UE der SOC „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

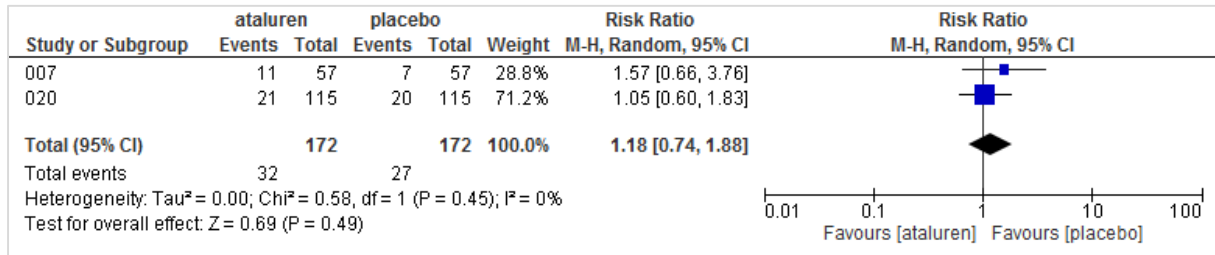


Abbildung 42: Meta-Analyse für UE „Sturz“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

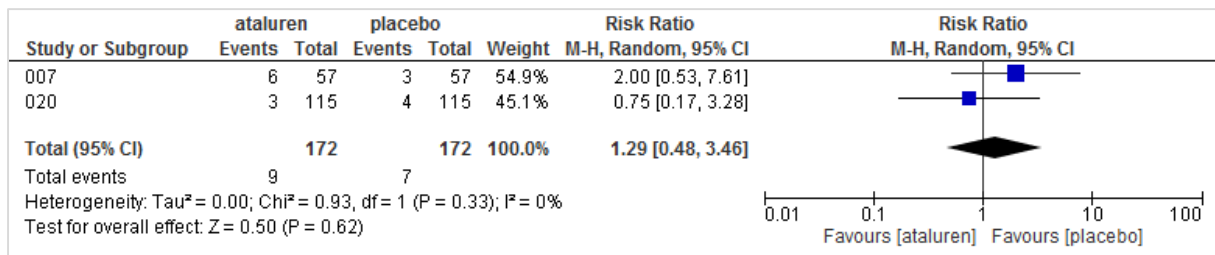


Abbildung 43: Meta-Analyse für UE „Kontusion“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

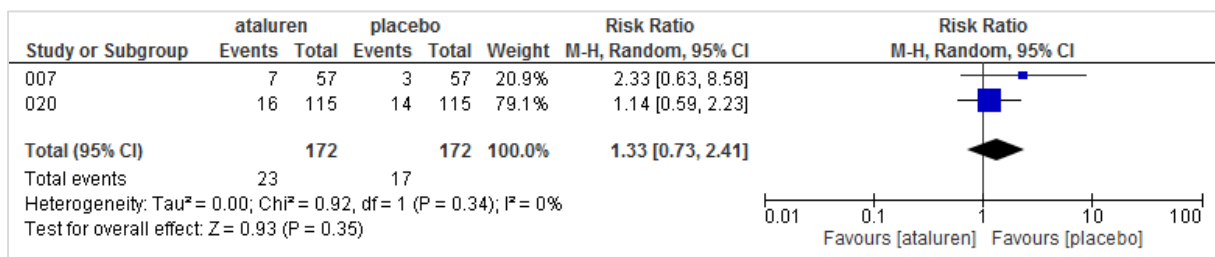


Abbildung 44: Meta-Analyse für UE der SOC „Stoffwechsel- und Ernährungsstörungen“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

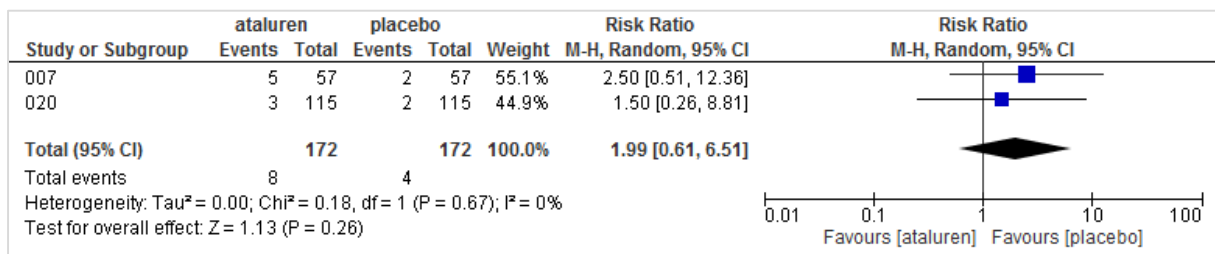


Abbildung 45: Meta-Analyse für UE „Appetit vermindert“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

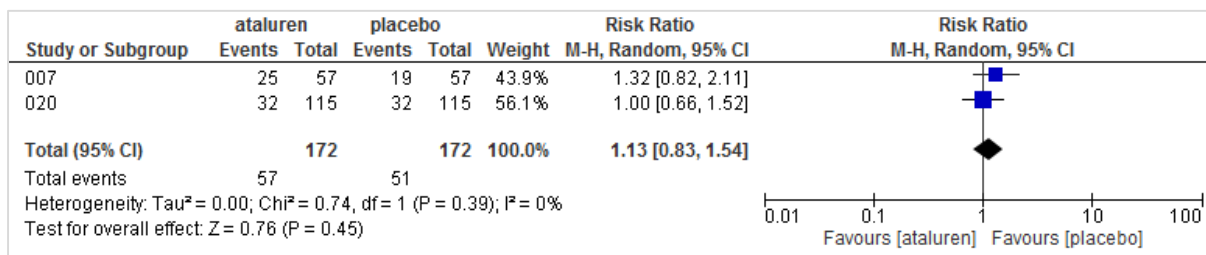


Abbildung 46: Meta-Analyse für UE der SOC „Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen“ aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

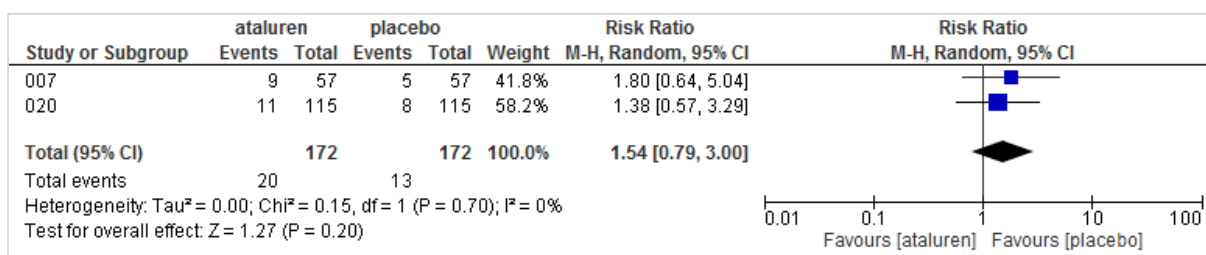


Abbildung 47: Meta-Analyse für UE „Rückenschmerzen“ (PT) aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

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.

#### 4.3.1.3.2 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 4.3.1.3.1.1.

*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.*

#### **4.3.1.3.2.1 Gehstrecke im 6MWT - RCT**

##### **Studie PTC124-GD-020-DMD**

Alle nachfolgend dargestellten Subgruppen-Analysen der Studie PTC124-GD-020-DMD waren prospektiv im SAP definiert (34), siehe auch Kapitel 4.2.5.5.

Die Interaktionstests (s. Tabelle 4-41) ergaben in der Studie PTC124-GD-020-DMD für die Stratifizierungsfaktoren „Dauer der Kortikosteroid-Vorbehandlung“ ( $\geq 12$  bzw.  $\geq 6 - < 12$  Monate;  $p=0.486$ ) und „BL-6MWD-Gruppe“ ( $\geq 350$  bzw.  $< 350$  m;  $p=0.482$ ) keine Hinweise oder Belege für unterschiedliche Effekte zwischen den Subgruppen ( $p>0.2$ ). Hinsichtlich der Subgruppen nach Alter ( $< 9$  Jahre vs.  $\geq 9$  Jahre) lieferte der Interaktionstest einen Hinweis ( $p=0.076$ ) auf unterschiedliche Effekte zwischen den Gruppen (7). Der globale statistische Test auf Interaktion zwischen den 6MWD-Subpopulationen ( $< 300$ ,  $\geq 300$  bis  $< 400$ ,  $\geq 400$  m) war signifikant ( $p=0.05$ ). Eine Bestätigung der Effektmodifikation in einer Meta-Analyse mit der zweiten Studie PTC124-GD-007-DMD war somit erforderlich, um die Effektmodifikation mit einer höheren Ergebnissicherheit zu bestätigen (siehe weiter unten, Meta-Analysen).

Ataluren zeigte sich in der Subpopulation mit einer BL-6MWD von  $\geq 300$  bis  $< 400$  m gegenüber Placebo hoch signifikant überlegen (Tabelle 4-41). Der Unterschied der Gehstrecke im 6MWT (LSM) zugunsten von Ataluren verglichen mit Placebo betrug 42.89 m [95% KI 11.75; 74.03],  $p=0.007$  und war deutlich größer als in der Gesamtpopulation (12.98 m; ITT). Bei Patienten mit einer initialen Gehstrecke unter 300 m nahm die Gehfähigkeit über 48 Woche unabhängig von der Behandlung deutlich ab, während bei den Patienten mit einer Gehstrecke von 400 m und mehr die Gehstrecke während der Behandlung weitgehend stabil blieb.

Tabelle 4-41: Änderung der 6MWD über 48 Wochen in prä-spezifizierten Subgruppen für PTC124-GD-020-DMD (ITT) - Interaktionstests

Gruppe	n BL/ Woche 48	M ( $\pm$ SD) BL (m)	M ( $\pm$ SD) Woche 48 (m)	$\Delta$ LSM ( $\pm$ SE) in Woche 48 (m) *	$\Delta$ LSM ( $\pm$ SE) vs. Placebo [95% KI] (m)	p-Wert vs. Placebo	p-Wert Inter- aktion	
<i>BL-6MWD &lt;300 m (Studienreport 020: Tabelle 19, Tabelle 14.2.1.2.2.2, Tabelle 14.2.1.3.3.2)</i>								
Placebo	21/19	233.50 (51.145)	96.78 (115.190)	-119.87 (18.837)			<b>0.050</b>	
Ataluren	24/22	256.85 (39.756)	129.15 (112.295)	-127.58 (17.027)	-7.71 (24.092) [- 54.93, 39.51]	0.749		
<i>BL-6MWD <math>\geq</math>300 m bis &lt;400 m (Studienreport 020: Tabelle 19, Tabelle 14.2.1.2.2.2, Tabelle 14.2.1.3.3.2)</i>								
Placebo	52/50	354.52 (29.712)	277.96 (120.719)	-69.93 (12.120)				
Ataluren	47/54	356.69 (26.672)	328.87 (70.306)	-27.04 (12.550)	<b>42.89</b> (15.886) [11.75, 74.03]	<b>0.007</b>		
<i>BL-6MWD <math>\geq</math>400 m (Studienreport 020: Tabelle 19, Tabelle 14.2.1.2.2.2, Tabelle 14.2.1.3.3.2)</i>								
Placebo	41/40	441.58 (31.583)	441.78 (51.490)	1.10 (12.829)				
Ataluren	43/42	434.96 (27.722)	426.93 (65.982)	-8.40 (12.821)	-9.51 (17.187) [-43.19, 24.18]	0.580		
<i>BL-6MWD &lt;350 m (Studienreport Tabelle 18, Tabelle 14.2.1.2.2.1, Tabelle 14.2.1.3.3.1)</i>								
Placebo	41/39	276.68 (58.416)	152.96 (120.325)	-116.75 (13.822)			0.482	
Ataluren	41/38	285.64 (46.554)	189.34 (115.672)	-94.86 (13.466)	21.89 (17.457) [-12.32, 56.11]	0.210		
<i>BL-6MWD <math>\geq</math>350 m (Studienreport Tabelle 18, Tabelle 14.2.1.2.2.1, Tabelle 14.2.1.3.3.1)</i>								
Placebo	73/70	412.32 (41.960)	392.04 (99.494)	-19.05 (10.117)				
Ataluren	73/71	409.87 (37.975)	399.67 (70.853)	-11.08 (10.313)	7.98 (13.014) [-17.53, 33.48]	0.540		

Gruppe	n BL/ Woche 48	M ( $\pm$ SD) BL (m)	M ( $\pm$ SD) Woche 48 (m)	$\Delta$ LSM ( $\pm$ SE) in Woche 48 (m) *	$\Delta$ LSM ( $\pm$ SE) vs. Placebo [95% KI] (m)	p-Wert vs. Placebo	p-Wert Inter- aktion	
<i>Alter &lt;9 Jahre (Studienreport 020: Tabelle 18, Tabelle 14.2.1.2.4, Tabelle 14.2.1.3.5)</i>								
Placebo	53/50	380.21 (70.580)	339.96 (136.017)	-47.79 (12.284)			0.076	
Ataluren	57/55	379.59 (69.922)	364.47 (119.561)	-31.53 (11.952)	16.26 (15.029) [-13.20, 45.71]	0.279		
<i>Alter <math>\geq</math>9 Jahre (Studienreport 020: Tabelle 18, Tabelle 14.2.1.2.4, Tabelle 14.2.1.3.5)</i>								
Placebo	61/59	349.05 (87.530)	278.14 (168.912)	-73.37 (12.007)				
Ataluren	57/54	350.79 (72.990)	287.52 (137.926)	-63.46 (12.168)	9.92 (14.489) [-18.48, 38.31]	0.494		
<i>Kortikoid-Vorbehandlung <math>\geq</math>6 - &lt;12 Monate (Studienreport 020: Tabelle 18, Tabelle 14.2.1.2.3, Tabelle 14.2.1.3.4)</i>								
Placebo	18/18	398.82 (45.145)	383.82 (75.461)	-42.36 (19.542)			0.486	
Ataluren	19/19	377.03 (71.930)	344.76 (144.428)	-45.48 (18.906)	-3.12 (25.636) [-53.37, 47.12]	0.903		
<i>Kortikoid-Vorbehandlung <math>\geq</math>12 Monate (Studienreport 020: Tabelle 18, Tabelle 14.2.1.2.3, Tabelle 14.2.1.3.4)</i>								
Placebo				-65.93 (8.379)				
Ataluren				-49.84 (8.357)	16.09 (11.425) [-6.30, 38.49]	0.159		

\* ANCOVA-Modell mit multipler Imputation

Post-hoc wurden weitere Sensitivitätsanalysen in verschiedenen BL-6MWD-Gruppen ausgeführt. Die Analysen nach BL-MWD von <250,  $\geq$ 250 bis <400 und  $\geq$ 400 m (p=0.254 Interaktionstest in Woche 48), <200,  $\geq$ 200 bis <400 und  $\geq$ 400 m (p=0.268 Interaktionstest in Woche 48), <300,  $\geq$ 300 bis <450 und >450 m (p=0.236 Interaktionstest in Woche 48) lieferten keine Hinweise auf zwischen den Subgruppen unterschiedliche Effekte im 6MWT. Dennoch zeigen die Analysen, dass die Patienten, die hinsichtlich ihrer Gehfähigkeit in einem mittleren Bereich der Krankheitsprogression liegen, von einer Therapie mit Ataluren



profitieren bzw. ein Behandlungseffekt bei diesen Patienten über eine Studiendauer von einem Jahr nachweisbar ist. Bei Patienten mit einer Gehstrecke zwischen  $\geq 200$  bis  $< 400$  m betrug der Unterschied zwischen Ataluren und Placebo in Woche 48 26.60 m (LSM, 95% KI -0.01, 53.22,  $p=0.050$ ), bei Patienten mit einer Gehstrecke von  $\geq 250$  bis  $< 400$  m 29.46 m (LSM, 95% KI 2.03, 56.88,  $p=0.035$ ) und bei Patienten mit einer Gehstrecke von  $\geq 300$  bis  $< 450$  m 24.39 (LSM, 95% KI -0.04, 48.83,  $p=0.05$ ) (8).

Die meisten der Patienten, die ihre Gehfähigkeit während der Studie verloren ( $n=9$  Ataluren,  $n=14$  Placebo), gehörten der Subpopulation mit dem schlechtesten Gehvermögen zu Beginn der Studie an (BL-6MWD  $< 300$  m). Keiner der Patienten der Ataluren-Gruppe, die zu Beginn der Studie  $\geq 300$  m im 6MWT bewältigen konnten, wurde während der Studie Rollstuhlpflichtig; in der Placebo-Gruppe waren es vier Patienten.

### Meta-Analysen der Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD

Die Auswertung von *a priori* geplanten Subgruppen in Meta-Analysen wurde durchgeführt, wenn ein Nachweis auf signifikante Interaktion zwischen den Subgruppen vorlag, und die Ergebnissicherheit der Subgruppeneffekte als hoch eingestuft wurde (39). Nur die Subgruppen basierend auf dem Effektmodifikator BL-6MWD ( $\geq 300$  m bis  $< 400$  m;  $< 300$  m;  $\geq 400$  m) qualifizierten für eine weitere Auswertung in Meta-Analysen (siehe Kapitel 4.2.5.5).

Die o. a. Subgruppen wurden daher in Meta-Analysen der beiden Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD eingeschlossen. Für die Studie PTC124-GD-007-DMD wurde als Analysepopulation die KITT-Population eingeschlossen, die Effektgrößen der Subgruppen wurden mittels der im SAP prädefinierten MMRM Methodik bestimmt (9–13). Für die Studie PTC124-GD-020-DMD wurde in den Subgruppenauswertungen die ITT Population, sowie das *a priori* definierte primäre ANVOVA-Modell eingesetzt (8).

Da die Definition der Schwellenwerte für den Effektmodifikator zu drei Kategorien führt ( $\geq 300$  bis  $< 400$  m;  $< 300$  m;  $\geq 400$  m), wurden nach den globalen Interaktionstests für die aggregierten Effektschätzer der drei Subgruppen in einem zweiten Schritt paarweise Interaktionstests durchgeführt, um signifikante Unterschiede zwischen den jeweiligen Subgruppen zu detektieren. Für die paarweisen Interaktionstests wurde die Signifikanzschwelle auf  $p=0.2$  angehoben, da hierdurch potentielle Probleme von nicht-transitiven Relationen vermieden werden (41).

Die Ergebnisse der Meta-Analysen für die drei Subgruppen im 6MWT sind in Abbildung 53 dargestellt. Der Gesamteffektschätzer über die beiden eingeschlossenen Studien für die Subgruppe mit BL 6MWD von  $\geq 300$  bis  $< 400$  m zeigte eine numerisch deutliche, signifikante Überlegenheit von Ataluren gegenüber Placebo ( $\Delta$  LSM: 45.33 m [95% KI: 20.16; 70.49];  $p=0.0004$ ). Die Heterogenität zwischen den beiden Studien war unbedeutend und nicht signifikant ( $p=0.79$ ,  $I^2=0.0\%$ ), daher war es gerechtfertigt, die Effektschätzer der beiden Studien in einer Meta-Analyse zusammenzufassen. Die Meta-Analyse für die Subgruppe mit BL 6MWD  $< 300$  m zeigte keinen signifikanten Therapieunterschied ( $\Delta$  LSM: 4.63 m [95% KI: -30.99; 40.24]), ebenso war der Gesamteffektschätzer der Subgruppe mit BL 6MWD

$\geq 400$  m nicht signifikant ( $\Delta$  LSM: 0.96 m [95% KI: -25.67; 27.59]). Da der globale Interaktionstest eine signifikante Heterogenität zwischen den drei aggregierten Subgruppen nachwies ( $p=0.04$ ;  $I^2=69.6\%$ ), wurden nachfolgend paarweise Interaktionstests der einzelnen Subgruppen durchgeführt.

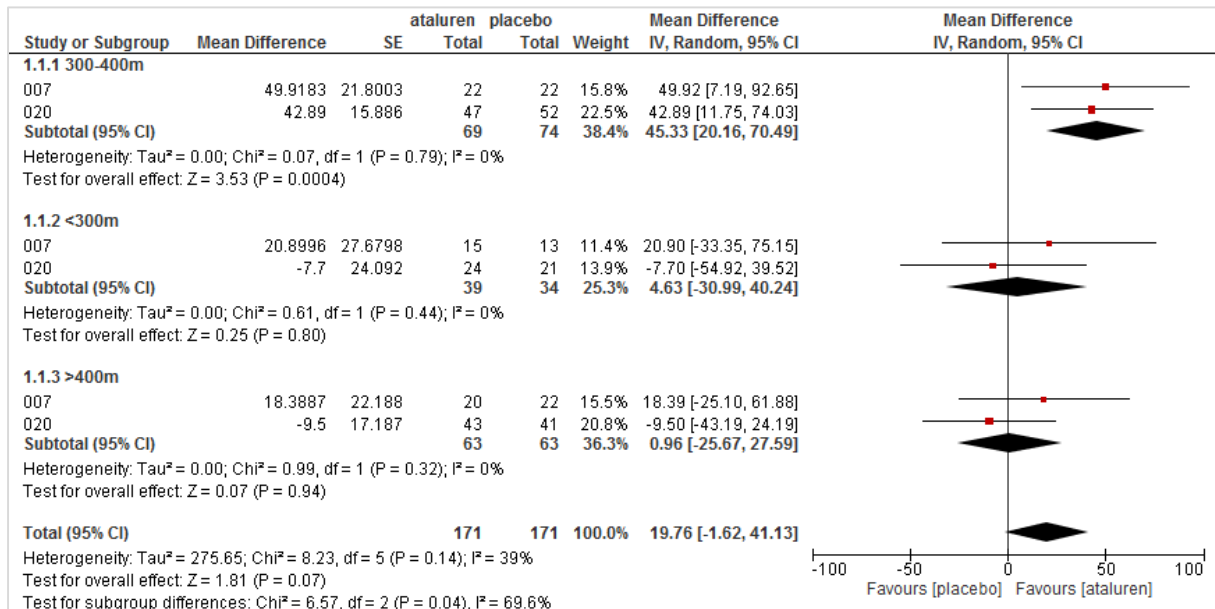


Abbildung 48: Meta-Analyse der Subgruppen (<300,  $\geq 300$  bis <400,  $\geq 400$  m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

In den paarweisen Interaktionstest wurden die aggregierten Effektschätzer folgender Subgruppen direkt miteinander verglichen: 300-400 m versus <300 m (Abbildung 54), 300-400 m versus  $\geq 400$  m (Abbildung 55), und <300 m versus  $\geq 400$  m (Abbildung 56). Wenn zwei Subgruppen untereinander homogen waren, konnten die beiden Subgruppen in eine gemeinsame Subgruppe zusammengefasst werden. Heterogene Subgruppen wurden nicht zusammengefasst und getrennt bewertet.

Für die drei Subgruppen im 6MWT ergaben sich dabei folgende Resultate. Die paarweisen Interaktionstests für die Subgruppen 300-400 m versus <300 m, und 300-400 m versus  $\geq 400$  m zeigten Heterogenität zwischen den Subgruppen ( $p < 0.2$ ). Im Gegensatz hierzu waren die aggregierten Effektschätzer in der Subgruppe <300 m und in der Subgruppe  $\geq 400$  m untereinander homogen ( $p \geq 0.2$ ). Aus diesem Grund wurden die Subgruppe <300 m und die Subgruppe  $\geq 400$  m in eine gemeinsame Subgruppe zusammengefasst. Für die Bewertung des Zusatznutzens wurde der aggregierte Effektschätzer der Subgruppe 300-400 m daher getrennt vom aggregierten Effektschätzer der zusammengefassten Subgruppen <300 m plus  $\geq 400$  m bewertet (Abbildung 52).

Zur Diskussion der biologischen Rationale der unterschiedlichen Subgruppeneffekte siehe Kapitel 4.4.2.4

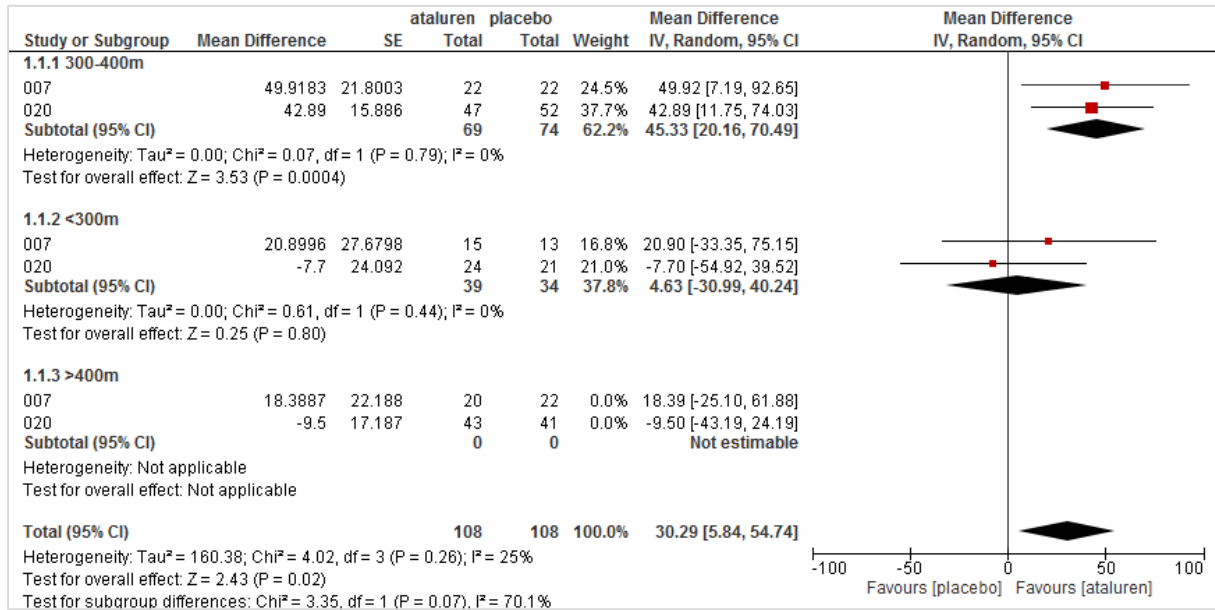


Abbildung 49: Paarweiser Interaktionst der Subgruppen (<300 m versus ≥300 bis <400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

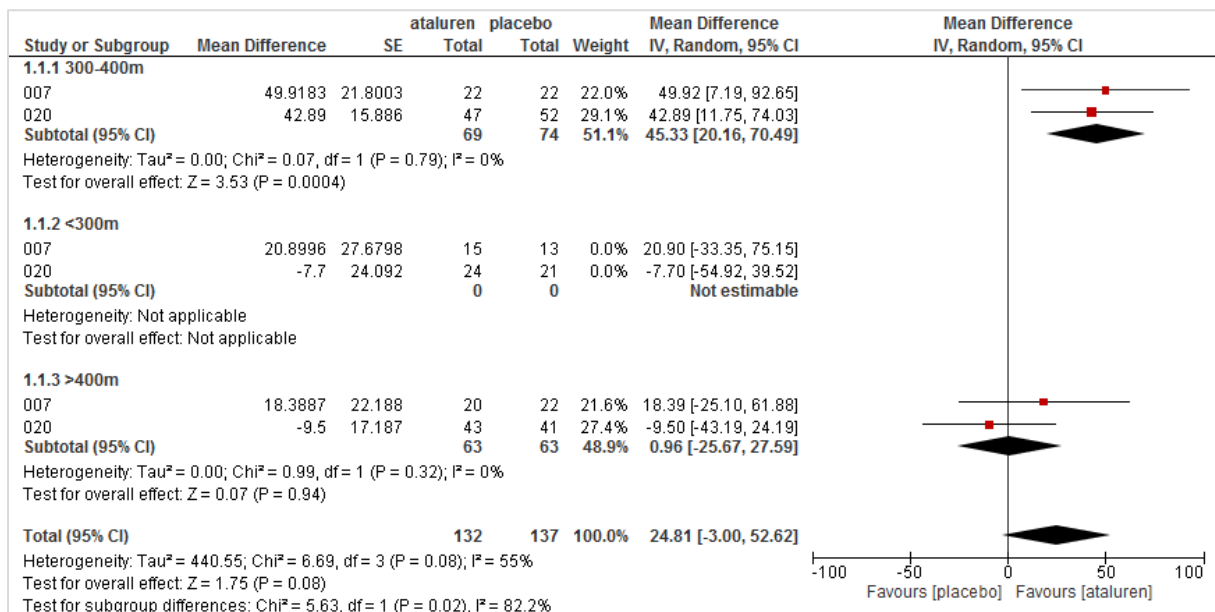


Abbildung 50: Paarweiser Interaktionst der Subgruppen (≥400 m versus ≥300 bis <400m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

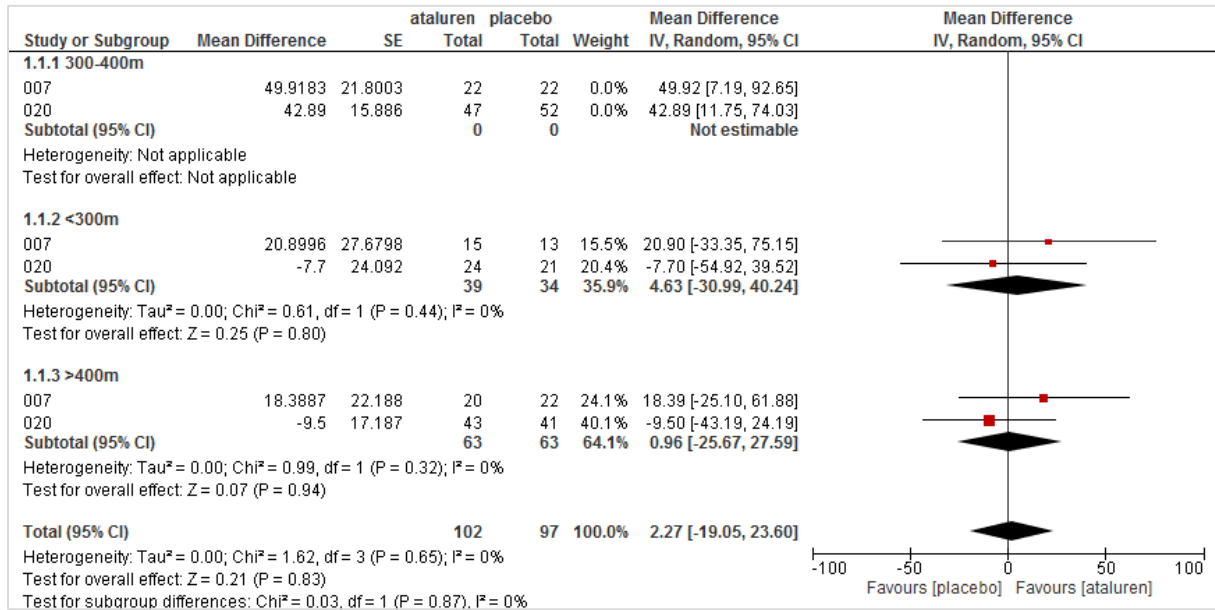


Abbildung 51: Paarweiser Interaktionstest der Subgruppen (<300 m versus ≥400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

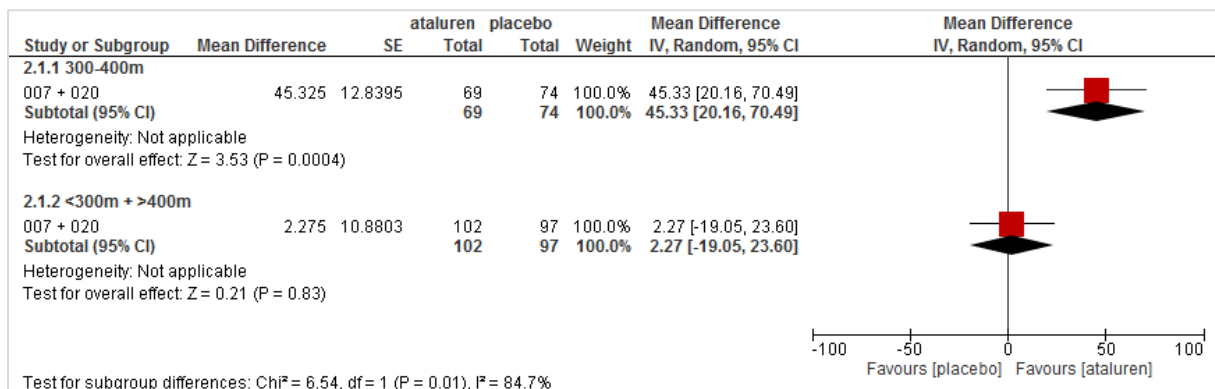


Abbildung 52: Gesamteffektschätzer der gepoolten Subgruppen (≥300 bis <400 m; <300 m plus ≥400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

Da für diesen Endpunkt noch kein validiertes, patientenindividuelles Responsekriterium existiert, wurde die Patientenrelevanz der Therapieunterschiede mittels einer Abschätzung der Effektgröße auf Basis des Effektgrößenmaßes Hedges' g adressiert. Das Konfidenzintervall des Therapieeffekts in der 300-400 m Subgruppe lag dabei deutlich über dem Schwellenwert von 0.2 für einen patientenrelevanten Effekt (H'g: 0.581 [95% KI: 0.246; 0.917]); s Tabelle 4-81, Kapitel 4.4.2.5.3.

#### 4.3.1.3.2.2 Zeit-Funktionstests - RCT

##### Studie PTC124-GD-020-DMD

Entsprechend IQWiG – Allgemeine Methoden Version 4.2 (7) ergaben die Tests auf Interaktion (Behandlung-Subgruppe) einen Hinweis auf unterschiedliche Effekte in den Subgruppen nach BL-6MWD  $\leq 350$  vs.  $> 350$  m (10 m gehen/laufen:  $p=0.137$ , 4 Stufen hinauf:  $p=0.131$ , 4 Stufen hinunter:  $p=0.135$ ) und  $< 300$ ,  $\geq 350$  bis  $< 400$  vs.  $\geq 400$  m (4 Stufen hinauf:  $p=0.080$ ). Hinsichtlich der Fähigkeit, vier Treppenstufen hinabzusteigen ergab sich bei den Subpopulationen nach BL-6MWD  $< 300$ ,  $\geq 350$  bis  $< 400$  vs.  $\geq 400$  m ein Beleg für unterschiedliche Effekte ( $p=0.037$ ), sodass es sich um einen tatsächlichen Unterschied zwischen den Subpopulationen handelt. Der Interaktionstest hinsichtlich der Fähigkeit, 10 m zu gehen/laufen, blieb mit  $p=0.208$  bei diesen Subgruppen knapp oberhalb des Grenzwerts von 0.2.

Die proximalen Muskelfunktionen wurden zu Studienbeginn unabhängig von den Testsystemen bei der Subpopulation mit guter Gehfähigkeit ( $\geq 400$  m) schneller ausgeführt als bei denjenigen mit bereits eingeschränktem Gehvermögen ( $< 300$  m). Die Subpopulation mit einer Gehstrecke von  $\geq 300$  bis  $< 400$  m nahm auch hinsichtlich der Fähigkeiten bei den TFTs eine Sonderstellung ein. Wie bereits beim 6MWT beobachtet, nahmen bei den Patienten mit eingeschränkter Gehfähigkeit ( $< 300$  m) die proximalen Muskelfunktionen über alle TFTs deutlicher ab als bei Patienten mit einer besseren Gehfähigkeit zu Beginn der Studie. Bei Patienten mit einer Gehstrecke über 400 m blieb die Muskelfunktion in beiden Gruppen unabhängig von der Behandlung und dem Testsystem über 48 Wochen weitgehend stabil. Die Konsistenz dieser Ergebnisse rechtfertigt die *a priori* geplante Analyse dieser Subpopulationen aus klinischer Sicht.

In der Subpopulation der Patienten mit einer BL-6MWD zwischen  $\geq 300$  und  $< 400$  m betrug der Unterschied der Fähigkeit, 10 m zu gehen oder laufen, im Mittel -1.84 sec (95% KI -3.79, 0.12,  $p=0.066$ ) zwischen Ataluren und Placebo (ITT; Tabelle 4-42). Das Ergebnis bestätigte die Primäranalyse und zeigte bei dieser Subgruppe einen deutlicheren Behandlungseffekt von Ataluren verglichen mit der Gesamtpopulation (-1.07 sec). Die mit Ataluren behandelten Patienten mit einer initialen Gehstrecke unter 300 m im 6MWT schnitten hinsichtlich ihrer Fähigkeit, 10 m zu gehen/laufen, etwas besser ab als die Patienten der Placebo-Gruppe, der Unterschied war nicht signifikant ( $p=0.066$ ). In der Subpopulation mit den längsten Gehstrecken ( $\geq 400$  m) änderte sich die Fähigkeit 10 zu gehen/laufen über den Studienverlauf in beiden Gruppen nicht.

Hinsichtlich der Fähigkeit, vier Stufen hinauf zu steigen, war Ataluren bei den Patienten, die zu Studienbeginn zwischen  $\geq 300$  und  $< 400$  m im 6MWT zurückgelegt hatten, Placebo signifikant überlegen (LSM -3.46 sec [95% KI: -5.71; -1.21],  $p=0.003$ ; ITT) (Tabelle 4-42). Der Effekt war wesentlich größer als in der Gesamtpopulation (-1.43 sec). Bei den Patienten mit bereits eingeschränktem Gehvermögen ( $< 300$  m) verschlechterte sich die Funktion unabhängig von der Behandlung, während die Funktionen bei den Patienten mit guter Gehfunktion ( $\geq 400$  m) während der Behandlung stabil blieben.

Wie schon bei den anderen TFTs war der Effekt hinsichtlich der Zeit, die benötigt wurde, um vier Stufen hinabzusteigen, in der Subgruppe mit einer BL-6MWD von  $\geq 300$  bis  $< 400$  m deutlicher ausgeprägt als in der Gesamtpopulation (Tabelle 4-42). Bei den Patienten der Ataluren-Gruppe nahm die Zeit, um diesen Test zu absolvieren, um 4.36 sec (LSM, 95% KI -6.64; -2.07; ITT) weniger ab als in der Placebo-Gruppe. Der Unterschied war mit  $p < 0.001$  in der ANCOVA-Analyse hoch signifikant. Die Ergebnisse der anderen Subpopulationen waren in ihrer Ausrichtung der Fähigkeit, vier Stufen hinaufzugehen, gleichgerichtet, d.h. rasche Entgleisung bei denjenigen mit einer Gehstrecke unter 300 m und stabile Funktion bei denjenigen mit einer Gehstrecke von  $\geq 400$  m.

Bei Patienten mit einer initialen Gehstrecke unter 350 m im 6MWT war Ataluren Placebo in allen TFTs signifikant überlegen (Tabelle 4-42). Bei dieser Subpopulation waren die Unterschiede zwischen Ataluren und Placebo in Woche 48 (LSM, 95% KI) wie folgt: Zeit, um 10 m zu gehen/laufen: -2.43 sec (-4.65, -0.20,  $p=0.033$ ;  $p=0.137$  Interaktion Subgruppen); Zeit, um vier Stufen hinauf zu gehen: -2.98 sec (-5.46, -0.49,  $p=0.019$ ;  $p=0.131$  Interaktion Subgruppen); Zeit, um vier Stufen hinab zu gehen: -3.61 sec [95% KI: -6.24; -0.98],  $p=0.007$ ;  $p=0.135$  Interaktion Subgruppen). Bei den Patienten mit einer Gehstrecke von  $\geq 350$  m im 6MWT zu Studienbeginn zeigten sich keine signifikanten Unterschiede zwischen Ataluren und Placebo und deren körperliche Fähigkeiten änderten sich über den Studienverlauf nur wenig.

Tabelle 4-42: Zeit für TFTs von BL bis Woche 48 in Subpopulationen nach BL-6MWD (ITT) für PTC124-GD-020-DMD; Ataluren versus Placebo (8)

Gruppe	n BL/ Woche 48	M (±SD) BL (sec)	M (±SD) Woche 48 (sec)	Δ LSM (± SE) in Woche 48 (sec)	Δ LSM (±SE) vs. Placebo [95% KI] (sec)	p-Wert vs. Placebo (p- Wert Interak- tion)
<i>Zeit, um 10 m zu gehen oder zu laufen, Subgruppe BL-6MWT &lt;300 m (Studienbericht 020: Tabelle 14.2.2.2.7, Tabelle 14.2.2.2.8)</i>						
Placebo	21/20	11.23 (2.824)	21.91 (9.314)	9.86 (1.173)		
Ataluren	24/22	10.81 (3.724)	18.60 (9.905)	7.11 (1.074)	-2.75 (1.497) [-5.69, 0.18]	0.066
<i>Zeit, um 10 m zu gehen oder zu laufen, Subgruppe BL-6MWT ≥300 bis &lt;400 m (Studienbericht 020: Tabelle 14.2.2.2.7, Tabelle 14.2.2.2.8)</i>						
Placebo	52/50	6.74 (1.717)	9.89 (6.644)	2.76 (0.762)		
Ataluren	47/45	6.27 (1.639)	7.37 (2.147)	0.92 (0.790)	-1.84 (0.999) [-3.79, 0.12]	0.066
<i>Zeit, um 10 m zu gehen oder zu laufen, Subgruppe BL-6MWT ≥400 m (Studienbericht 020: Tabelle 14.2.2.2.7, Tabelle 14.2.2.2.8)</i>						
Placebo	41/40	4.60 (1.021)	4.94 (1.237)	0.31 (0.806)		
Ataluren	43/42	4.69 (0.939)	5.20 (1.418)	0.51 (0.807)	0.21 (1.078) [-1.91, 2.32]	0.848 (0.208)
<i>Zeit, um 10 m zu gehen oder zu laufen, Subgruppe BL-6MWT &lt;350 m (Studienbericht 020: Tabelle 14.2.2.2.9, Tabelle 14.2.2.2.10)</i>						
Placebo	41/40	9.46 (2.873)	16.99 (9.457)	6.94 (0.899)		
Ataluren	41/38	9.30 (3.456)	14.41 (9.044)	4.51 (0.876)	-2.43 (1.136) [-4.65, -0.20]	<b>0.033</b>
<i>Zeit, um 10 m zu gehen oder zu laufen, Subgruppe BL-6MWT ≥350 m (Studienbericht 020: Tabelle 14.2.2.2.9, Tabelle 14.2.2.2.10)</i>						
Placebo	73/70	5.31 (1.582)	6.44 (4.503)	0.88 (0.657)		
Ataluren	73/71	5.13 (1.354)	5.81 (1.840)	0.58 (0.673)	-0.30 (0.846) [-1.96, 1.36]	0.721 (0.137 Hinweis)

Gruppe	n BL/ Woche 48	M (±SD) BL (sec)	M (±SD) Woche 48 (sec)	Δ LSM (± SE) in Woche 48 (sec)	Δ LSM (±SE) vs. Placebo [95% KI] (sec)	p-Wert vs. Placebo (p- Wert Interak- tion)
<i>Zeit, um 4 Treppenstufen hinaufzusteigen, Subgruppe BL-6MWT &lt;300 m (Studienbericht 020: Tabelle 14.2.2.3.7, Tabelle 14.2.2.3.8)</i>						
Placebo	19/17	14.18 (8.674)	23.66 (9.027)	9.15 (1.379)		
Ataluren	21/20	13.48 (7.176)	22.32 (9.370)	8.68 (1.269)	-0.47 (1.770) [-3.94, 3.00]	0.790
<i>Zeit, um 4 Treppenstufen hinaufzusteigen, Subgruppe BL-6MWT ≥300 bis &lt;400 m (Studienbericht 020: Tabelle 14.2.2.3.7, Tabelle 14.2.2.3.8)</i>						
Placebo	51/49	6.10 (3.072)	11.97 (9.942)	5.73 (0.878)	-	-
Ataluren	47/45	5.16 (1.959)	7.68 (4.852)	2.27 (0.907)	-3.46 (1.150) [-5.71, -1.21]	<b>0.003</b>
<i>Zeit, um 4 Treppenstufen hinaufzusteigen, Subgruppe BL-6MWT ≥400 m (Studienbericht Tabelle 14.2.2.3.7, Tabelle 14.2.2.3.8)</i>						
Placebo	41/40	3.09 (1.446)	3.41 (1.836)	0.37 (0.928)		
Ataluren	43/42	3.06 (1.239)	3.59 (2.376)	0.54 (0.927)	0.17 (1.240) [-2.26, 2.60]	0.893 (0.080 Hinweis)
<i>Zeit, um 4 Treppenstufen hinaufzusteigen, Subgruppe BL-6MWT &lt;350 m (Studienbericht 020: Tabelle 14.2.2.3.9, Tabelle 14.2.2.3.10)</i>						
Placebo	39/36	10.65 (7.326)	20.26 (9.759)	9.44 (1.003)		
Ataluren	38/36	10.24 (6.518)	16.89 (9.773)	6.44 (0.978)	-2.98 (1.269) [-5.46, -0.49]	<b>0.019</b>
<i>Zeit, um 4 Treppenstufen hinaufzusteigen, Subgruppe BL-6MWT ≥350 m (Studienbericht 020: Tabelle 14.2.2.3.9, Tabelle 14.2.2.3.10)</i>						
Placebo	72/70	4.05 (2.301)	5.65 (6.415)	1.57 (0.724)		
Ataluren	73/71	3.67 (1.668)	4.71 (3.591)	1.00 (0.738)	-0.57 (0.931) [-2.39, 1.25]	0.541 (0.131 Hinweis)



Gruppe	n BL/ Woche 48	M ( $\pm$ SD) BL (sec)	M ( $\pm$ SD) Woche 48 (sec)	$\Delta$ LSM ( $\pm$ SE) in Woche 48 (sec)	$\Delta$ LSM ( $\pm$ SE) vs. Placebo [95% KI] (sec)	p-Wert vs. Placebo (p- Wert Interak- tion)
<i>Zeit, um 4 Treppenstufen hinabzusteigen, Subgruppe BL-6MWT &lt;300 m (Studienbericht 020: Tabelle 14.2.2.4.7, Tabelle 14.2.2.4.8)</i>						
Placebo	16/17	10.48 (8.429)	22.76 (9.878)	9.38 (1.476)		
Ataluren	21/21	11.88 (8.088)	20.99 (9.755)	8.41 (1.277)	-0.97 (1.828) [-4.55, 2.61]	0.595
<i>Zeit, um 4 Treppenstufen hinabzusteigen, Subgruppe BL-6MWT <math>\geq</math>300 bis &lt;400 m (Studienbericht 020: Tabelle 14.2.2.4.7, Tabelle 14.2.2.4.8)</i>						
Placebo	51/49	4.63 (2.753)	9.82 (9.696)	4.90 (0.894)		
Ataluren	47/45	4.03 (1.645)	5.17 (2.529)	0.54 (0.915)	-4.36 (1.165) [-6.64, -2.07]	<b>&lt;0.001</b>
<i>Zeit, um 4 Treppenstufen hinabzusteigen, Subgruppe BL-6MWT <math>\geq</math>400 m (Studienbericht 020: Tabelle 14.2.2.4.7, Tabelle 14.2.2.4.8)</i>						
Placebo	41/40	2.74 (1.084)	2.81 (1.246)	-0.08 (0.938)		
Ataluren	43/42	2.76 (0.803)	2.90 (1.195)	-0.21 (0.935)	-0.31 (1.251) [-2.58, 2.32]	0.917 <b>(0.037 Beleg)</b>
<i>Zeit, um 4 Treppenstufen hinabzusteigen, Subgruppe BL-6MWT &lt;350 m (Studienbericht 020: Tabelle 14.2.2.4.9, Tabelle 14.2.2.4.10)</i>						
Placebo	36/36	7.84 (6.578)	18.53 (11.027)	8.89 (1.081)		
Ataluren	38/37	8.40 (7.182)	14.54 (10.607)	5.28 (1.014)	-3.61 (1.342) [-6.24, -0.98]	<b>0.007</b>
<i>Zeit, um 4 Treppenstufen hinabzusteigen, Subgruppe BL-6MWT <math>\geq</math>350 m (Studienbericht 020: Tabelle 14.2.2.4.9, Tabelle 14.2.2.4.10)</i>						
Placebo	72/70	3.24 (1.505)	4.48 (5.328)	0.99 (0.752)		
Ataluren	73/71	3.27 (1.430)	3.63 (1.890)	-0.07 (0.766)	-1.07 (0.966) [-2.96, 0.82]	0.268 (0.135 Hinweis)

ANCOVA mit multipler Imputation; Faktoren: Stratifizierungsfaktoren Alter, BL-6MWD-Gruppe (<300 m,  $\geq$ 300 m bis <400 m und  $\geq$ 400 m bei Analyse nach diesen 6MWD-Gruppen), Dauer der Kortikosteroid-Vorbehandlung, Behandlung, BL-TFT und Interaktion zwischen Behandlung und Subgruppe

### Fazit

In den TFTs zeigte Ataluren durchgängig einen positiven Behandlungseffekt verglichen mit Placebo. Der Unterschied war in der Gesamtpopulation allerdings nur hinsichtlich der Fähigkeit, vier Treppenstufen hinab zu gehen, signifikant. Wie bereits im 6MWT beobachtet, waren die Unterschiede bei den Patienten mit einer initialen Gehstrecke zwischen  $\geq$ 300 m und <400 m, bei denen nach jüngsten Erkenntnissen u.a der Ataluren-Studien ein Behandlungseffekt über 48 Wochen am ehesten nachweisbar ist (siehe Kapitel 4.4.2), unabhängig von dem Testsystem wesentlich deutlicher ausgeprägt als in der heterogenen Gesamtpopulation (10 m gehen/laufen: -1.07 vs. -1.84 sec, 4 Stufen hinauf: -1.43 vs. -3.64 sec, 4 Stufen hinunter: -1.97 vs. 4.36 sec). In Übereinstimmung mit den 6MWT-Ergebnissen verschlechterte sich die proximale Muskelfunktion unabhängig von dem Testsystem bei Patienten mit bereits eingeschränkter Gehfunktion (<300 m) rasch, während Patienten mit noch guter Gehfähigkeit ( $\geq$ 400 m) über die verschiedenen Tests weitgehend stabil blieben. Hinsichtlich der Fähigkeit, vier Stufen hinauf zu gehen, ergab sich ein Hinweis und

hinsichtlich der Fähigkeit, vier Stufen hinabzugehen, ein Beleg für unterschiedliche Effekte zwischen den Subgruppen. Zusammen mit dem Beleg eines unterschiedlichen Effekts auf die 6MWD liefern die Daten eine robuste Evidenz für einen zwischen den Subpopulationen (<300 m, ≥300 bis <400, ≥400 m) unterschiedlichen Effekt.

Die Subgruppen-Analysen nach BL-6MWD von <350 vs. ≥350 m zeigten durchgängig einen signifikanten Behandlungseffekt von Ataluren bei Patienten mit schlechterer Gehfähigkeit, und keine nennenswerte Änderung bei Patienten mit der besseren Gehfähigkeit zu Beginn der Studie. Die Interaktionstests ergaben Hinweise auf einen zwischen den Subpopulationen unterschiedlichen Effekt ( $0.05 < p < 0.2$ ). In ihrer Gesamtheit zeigen die Ergebnisse der Studie übereinstimmend über verschiedene Testsysteme, dass es sich bei den Subpopulationen mit einer Gehstrecke von <300, ≥300 bis <400 bzw. ≥400 m um klinisch distinkte Krankheitsstadien mit deutlich unterschiedlichem Krankheitsverlauf handelt. Daher erscheint es plausibel, dass der Effekt bei Patienten mit einer Gehstrecke von <350 vs. ≥350 m durch den Effekt bei den Subpopulationen mit einer Gehstrecke von <300, ≥300 bis <400 bzw. ≥400 m überlagert bzw. bedingt wird.

### **Meta-Analysen der Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD**

#### *Zeit, um 4 Treppenstufen hinaufzusteigen*

Die Ergebnisse der Meta-Analysen der drei Subgruppen für den Endpunkt Zeit, um 4 Treppenstufen hinaufzusteigen sind in Abbildung 53 dargestellt. Der Gesamteffektschätzer über die beiden eingeschlossenen Studien für die Subgruppe mit BL 6MWD von ≥300 bis <400 m zeigte eine numerisch deutliche, signifikante Überlegenheit von Ataluren gegenüber Placebo ( $\Delta$  LSM: -3.47 sec [95% KI: -5.35; -1.59];  $p=0.0003$ ). Die Heterogenität zwischen den beiden Studien war unbedeutend und nicht signifikant ( $p=0.96$ ,  $I^2=0.00\%$ ), daher war es gerechtfertigt, die Effektschätzer der beiden Studien in einer Meta-Analyse zusammenzufassen. Die Meta-Analyse für die Subgruppe mit BL 6MWD < 300 m zeigte einen Trend zur Überlegenheit von Ataluren ( $\Delta$  LSM: -2.30 sec [95% KI: -6.22; 1.61]), ebenso war der Gesamteffektschätzer der Subgruppe mit BL 6MWD ≥400 m nicht signifikant ( $\Delta$  LSM: 0.05 sec [95% KI: -2.04; 1.95]). Da der globale Interaktionstest eine signifikante Heterogenität zwischen den drei aggregierten Subgruppen nachwies ( $p=0.049$ ;  $I^2=66.86\%$ ), wurden nachfolgend paarweise Interaktionstests der einzelnen Subgruppen durchgeführt.

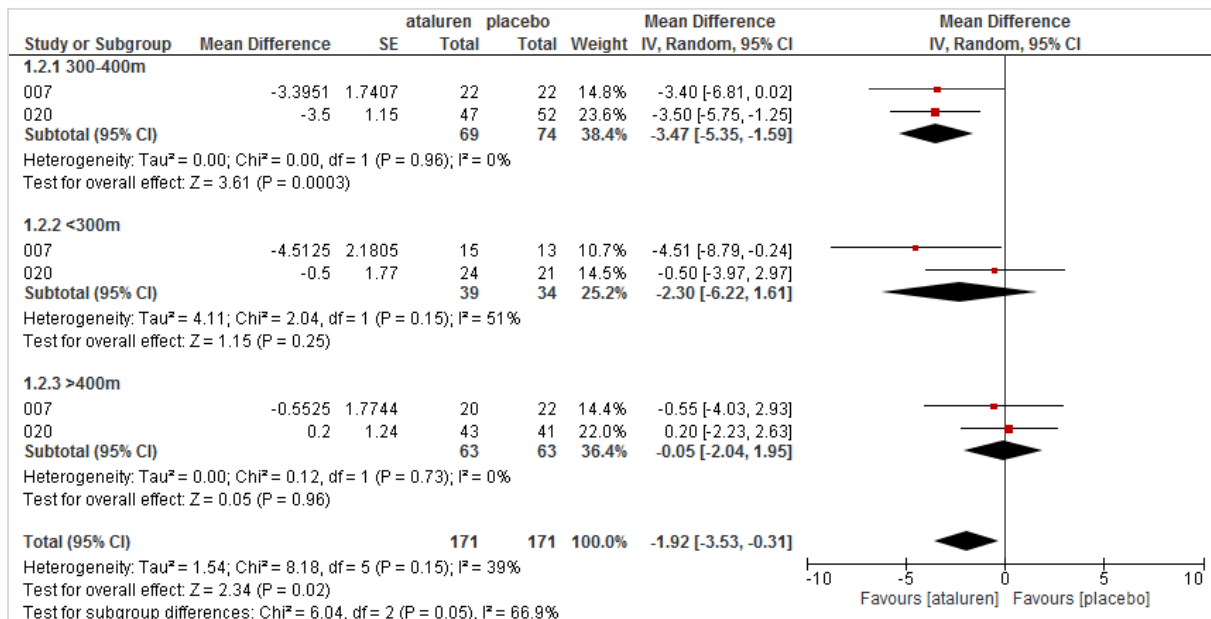


Abbildung 53: Meta-Analyse der Subgruppen (<300, ≥300 bis <400, ≥400 m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

In den paarweisen Interaktionstest wurden die aggregierten Effektschätzer folgender Subgruppen direkt miteinander verglichen: 300-400 m versus <300 m (Abbildung 54), 300-400 m versus ≥400 m (Abbildung 55), und <300 m versus ≥400 m (Abbildung 56). Wenn zwei Subgruppen untereinander homogen waren, konnten die beiden Subgruppen in eine gemeinsame Subgruppe zusammengefasst werden. Heterogene Subgruppen wurden nicht zusammengefasst und getrennt bewertet.

Für die drei Subgruppen ergaben sich dabei folgende Resultate. Der paarweise Interaktionstest für die Subgruppe 300-400 m versus <300 m zeigte keine Heterogenität zwischen den Subgruppen; der paarweise Interaktionstest 300-400 m versus ≥400 m zeigte jedoch Heterogenität zwischen den Subgruppen ( $p < 0.2$ ). Die aggregierten Effektschätzer der Subgruppe <300 m und in der Subgruppe ≥400 m waren untereinander homogen ( $p \geq 0.2$ ). Für diesen seltenen Fall einer nicht-transitiven Relation ist die Entscheidung, welche Subgruppen gepoolt werden können aus rein statistischer Sicht nicht eindeutig zu treffen. Aus Gründen der Konsistenz mit der Vorgangsweise für den 6MWT wurden die Subgruppe <300 m und die Subgruppe ≥400 m in eine gemeinsame Subgruppe zusammengefasst. Für die Bewertung des Zusatznutzens wurde der aggregierte Effektschätzer der Subgruppe 300-400 m daher getrennt vom aggregierten Effektschätzer der zusammengefassten Subgruppen <300 m plus ≥400 m bewertet (Abbildung 57).

Zur Diskussion der biologischen Rationale der unterschiedlichen Subgruppeneffekte siehe Kapitel 4.4.2.4

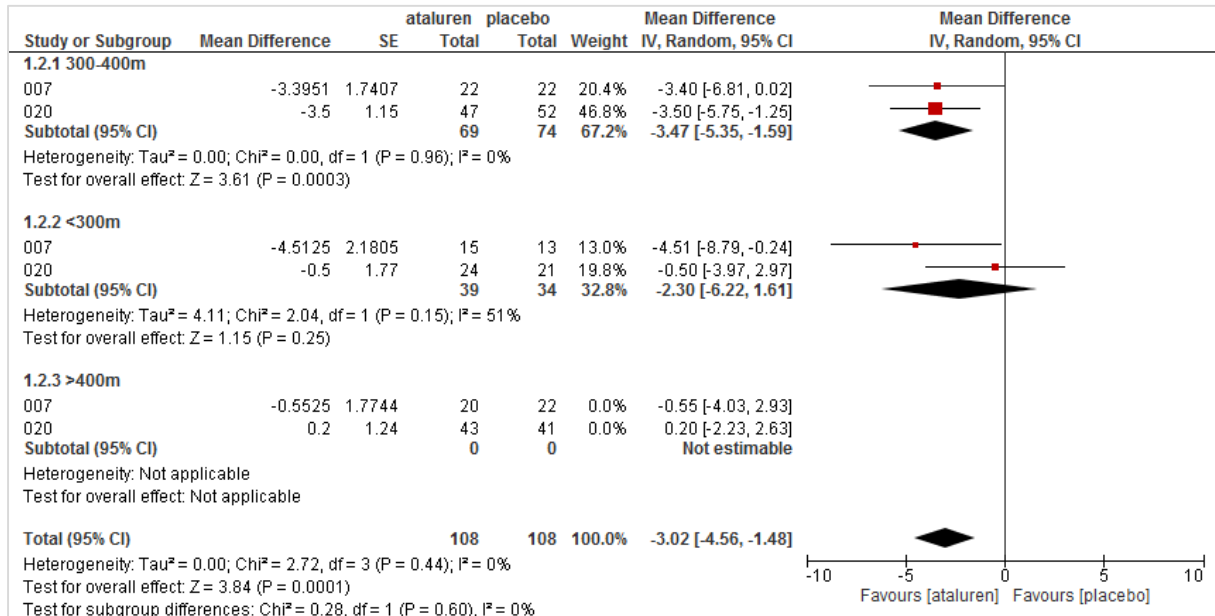


Abbildung 54: Paarweiser Interaktionst der Subgruppen (<300 m versus ≥300 bis <400 m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

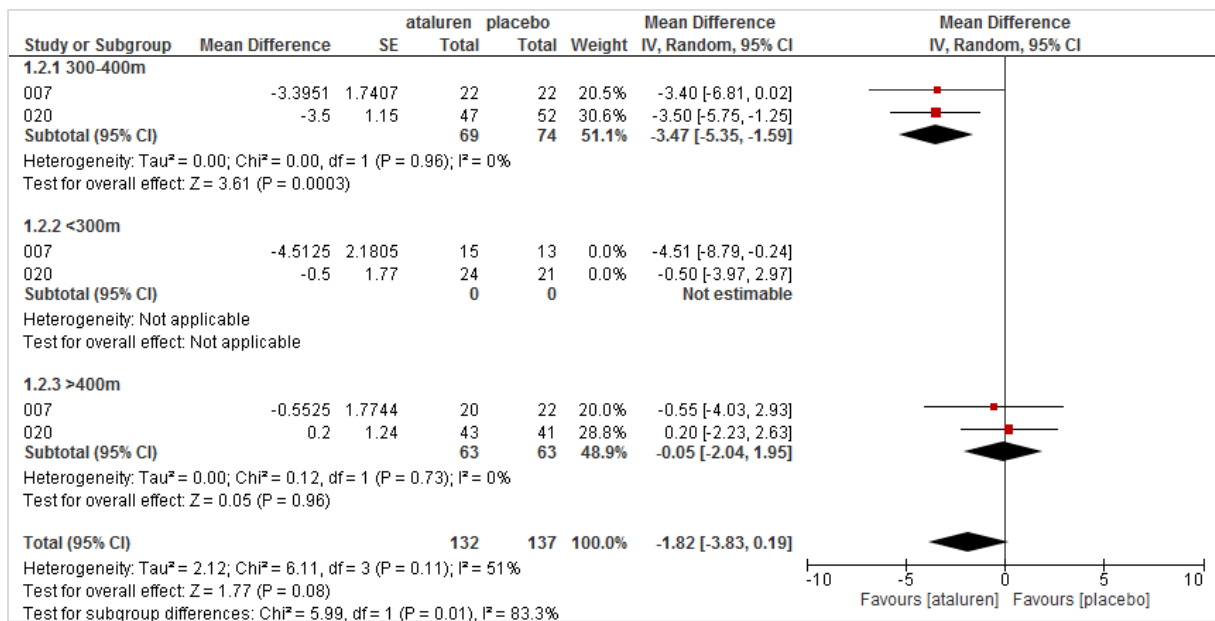


Abbildung 55: Paarweiser Interaktionst der Subgruppen (≥400 m versus ≥300 bis <400m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

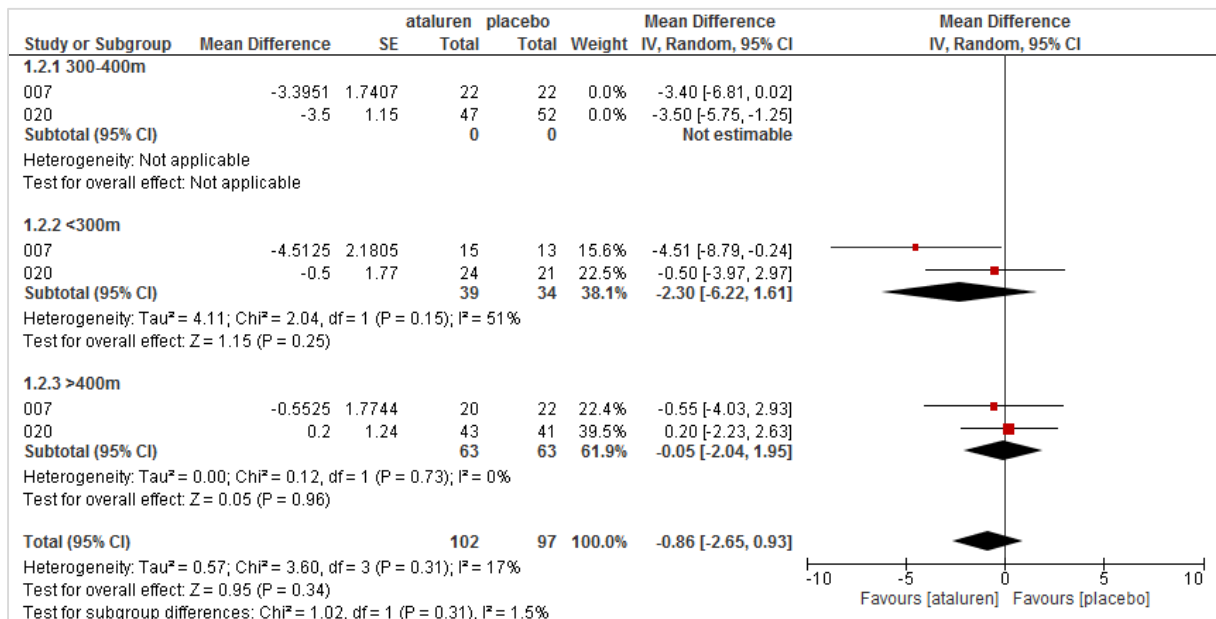


Abbildung 56: Paarweiser Interaktionstest der Subgruppen (<300 m versus ≥400 m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

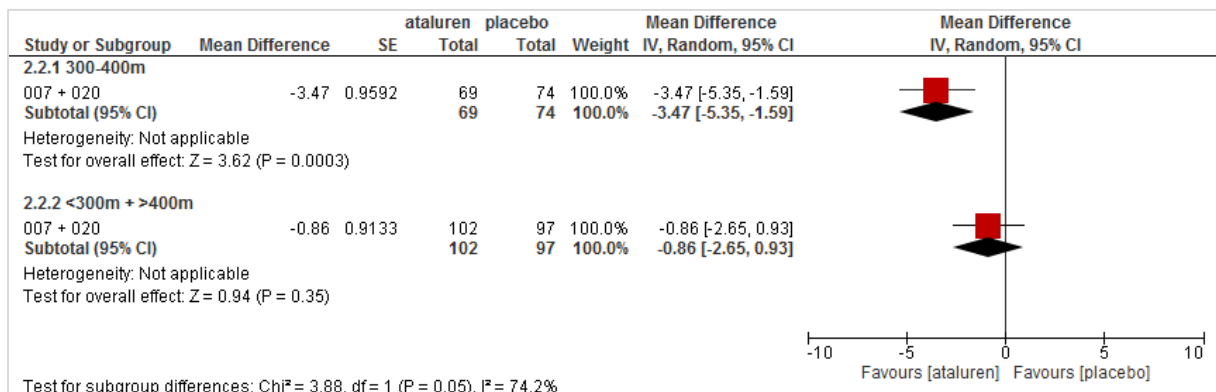


Abbildung 57: Gesamteffektschätzer der gepoolten Subgruppen (≥300 bis <400 m; <300 m plus ≥400 m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

Ein Therapieunterschied im 4 Treppensteigttest von 2.1-2.2 Sekunden wird für diese Patientenpopulation als Patienten-relevant eingeschätzt (18). Da für diesen Endpunkt jedoch noch kein formell validiertes Resoponsekriterium existiert, wurde die Patientenrelevanz der Therapieunterschiede mittels einer Abschätzung der Effektgröße auf Basis des Effektgrößenmaßes Hedges' g adressiert. Das Konfidenzintervall des Therapieeffekts in der 300-400 m Subgruppe lag dabei deutlich über dem Schwellenwert von 0.2 für einen patientenrelevanten Effekt (H'g: -0.598 [95% KI: -0.934; -0.263]); s. Tabelle 4-81, Kapitel 4.4.2.5.3.

*Zeit, um 4 Treppenstufen hinabzusteigen*

Die Ergebnisse der Meta-Analysen der drei Subgruppen für den Endpunkt Zeit, um 4 Treppenstufen hinabzusteigen sind in Abbildung 58 dargestellt. Der Gesamteffektschätzer über die beiden eingeschlossenen Studien für die Subgruppe mit BL 6MWD von  $\geq 300$  bis  $< 400$  m zeigte eine numerisch deutliche, signifikante Überlegenheit von Ataluren gegenüber Placebo ( $\Delta$  LSM: -4.31 sec [95% KI: -6.25; -2.37];  $p < 0.0001$ ). Die Heterogenität zwischen den beiden Studien war unbedeutend und nicht signifikant ( $p = 0.89$ ,  $I^2 = 0.00\%$ ), daher war es gerechtfertigt, die Effektschätzer der beiden Studien in einer Meta-Analyse zusammenzufassen. Die Meta-Analyse für die Subgruppe mit BL 6MWD  $< 300$  m zeigte keinen signifikanten Therapieunterschied ( $\Delta$  LSM: -0.90 sec [95% KI: -3.72; 1.93]), ebenso war der Gesamteffektschätzer der Subgruppe mit BL 6MWD  $\geq 400$  m nicht signifikant ( $\Delta$  LSM: -0.18 sec [95% KI: -2.23, 1.87]). Da der globale Interaktionstest eine signifikante Heterogenität zwischen den drei aggregierten Subgruppen nachwies ( $p = 0.011$ ;  $I^2 = 77.93\%$ ), wurden nachfolgend paarweise Interaktionstests der einzelnen Subgruppen durchgeführt.

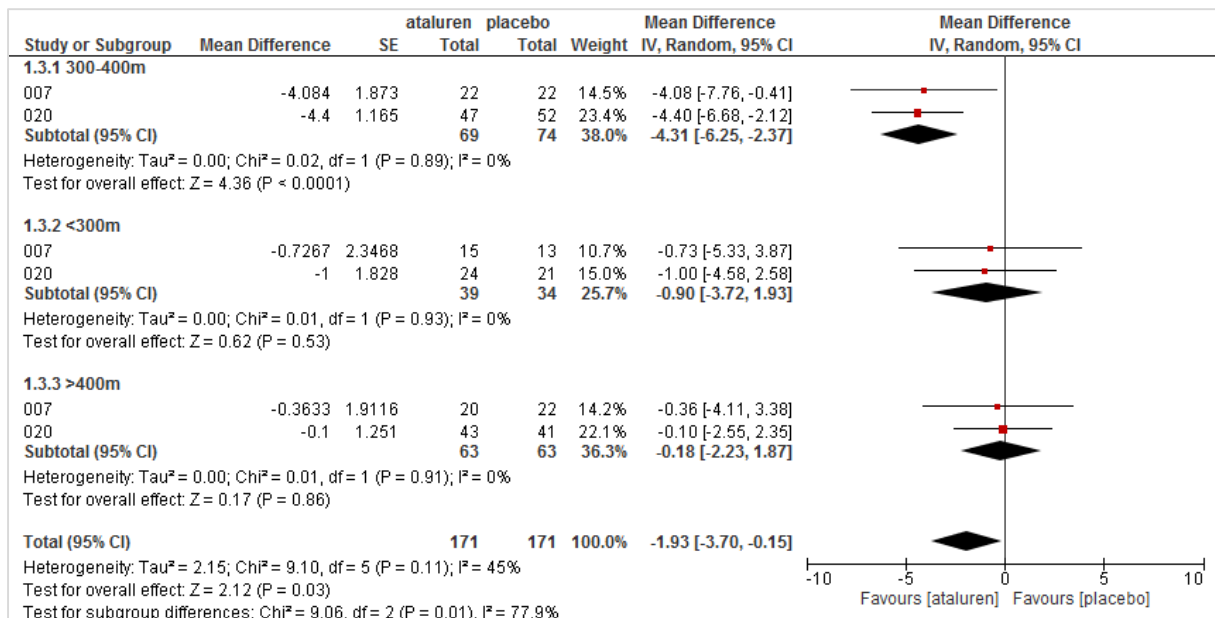


Abbildung 58: Meta-Analyse der Subgruppen ( $< 300$ ,  $\geq 300$  bis  $< 400$ ,  $\geq 400$  m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

In den paarweisen Interaktionstest wurden die aggregierten Effektschätzer folgender Subgruppen direkt miteinander verglichen: 300-400 m versus  $< 300$  m (Abbildung 59), 300-400 m versus  $\geq 400$  m (Abbildung 60), und  $< 300$  m versus  $\geq 400$  m (Abbildung 61). Wenn zwei Subgruppen untereinander homogen waren, konnten die beiden Subgruppen in eine gemeinsame Subgruppe zusammengefasst werden. Heterogene Subgruppen wurden nicht zusammengefasst und getrennt bewertet.

Für die drei Subgruppen ergaben sich dabei folgende Resultate. Die paarweisen Interaktionstests für die Subgruppen 300-400 m versus <300 m, und 300-400 m versus  $\geq 400$  m zeigten Heterogenität zwischen den Subgruppen ( $p < 0.2$ ). Im Gegensatz hierzu waren die aggregierten Effektschätzer in der Subgruppe <300 m und in der Subgruppe  $\geq 400$  m untereinander homogen ( $p \geq 0.2$ ). Aus diesem Grund wurden die Subgruppe <300 m und die Subgruppe  $\geq 400$  m in eine gemeinsame Subgruppe zusammengefasst. Für die Bewertung des Zusatznutzens wurde der aggregierte Effektschätzer der Subgruppe 300-400 m daher getrennt vom aggregierten Effektschätzer der zusammengefassten Subgruppen <300 m plus  $\geq 400$  m bewertet (Abbildung 62).

Zur Diskussion der biologischen Rationale der unterschiedlichen Subgruppeneffekte siehe Kapitel 4.4.2.4

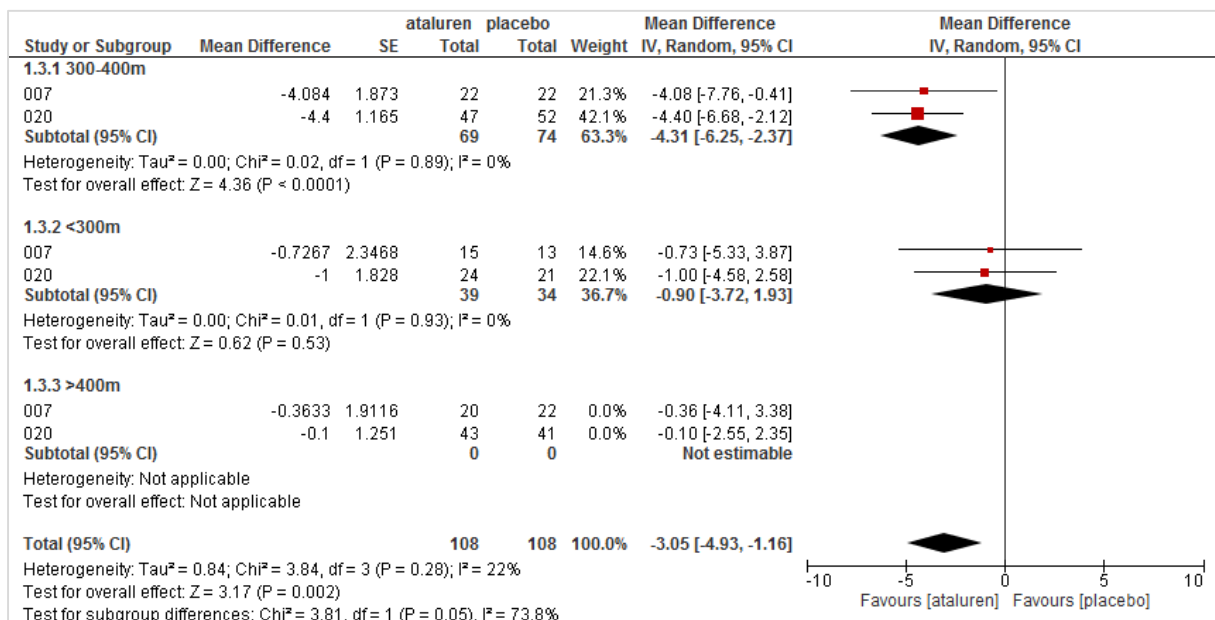


Abbildung 59: Paarweiser Interaktionstest der Subgruppen (<300 m versus  $\geq 300$  bis <400 m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

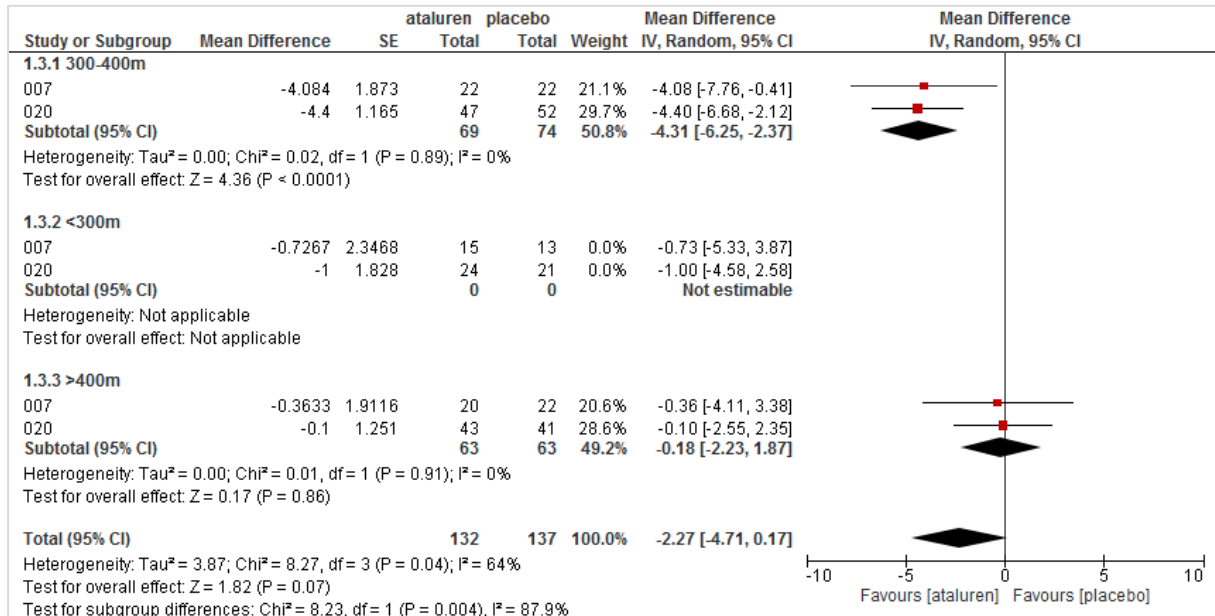


Abbildung 60: Paarweiser Interaktionstest der Subgruppen ( $\geq 400$  m versus  $\geq 300$  bis  $< 400$  m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

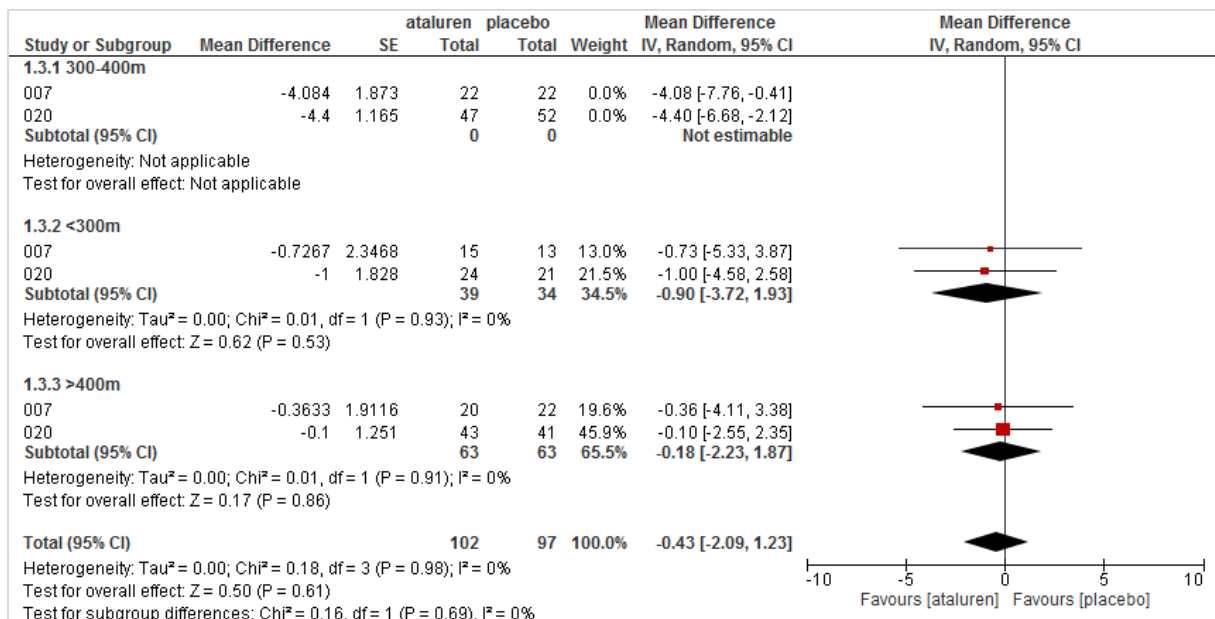


Abbildung 61: Paarweiser Interaktionstest der Subgruppen ( $< 300$  m versus  $\geq 400$  m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo



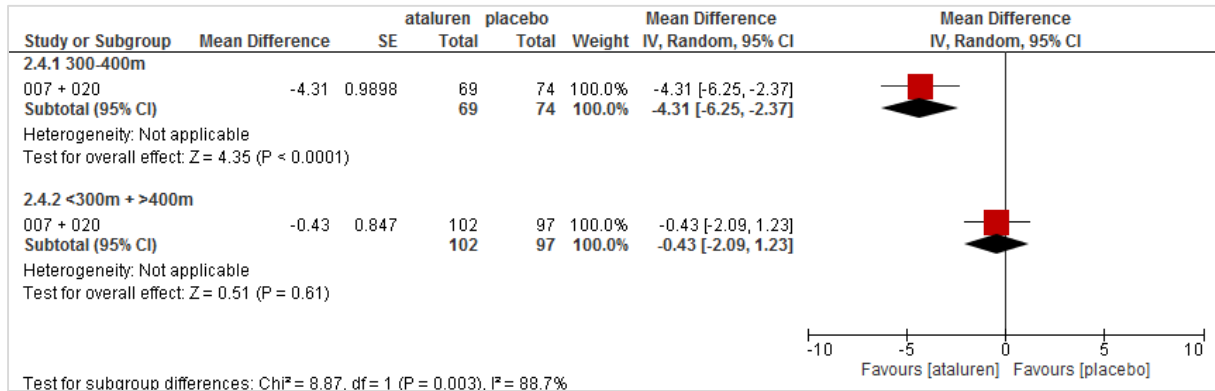


Abbildung 62: Gesamteffektschätzer der gepoolte Subgruppen ( $\geq 300$  bis  $< 400$  m;  $< 300$  m plus  $\geq 400$  m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

Da für diesen Endpunkt noch kein formell validiertes Resoponsekriterium existiert, wurde die Patientenrelevanz der Therapieunterschiede mittels einer Abschätzung der Effektgröße auf Basis des Effektgrößenmaßes Hedges' g adressiert. Das Konfidenzintervall des Therapieeffekts in der 300-400 m Subgruppe lag dabei deutlich über dem Schwellenwert von 0.2 für einen patientenrelevanten Effekt (H'g: -0.720 [95% KI: -1.060; -0.381]); s. Tabelle 4-81, Kapitel 4.4.2.5.3.

#### *Zeit, um 10 m zu gehen oder zu laufen*

Die Ergebnisse der Meta-Analysen für die drei Subgruppen des Endpunkts „Zeit, um 10 m zu gehen oder zu laufen“ sind in Abbildung 63 dargestellt. Der Gesamteffektschätzer über die beiden eingeschlossenen Studien für die Subgruppe mit BL 6MWD von  $\geq 300$  bis  $< 400$  m zeigte eine numerisch deutliche, signifikante Überlegenheit von Ataluren gegenüber Placebo ( $\Delta$  LSM: -2.18 sec [95% KI: -3.82; -0.54];  $p=0.009$ ). Die Heterogenität zwischen den beiden Studien war mittelmäßig und nicht signifikant ( $p=0.48$ ,  $I^2=0.0\%$ ), daher war es gerechtfertigt, die Effektschätzer der beiden Studien in der Meta-Analyse zusammenzufassen. Die Meta-Analyse für die Subgruppe mit BL 6MWD  $< 300$  m zeigte keinen signifikanten Therapieunterschied ( $\Delta$  LSM: -2.18 sec [95% KI: -4.50; 0.14]), ebenso war der Gesamteffektschätzer der Subgruppe mit BL 6MWD  $\geq 400$  m nicht signifikant ( $\Delta$  LSM: 0.09 sec [95% KI: -1.65; 1.83]). Da der globale Interaktionstest keine signifikante Heterogenität zwischen den drei aggregierten Subgruppen nachwies ( $p=0.13$ ;  $I^2=69.6\%$ ), wurden auch keine paarweisen Interaktionstests der einzelnen Subgruppen durchgeführt. Für diesen Endpunkt wird daher der Gesamteffektschätzer aus der Meta-Analyse der KITT/ITT Population für die Nutzenbewertung herangezogen (s. Abbildung 17). Die Subgruppen werden nicht separat bewertet.

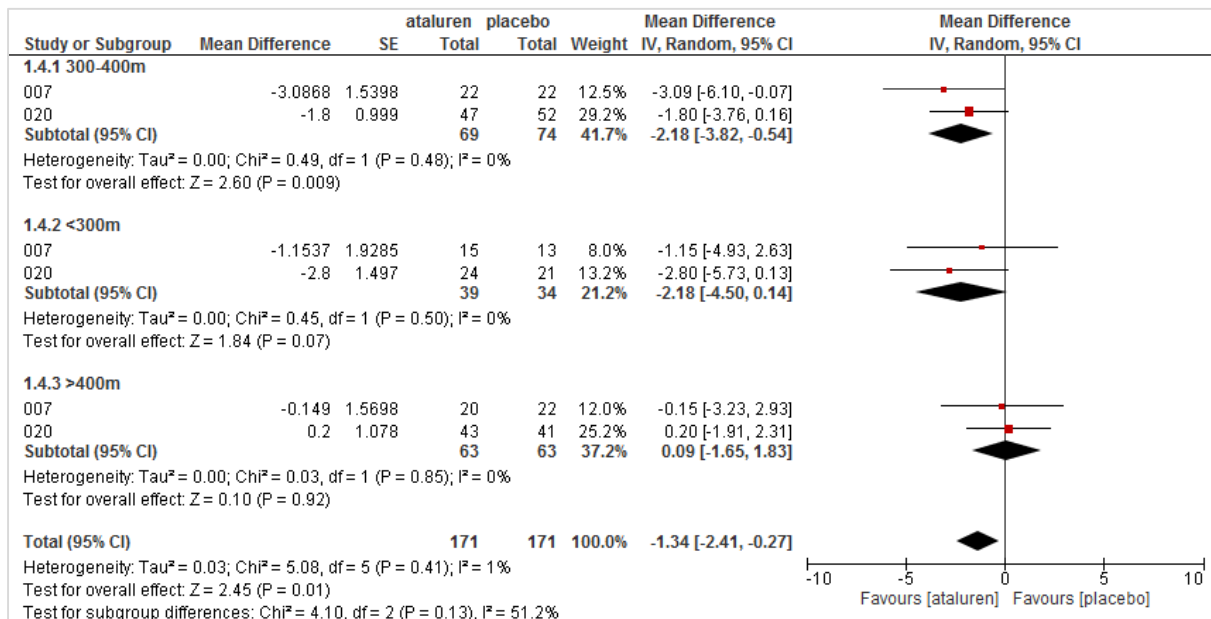


Abbildung 63: Meta-Analyse der Subgruppen (<300, ≥300 bis <400, ≥400 m) für Zeit, um 10 m zu gehen/laufen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

#### 4.3.1.3.2.3 North Star Ambulatory Assessment (NSAA) – RCT

##### Studie PTC124-GD-020-DMD

Die Interaktionstests (Behandlung-Subgruppe) erbrachten keine Hinweise oder Belege für unterschiedliche Effekte zwischen den BL-6MWD-Subgruppen. Der Interaktionstest lag mit  $p=0.207$  nur geringfügig oberhalb des Schwellenwertes für einen Hinweis auf unterschiedliche Effekte zwischen den BL-6MWD-Gruppen (<300 m, ≥300 bis <400 m, ≥400 m).

Auch in der Analyse des NSAA waren die körperlichen Fähigkeiten in Übereinstimmung mit den zuvor dargestellten Testsystemen bei Patienten mit guter Gehfähigkeit besser als bei Patienten, deren Gehfähigkeit bereits zu Studienbeginn deutlich eingeschränkt war. Bei den Patienten mit einer initialen Gehstrecke von ≥300 bis <400 m im 6MWT war der Unterschied zugunsten von Ataluren wie bereits bei den anderen Tests höher als in der Gesamtpopulation und erreichte statistische Signifikanz (Tabelle 4-43;  $\Delta$  LSM: 4.26 [95% KI: 0.17; 8.35],  $p=0.041$ , transformierte Werte), allerdings überschritt der p-Wert des Interaktionstests die Schwelle von  $p<0.2$  wie bereits oben angemerkt. Wiederum zeigte Ataluren bei Patienten mit einer initialen Gehstrecke unter 300 m bzw. von ≥400 m und höher keinen Effekt; bei den Letztgenannten änderten sich die Fähigkeiten während der Studie nur geringfügig. Wie bereits bei den anderen Testsystemen zeigte Ataluren bei den Patienten mit einer Gehstrecke unter 350 m einen numerisch größeren Effekt als bei Patienten, die zu Studienbeginn 350 m oder mehr im 6MWT bewältigen konnten (Tabelle 4-43).

Tabelle 4-43: Änderung NSAA von BL bis Woche 48 (ITT; BL-6MWD- Subpopulationen) für PTC124-GD-020-DMD; Ataluren versus Placebo (8)

Gruppe	n BL/ Woche 48	M (±SD) BL (sec)	M (±SD) in Woche 48 (sec)	LSM (± SE) Woche 48 (sec)	Δ vs. Placebo, M (±SE) [95% KI] (sec)	p-Wert vs. Placebo (p-Wert Interak- tion)
<i>Änderung des NSAA-Gesamtscores (nicht transformiert): 6MWD &lt;300 m (Studienbericht 020: Tabelle 14.2.3.5, Tabelle 14.2.3.6)</i>						
Placebo	21/19	11.9 (5.63)	7.1 (5.72)	-4.29 (0.960)	-	-
Ataluren	24/18	11.9 (3.65)	8.7 (5.85)	-3.91 (0.886)	0.37 (1.222) [-2.02, 2.77]	0.760
<i>Änderung des NSAA-Gesamtscores (nicht transformiert): 6MWD ≥300 bis &lt;400 m (Studienbericht 020: Tabelle 14.2.3.5, Tabelle 14.2.3.6)</i>						
Placebo	52/49	20.6 (5.92)	16.1 (7.89)	-4.28 (0.627)	-	-
Ataluren	47/45	21.4 (5.54)	18.3 (6.81)	-2.57 (0.645)	1.70 (0.814) [0.11, 3.30]	<b>0.037</b>
<i>Änderung des NSAA-Gesamtscores (nicht transformiert): 6MWD ≥400 m (Studienbericht 020: Tabelle 14.2.3.5, Tabelle 14.2.3.6)</i>						
Placebo	41/40	28.5 (4.64)	26.7 (6.36)	-1.57 (0.657)	-	-
Ataluren	43/42	28.7 (4.26)	26.9 (5.08)	-1.68 (0.659)	-0.11 (0.876) [-1.83, 1.60]	0.896 (0.291)
<i>Änderung des NSAA-Gesamtscores (nicht transformiert): 6MWD &lt;350 m (Studienbericht 020: Tabelle 14.2.3.9, Tabelle 14.2.3.10)</i>						
Placebo	41/39	15.1 (6.32)	9.4 (6.58)	-5.41 (0.695)	-	-
Ataluren	41/34	14.7 (5.31)	11.2 (6.26)	-3.95 (0.681)	1.46 (0.878) [-0.26, 3.18]	0.097
<i>Änderung des NSAA-Gesamtscores (nicht transformiert): 6MWD ≥350 m (Studienbericht: Tabelle 14.2.3.9, Tabelle 14.2.3.10)</i>						
Placebo	73/69	25.7 (6.19)	23.5 (7.66)	-2.20 (0.512)	-	-
Ataluren	73/71	26.4 (5.45)	24.4 (6.36)	-1.79 (0.519)	0.41 (0.655) [-0.87, 1.70]	0.527 (0.336)

Gruppe	n BL/ Woche 48	M (±SD) BL (sec)	M (±SD) in Woche 48 (sec)	LSM (± SE) Woche 48 (sec)	Δ vs. Placebo, M (±SE) [95% KI] (sec)	p-Wert vs. Placebo (p-Wert Interak- tion)
<i>Änderung des transformierten NSAA-Gesamtscores: 6MWD &lt;300 m (Studienbericht 020: Tabelle 14.2.3.7, Tabelle 14.2.3.8)</i>						
Placebo	21/19	38.4 (11.54)	24.8 (15.65)	-12.04 (2.476)	-	-
Ataluren	24/19	39.3 (7.26)	29.3 (14.82)	-11.38 (2.297)	0.65 (3.183) [-5.59, 6.89]	0.837
<i>Änderung des transformierten NSAA-Gesamtscores: 6MWD ≥300 bis &lt;400 m (Studienbericht 020: Tabelle 14.2.3.7, Tabelle 14.2.3.8)</i>						
Placebo	52/49	56.8 (12.64)	47.0 (17.72)	-8.97 (1.603)	-	-
Ataluren	47/45	58.3 (12.58)	52.3 (13.88)	-4.71 (1.652)	4.26 (2.087) [0.17, 8.35]	<b>0.041</b>
<i>Änderung des transformierten NSAA-Gesamtscores: 6MWD ≥400 m (Studienbericht 020: Tabelle 14.2.3.7, Tabelle 14.2.3.8)</i>						
Placebo	41/40	75.7 (13.11)	71.5 (16.34)	-3.26 (1.686)	-	-
Ataluren	43/42	75.9 (12.71)	71.4 (14.09)	-4.32 (1.691)	-1.06 (2.244) [-5.46, 3.34]	0.637 <b>(0.207)</b>
<i>Änderung des transformierten NSAA-Gesamtscores: 6MWD &lt;350 m (Studienbericht 020: Tabelle 14.2.3.11, Tabelle 14.2.3.12)</i>						
Placebo	41/39	45.1 (13.16)	31.2 (16.52)	-12.79 (1.801)	-	-
Ataluren	41/35	44.6 (10.38)	36.0 (15.00)	-9.44 (1.778)	3.25 (2.292) [-1.14, 7.85]	0.143
<i>Änderung des transformierten NSAA-Gesamtscores: 6MWD &lt;350 m (Studienbericht 020: Tabelle 14.2.3.11, Tabelle 14.2.3.12)</i>						
Placebo	73/69	68.7 (15.21)	64.0 (18.49)	-4.46 (1.324)	-	-
Ataluren	73/71	70.1 (14.50)	65.5 (15.42)	-4.02 (1.345)	0.45 (1.693) [-2.87, 3.76]	0.792 (0.305)

ANCOVA mit multipler Imputation; Faktoren: Stratifizierungsfaktoren Alter, BL-6MWD-Gruppe (<300, ≥300 bis <400, ≥400 m; <350 und ≥350 m bei Subgruppe nach Stratifizierungsfaktor), Kortikosteroid-Vorbehandlung, Behandlung, BL-NSAA Gesamtscore (nicht transformiert oder transformiert), Interaktion zwischen Behandlung und Subgruppe.

Die Aktivität, den Kopf im Liegen zu heben, wurde in der Analyse der transformierten Werte nicht eingeschlossen. Bei Patienten, die ihre Gehfähigkeit verloren hatten, wurde ein Wert von 0 eingegeben.

## Fazit

Die Analyse des NSAA bestätigt die Ergebnisse der vorherigen Tests und zeigt einen moderaten, nicht signifikanten Behandlungseffekt von Ataluren in der heterogenen Gesamtpopulation und einen signifikanten Effekt bei den Patienten mit einer initialen Gehstrecke zwischen ≥300 und <400 m. Der Test auf Interaktion zwischen den 6MWD-Subpopulationen lag allerdings oberhalb des Schwellenwerts von  $p < 0.2$ , sodass die Gesamtpopulation zur Bewertung herangezogen werden muss. Selbst in der Subpopulation mit einer BL-6MWD von ≥300 bis <400 m blieb der Unterschied zwischen Ataluren und Placebo (4.26 Punkte) unterhalb der von Mayhew et al. (2013) aufgrund einer verteilungsbasierten Methode vorgeschlagenen mittleren MCID von 8.8 bzw. 6.9 Punkten (transformierte Daten) (25).

Bei der Subpopulation mit einer Gehstrecke von 400 m oder mehr änderte sich der Score auch in der Placebo-Gruppe nur wenig, sodass bei diesen Patienten ein Behandlungseffekt über 48 Wochen kaum nachweisbar war. Bei den Patienten mit bereits deutlichen Einschränkungen (6MWD <300 m) verschlechterte sich der transformierte NSAA-Score demgegenüber unabhängig von der Behandlung deutlich. Somit spiegelt auch die Analyse des NSAA in ihrer Ausrichtung die Ergebnisse der anderen Tests wider und zeigt dieses Testsystem in der Subpopulation mit einer Gehstrecke von  $\geq 300$  bis <400 m über die Dauer der Studie auch am sensitivsten anspricht.

#### **4.3.1.3.2.4 Gesundheitsbezogene Lebensqualität (PRO) – RCT**

##### **Studie PTC124-GD-020-DMD**

##### **Subgruppen-Analysen - RCT**

Es ergaben sich keine Hinweise oder Belege für unterschiedliche Effekte zwischen den BL-6MWD-Subgruppen ( $p > 0.2$  Interaktion Behandlung-Subgruppe). Die Einschätzung der Patienten oder Eltern/Betreuer hinsichtlich ihrer Gesundheitsbezogenen Lebensqualität nahm erwartungsgemäß über die 6MWD-Gruppen sukzessive von denjenigen mit guter Gehfähigkeit ( $\geq 400$  m) bis zu denjenigen mit schon eingeschränktem Gehvermögen (<300 m) ab. Während die Patienten (oder deren Eltern/Betreuer) mit einer Gehstrecke von 400 m oder mehr sich zu Beginn der Studie in ihren Aktivitäten noch relativ wenig eingeschränkt fühlten (91.0 von 100 Punkten; Domäne Transfer/Basismobilität), berichteten die Patienten oder die Eltern/Betreuer von Patienten mit einer Gehstrecke unter 300 m bereits von deutlichen Einschränkungen bei ihren Alltagsaktivitäten. Ein ähnlicher Trend zeigte sich hinsichtlich der Domäne Sport/körperliche Funktion, wobei die Ausgangswerte wahrscheinlich aufgrund der höheren Belastung verglichen mit der Domäne Transfer/Basismobilität über alle 6MWD-Gruppen niedriger (schlechter) waren. Die Subpopulation mit einer BL-6MWD von  $\geq 300$  bis <400 m nahm auch aus Sicht der Patienten/Betreuer eine Sonderstellung ein, bei denen die Gesundheitsbezogene Lebensqualität über 48 Wochen nicht weitgehend unverändert blieb oder sich nur wenig änderte (6MWD  $\geq 400$  m) und bei denen diese über die Dauer der Studie zwar messbar aber weniger deutlich als bei Patienten mit schon ausgeprägten Einschränkungen beim Gehen abnahm (BL-6MWD <300 m; Tabelle 4-44).

Die Effekte von Ataluren waren in der Subpopulation mit der mittleren Ausgangsgehstrecke ( $\geq 300$  bis <400 m) tendenziell etwas höher als in den anderen 6MWD-Gruppen (Tabelle 4-44). In den Analysen stratifiziert nach BL-6MWD-Gruppe ergaben sich hinsichtlich beider PODCI-Domänen weder bei den Patienten mit einer Gehstrecke unter 350 m noch bei den Patienten mit einer Gehstrecke von  $\geq 350$  m signifikante Unterschiede zwischen Ataluren und Placebo. Bei Patienten mit einer Gehstrecke unter 350 m zeigte sich hinsichtlich der Domäne Sport/körperliche Funktion ein positiver Trend zugunsten von Ataluren ( $p = 0.059$ ) (Daten nicht gezeigt).

Tabelle 4-44: Änderung des PODCI von BL bis Woche 48 nach BL-6MWD-Gruppe (ITT) für PTC124-GD-020-DMD; Ataluren versus Placebo (8)

Gruppe	n BL/ Woche 48	M (±SD) BL (Score)	M (±SD) Woche 48 (Score)	Δ LSM (± SE) in Woche 48 (Score)	Δ LSM (±SE) vs. Placebo [95% KI] (Score)	p-Wert vs. Placebo (p-Wert Interaktion)
<i>Änderung des PODCI Transfer/Basismobilität: BL-6MWD &lt;300 m (Studienbericht 020: Tabelle 14.2.4.1.3, Tabelle 14.2.4.1.4)</i>						
Placebo	21/20	67.1 (16.45)	45.5 (26.91)	-20.32 (3.308)	-	-
Ataluren	24/23	71.9 (16.85)	50.7 (27.89)	-21.18 (3.017)	-0.86 (4.210) [-9.12, 7.39]	0.837
<i>Änderung des PODCI Transfer/Basismobilität: BL-6MWD ≥300 bis &lt;400 m (Studienbericht 020: Tabelle 14.2.4.1.3, Tabelle 14.2.4.1.4)</i>						
Placebo	52/50	79.6 (14.93)	71.0 (21.89)	-8.14 (2.155)	-	-
Ataluren	47/44	83.5 (10.30)	79.2 (14.81)	-4.60 (2.251)	3.54 (2.840) [-2.03, 9.11]	0.213
<i>Änderung des PODCI Transfer/Basismobilität: BL-6MWD ≥400 m (Studienbericht 020: Tabelle 14.2.4.1.3, Tabelle 14.2.4.1.4)</i>						
Placebo	41/40	91.0 (8.84)	87.6 (12.21)	-3.59 (2.276)	-	-
Ataluren	43/42	91.0 (7.48)	89.5 (9.56)	-1.53 (2.278)	2.06 (3.045) [-3.90, 8.03]	0.498 (0.615)
<i>Änderung des PODCI Sport/körperliche Funktion: BL-6MWD &lt;300 m (Studienbericht: Tabelle 14.2.4.2.3, Tabelle 14.2.4.2.4)</i>						
Placebo	21/20	42.0 (24.71)	26.1 (21.50)	-13.74 (3.563)	-	-
Ataluren	24/23	44.9 (19.44)	33.0 (17.29)	-10.92 (3.259)	2.83 (4.534) [-6.06, 11.71]	0.533
<i>Änderung des PODCI Sport/körperliche Funktion: BL-6MWD ≥300 bis &lt;400 m (Studienbericht: Tabelle 14.2.4.2.3, Tabelle 14.2.4.2.4)</i>						
Placebo	52/50	53.8 (18.58)	43.3 (22.86)	-9.66 (2.314)	-	-
Ataluren	47/44	54.3 (15.71)	48.9 (20.17)	-5.17 (2.432)	4.49 (3.051) [-1.49, 10.47]	0.141
<i>Änderung des PODCI Sport/körperliche Funktion: BL-6MWD ≥400 m (Studienbericht: Tabelle 14.2.4.2.3, Tabelle 14.2.4.2.4)</i>						
Placebo	41/40	66.0 (16.77)	65.6 (20.50)	-0.49 (2.451)	-	-
Ataluren	43/42	64.5 (18.43)	63.3 (20.91)	-1.37 (2.457)	-0.89 (3.285) [-7.32, 5.55]	0.787 (0.464)

ANCOVA mit multipler Imputation; Faktoren: Stratifizierungsfaktoren Alter, BL-6MWD-Gruppe (<300, ≥300 bis <400, ≥400 m), Kortikosteroid-Vorbehandlung, Behandlung, BL-PODCI; Interaktion zwischen Behandlung und Subgruppe.

### Fazit

Der PODCI wird als valides und klinisch relevantes Instrument bei DMD eingestuft, wobei nicht geklärt ist, ob es auf eine Behandlung anspricht (24).

Der Behandlungseffekt von Ataluren ermittelt mittels PODCI erreichte in keiner der Analysen verglichen mit Placebo das Signifikanzniveau und die Interaktionstests lieferten keine Hinweise auf zwischen den BL-6MWD-Gruppen unterschiedliche Effekte. In der Subpopulation mit einer Gehstrecke von ≥300 bis <400 m war der Behandlungseffekt in Überein-

stimmung mit der Mehrzahl der übrigen Testsysteme etwas deutlicher, was verglichen mit den anderen 6MWD-Gruppen wahrscheinlich auf eine höhere Sensitivität gegenüber dem Testsystem zurückzuführen ist. Hinsichtlich der Domäne Sport/körperliche Funktion erreichte der Unterschied zwischen Ataluren und Placebo (4.49 Punkte) bei der Subpopulation mit einer Gehstrecke von  $\geq 300$  bis  $< 400$  m nahezu die von Henricson et al. (2013) vorgeschlagene MCID von 4.5 Punkten. Zur Schätzung der MCID wurde eine indirekte Methode (1/3 SD) eingesetzt (20), sodass Unsicherheiten hinsichtlich der Validität der MCID bestehen (3). Die Analyse des PODCI hat keinen Einfluss auf die Bewertung des Ausmaßes des Zusatznutzens von Ataluren.

#### 4.3.1.3.3 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.*

Für Arzneimittel zur Behandlung seltener Leiden müssen die Abschnitte 4.3.1 und 4.3.2 nicht ausgefüllt werden (17). Das Ausmaß des Zusatznutzens von Ataluren wird aus den Meta-Analysen der Studien PTC124-GD-007-DMD (siehe Nutzenbewertungsverfahren zum Wirkstoff Ataluren, Beschlussfassung: 21.05.2015) und PTC124-GD-020-DMD abgeleitet, deren Ergebnisse in Kapitel 4.4.2 tabellarisch zusammengefasst werden.

### 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 4-45: 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 4.3.1.3.1.1.**



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

Tabelle 4-46: 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 4-47: Operationalisierung von <Endpunkt xxx>

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 4-48: 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 4-49: Ergebnisse für &lt;Endpunkt xxx&gt; aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 4.3.1.3.1.1)
<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 4.3.1.3.2.*

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*

Nicht zutreffend.

##### **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 4-50: 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 4-51: Operationalisierung von <Endpunkt xxx>

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 4-52: Verzerrungsaspekte für <Endpunkt xxx> – 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 4.3.1.3.2.**

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 4-53: 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.

#### **4.3.2.3.2 Subgruppenanalysen – weitere Untersuchungen**

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

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.*

Die Resultate der Einzelstudien PTC124-GD-007-DMD und PTC124-GD-020-DMD entsprechen der Evidenzstufe Ib der in der VerFO-GBA dargestellten Evidenzhierarchie. Die Studienqualität ist in beiden Fällen als sehr gut anzusehen. Da grundlegende Kriterien wie die Verblindung, adäquate Randomisierungsmethoden und eine Sicherstellung der verdeckten

Zuteilung der Patienten zu den Studienarmen in den beiden Studien erfüllt sind, kann das Verzerrungspotential im Fall beider Studien mit niedrig bewertet werden.

In den zur Nutzenbewertung relevanten Endpunkten zeigte sich kein Hinweis auf Heterogenität zwischen den Studien, so dass hier Meta-Analysen der Evidenzstufe Ia berechnet werden konnten.

#### **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**

DMD ist eine schwere, progredient verlaufende, X-chromosomal vererbte Erkrankung, die durch Mutationen im Dystrophin-Gen verursacht wird. Bei etwa 10% der Patienten kann eine Nonsense-Mutation identifiziert werden, die in der mRNA ein vorzeitiges Stop-Codon erzeugt. Dieses Stop-Codon führt zum vorzeitigen Abbruch der Protein-Translation, sodass kein funktionales Dystrophin in der Skelettmuskulatur, der glatten Muskulatur, und Herzmuskulatur gebildet wird. Im Verlauf der Erkrankung wird mehr und mehr Muskelmasse



durch Fett und fibröses Gewebe ersetzt, der Verlust des gesunden Muskelgewebes ist irreversibel. Die Erkrankung führt zu einem raschen Verlust der körperlichen Funktionen mit nachfolgendem Atem- und Herzversagen sowie Tod im jungen Erwachsenenalter (üblicherweise <30 Jahre) (u.a. Kole, Leppert 2012 zitiert nach (53)).

Translarna (Ataluren) erhielt im Jahr 2014 eine auflagengebundene europäische Zulassung (Zulassungsnummer: EU/1/13/902/003), die den pharmazeutischen Unternehmer zum Abschluss einer konfirmatorischen Studie verpflichtete (1). Zulassungsinhaberin ist PTC Therapeutics International Limited, Irland. Translarna ist zur Behandlung der Duchenne-Muskeldystrophie infolge einer Nonsense-Mutation im Dystrophin-Gen bei gehfähigen Patienten im Alter ab fünf Jahren zugelassen. Bei nicht gehfähigen Patienten wurde keine Wirksamkeit nachgewiesen. Das Vorliegen einer Nonsense-Mutation im Dystrophin-Gen ist durch Gentest nachzuweisen. Durch Ataluren wird das vorzeitige Stop-Codon überlesen, sodass die Translation bis zu dem normalen Stop-Codon fortschreiten kann und funktionales Dystrophin entsteht.

Zur kurativen Therapie der DMD stand bis zur Zulassung von Ataluren kein Arzneimittel zur Verfügung. Die Behandlung der DMD bestand zuvor aus symptomatische Maßnahmen zur Unterstützung der kardialen und respiratorischen Funktionen, sowie medikamentösen Interventionen, insbesondere Kortikosteroide zur Verbesserung der Muskelstärke und -funktion. Orthopädische Korrekturen und physiotherapeutische Interventionen waren und sind ebenfalls Teil der Therapie (32). Kortikosteroide sind in dieser Indikation in Europa nicht zugelassen. Sie haben z.T. signifikante Nebenwirkungen und hinsichtlich der/des optimalen Dosisregimes besteht kein Konsens (2). 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 (6). 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 (6). Der Verlust der Gehfähigkeit ist als schwerwiegendes Symptom der DMD für die in Frage stehende Population anzusehen, zumal die Gehfähigkeit Voraussetzung der Anwendung von Ataluren ist. Der Verlust der Gehfähigkeit führt deterministisch zu schwerwiegenden Folgekomplikationen. Diese Schlussfolgerung wird z.B. durch eine Publikation von Humbertclaude et al. (2012) bei Kindern mit DMD gestützt (33).

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 (6). Für die Bewertung des Ausmaßes des Zusatznutzens von Ataluren wurden die patientenrelevanten Zielgrößen Morbidität (Nutzendimensionen: Verbesserung des Gesundheitszustands [Mobilität und körperliche Funktion]) sowie die Verträglichkeit herangezogen. Die Zielgröße Mortalität konnte in den vorliegenden 48-Wochen-Studien nicht evaluiert werden, da das zugelassene Anwendungsgebiet die nicht-ambulatorische Phase nicht mit einschließt. Die Gesundheitsbezogene Lebensqualität wurde in beiden Studien erfasst, allerdings mit unterschiedlichen Instrumenten. Tabelle 4-54 gibt einen Überblick der in dieser Nutzenbewertung berücksichtigten patientenrelevanten Endpunkte. Der 6MWT, der in beiden Ataluren-Studien als primärer Endpunkt erhoben wurde, 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 4-54: Darstellung der patientenrelevanten Endpunkte

Patientenrelevante Zielgrößenkategorie	Endpunktparameter
Krankheitsbedingte Symptome (Morbidität)	<ul style="list-style-type: none"> <li>– Gehstrecke im 6MWT bis Woche 48</li> <li>– Zeit bis zu einer mindestens 10%igen Verschlechterung der Gehstrecke im 6MWT</li> <li>– TFTs: Zeit (Maximum 30 sec), die benötigt wird, um <ul style="list-style-type: none"> <li>• vier Treppenstufen hinauf oder hinunter zu gehen</li> <li>• 10 m zu gehen oder zu laufen</li> </ul> </li> <li>– Körperliche Funktion mittels NSAA<sup>1</sup> inklusive Zeit, die benötigt wird, um aus der Rückenlage aufzustehen</li> <li>– Aktivitäten des täglichen Lebens und Krankheitssymptome (PRO)<sup>1</sup></li> </ul>
Gesundheitsbezogen Lebensqualität	– PODCI <sup>1</sup>
Verträglichkeit	– Nicht schwerwiegende (bzw. schwerwiegende) unerwünschte Ereignisse

<sup>1</sup> Das NSAA und PODCI wurden nur in der hier gegenständlichen Studie 020 erfasst und das Instrument zur Erfassung von Alltagsaktivitäten unterschied sich zwischen den Studien deutlich. Diese Endpunkte konnten daher nicht meta-analytisch ausgewertet werden.

Das Ausmaß des Zusatznutzens von Ataluren in der zugelassenen Indikation bei Patienten mit nmDMD wurde bereits durch den G-BA auf Basis der vorausgegangenen randomisierten, Placebo-kontrollierten Phase 2b-Studie mit einem geringen Zusatznutzen bewertet (siehe veröffentlichtes Nutzendossier (38)). Zwischenzeitlich liegen die Ergebnisse einer confirmatorischen Placebo-kontrollierten Phase 3 Studie vor, deren Design und Ergebnisse in Kapitel 4.3.1 präsentiert wurden. Die Aussagen zum Ausmaß des Zusatznutzens von Ataluren werden aus den Meta-Analysen der beiden Studien abgeleitet.

In den Meta-Analysen wurden folgende patientenrelevanten Endpunkte ausgewertet:

- Gehstrecke im 6MWT (s. Kapitel 4.3.1.3.1.1),
- Zeit bis zur 10%igen Verschlechterung im 6MWT (s. Kapitel 4.3.1.3.1.1) und
- TFTs (4 Treppenstufen hinauf- oder hinunter gehen, 10 m gehen/laufen) (s. Kapitel 4.3.1.3.1.2)
- Verträglichkeit (s. Kapitel 4.3.1.3.1.6)

Die Nebenwirkungsraten beider Studien wurden jeweils zusammengefasst in Meta-Analysen ausgewertet. Weder die Ergebnisse der Studie 007 (siehe frühere Nutzenbewertung) noch der Studie 020 (siehe Kapitel 4.3.1.3.1.6) erlauben eine Quantifizierung des Ausmaßes des Zusatznutzens von Ataluren hinsichtlich der Gesundheitsbezogenen Lebensqualität. Eine meta-analytische Auswertung war aufgrund der zwischen den Studien unterschiedlichen Messinstrumente nicht möglich. Dies trifft auch auf die Aktivitäten des täglichen Lebens zu. Das NSAA wurde nur in der Studie 020 eingesetzt. 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 Aussagekraft der Nachweise

Bei beiden zur Nutzenbewertung herangezogenen Studien (siehe separates Dossier zur Nutzenbewertung für Studie 007) handelt es sich um randomisierte, Placebo-kontrollierte Doppelblindstudien der Evidenzstufe 1b. Die *a priori* festgelegten Endpunkte wurden adäquat operationalisiert und das Verzerrungspotential wurde auf Studienebene und auf der Endpunktebene als niedrig eingestuft, so dass von einer hohen qualitativen Ergebnissicherheit ausgegangen werden kann. Da es sich bei den Meta-Analysen um eine systematische Übersicht von Studien der Evidenzstufe 1b handelt, ergibt sich laut VerFO in der Gesamtschau die Evidenzstufe 1a (6). In die Meta-Analyse wurden alle mit Ataluren durchgeführten RCTs eingeschlossen (PTC124-GD-007-DMD und PTC124-GD-020-DMD).

Entsprechend den „Allgemeinen Methoden - Version 4.2“ des IQWiG kann aus einer Meta-Analyse von mindestens zwei homogenen Studien mit hoher qualitativer Ergebnissicherheit ein Beleg abgeleitet werden. Im Fall homogener Ergebnisse muss der gemeinsame Effektschätzer statistisch signifikant sein, um gemäß der vorhandenen Ergebnissicherheit einen Beleg, einen Hinweis oder einen Anhaltspunkt zu liefern (7). Voraussetzung ist, dass die Studien sinnhaft zusammengefasst werden können und nicht möglichen Verzerrungen unterliegen (7,37).

Da für die vorliegende Nutzenbewertung Subgruppenanalysen eine wesentliche Rolle spielen, wird im Folgenden kurz auf diese Thematik eingegangen. Subgruppenanalysen werden in der methodischen Literatur oftmals kritisch diskutiert. Gibt es allerdings *a priori* Kenntnis von einem möglichen Effektmodifikator, ist es laut IQWiG - Allgemeine Methoden sogar erforderlich, eine mögliche Heterogenität bezüglich des Effekts in den verschiedenen Patientengruppen vorab zu untersuchen. Lässt sich diese nachweisen, ist der über alle

Patienten hinweg geschätzte Gesamteffekt nicht sinnvoll interpretierbar (7). Eine Effektmodifikation durch die Baseline-Gehstrecke im 6MWT war auf Grund der Ergebnisse der Studie PTC124-GD-007-DMD schon bekannt. Die Effektmodifikation durch die Baseline-Gehstrecke beruht jedoch nicht direkt auf einer unterschiedlichen Wirksamkeit von Ataluren in Abhängigkeit der Baseline-Gehstrecke, sondern auf der unterschiedlichen Sensitivität des 6MWT, einen Behandlungseffekt bei Patienten in bestimmten Phasen der Krankheitsprogression messen zu können.

Je nach Fragestellung können im Rahmen systematischer Auswertungen Subgruppen-Analysen notwendig sein. Das IQWiG interpretiert im Rahmen von Meta-Analysen die Ergebnisse eines Heterogenitäts- oder Interaktionstests bezüglich wichtiger Subgruppen in der Regel wie folgt: Ein zum Niveau  $\alpha = 0.05$  signifikantes Ergebnis wird als Beleg unterschiedlicher Effekte, ein zum Niveau  $\alpha = 0.20$  signifikantes Ergebnis wird als Hinweis auf unterschiedliche Effekte in den Gruppen gewertet. Liegt ein Beleg für unterschiedliche Effekte in den Subgruppen vor, so werden die Ergebnisse der Subgruppen separat dargestellt. Die in diesem Zusammenhang besonders relevante Subgruppen-Analyse (BL-6MWD <300 m,  $\geq 300$  bis <400 m und  $\geq 400$  m) war im SAP der Studie 020 prospektiv definiert.

#### 4.4.2.3 Validität und Patientenrelevanz der Endpunkte

Die Validität und Patientenrelevanz der meisten Endpunkte der Studie 020 sowie aller Endpunkte der Meta-Analyse wurden bereits in der vorausgegangenen Nutzenbewertung der Ataluren-Studie 007 durch den G-BA bewertet (3).

Als Ergebnis eines Meetings im Jahre 2014 (TREAT-NMD), an dem international ausgewiesene Experten, Patienten und Vertreter der Industrie teilnahmen, wurde die Eignung und Patientenrelevanz verschiedener Instrumente als Endpunkte in Phase 2/3-Studien bei Patienten mit DMD bewertet (24). Tabelle 4-55 zeigt den Status der hier relevanten Endpunkte, soweit sie bewertet wurden. Das NSAA, die TFTs und der 6MWT erfüllten alle Kriterien.

Tabelle 4-55: Entscheidungshilfen zum Einschluss klinischer Endpunkte in Phase 2/3-Studien bei DMD (24)

Endpunkt / Status	NSAA	TFTs	6MWT	PODCI
Subgruppe (Alter)	3.5 Jahre bis gehunfähig	4 Jahre bis gehunfähig	5 Jahre bis gehunfähig	3-21 Jahre
Konzeption passt zu DMD	√	√	√	√
Reliabilität	√	√	√	√
Validierung mit anderen Endpunkten	√	√	√	√
Normative Daten	√	√	√	√

<b>Endpunkt / Status</b>	<b>NSAA</b>	<b>TFTs</b>	<b>6MWT</b>	<b>PODCI</b>
Fortlaufende Studien zum natürlichen Verlauf	√	√	√	√
Multizenter-Studien	√	√	√	√
Ansprechen auf Behandlung	√	√	√	?
Klinisch sinnvoll	√	√	√	√

Da der molekulare Defekt der nmDMD direkt die Muskulatur betrifft (Skelettmuskulatur, Atemmuskulatur, Herzmuskel, etc.), stellt die Gehfähigkeit ein direktes Maß für die Schwere der Erkrankung bzw. für die deterministische Krankheitsprogression dar. Die verminderte Gehfähigkeit von DMD-Patienten ist eines der am frühesten auftretenden Symptome und ist als sensitives Maß der Krankheitsprogression zu werten. Die Krankheitsprogression bei 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).

Die EMA stuft eine Verbesserung der körperlichen Funktion oder zumindest eine Verzögerung der Krankheitsprogression als den wichtigsten Endpunkt bei Patienten mit DMD ein. Auch der Erhalt täglicher Aktivitäten wird von der EMA als wichtiges Therapieziel eingeschätzt (2). Auch das NICE kommt in einer Evaluation von Ataluren zu dem Schluss, dass die Fähigkeit zu gehen für Kinder mit DMD und deren Angehörigen von außerordentlicher Bedeutung ist (53). Die Validität des 6MWT als primärer Endpunkt wurde in einer Reihe internationaler klinischer Multizenter-Studien und in longitudinalen Studien zum natürlichen Verlauf bei DMD bestätigt (27,30,54,55). Der 6MWT wird als sensitiver eingestuft, um die Krankheitsprogression zu erfassen, als Tests zur Muskelstärke oder TFTs (24). Der 6MWT ist bei DMD validiert und korreliert mit den relevanten Pathomechanismen, der Krankheitsprogression, der Muskelstärke, der Ausdauer während kontinuierlicher Aktivität und den grobmotorischen Fähigkeiten. Bei DMD-Patienten mit einer Gehstrecke von mindestens 330 m im 6MWT ist das Risiko, innerhalb von zwei Jahren die Gehfähigkeit zu verlieren, gering. Die FDA schätzt in einem aktuellen Richtlinien-Entwurf („Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment: Guidance for Industry“, Juni 2015) den 6MWT oder kürzere Test-Versionen als geeignet ein, um bei Kindern ab einem Alter von 5-6 Jahren, die Stärke und Ausdauer zu erfassen (29). Auch die EMA kommt in der aktuellen Richtlinie zu dem Schluss, dass der 6MWT ein bei Kindern ab fünf Jahren validiertes Instrument ist, welches die für gehfähige DMD-Patienten relevanten Aspekte der Muskelfunktion und Ausdauer erfasst (2). Bei älteren Patienten stellt der „floor“ Effekt eines Verlustes der Gehfähigkeit allerdings eine Herausforderung dar. Die

FDA merkt an, dass die Analyse des 6MWT sehr stark durch den Ein- oder Ausschluss von Patienten beeinflusst wird, die während der Studie ihre Gehfähigkeit verlieren (29). Probleme können laut EMA z.B. der Einfluss des Alters zu Studienbeginn, Lerneffekte oder eine Interferenz mit dem Wachstum darstellen (2). Studien liefern allerdings keine Hinweise auf einen Lern- oder Placeboeffekt im 6MWT, welcher die Krankheitsprogression erfasst (24). In der früheren Bewertung des Nutzendossiers zu Ataluren bewertet der G-BA die Patientenrelevanz des 6MWT als Surrogat für „Meilensteine“ des Funktionsverlusts entsprechend der Expertenempfehlungen als nachvollziehbar (3), welche jüngst (TREAT-NMD-Meeting 2014) bestätigt wurden (24).

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 (19). Eine dauerhafte 10%ige Verschlechterung der 6MWD über ein Jahr ist klinisch relevant und ein Prognosefaktor für den Verlust der Gehfähigkeit (19,54). Die Zeit bis zur 10%igen Verschlechterung im 6MWT ist somit ein direktes Maß für die Veränderung der schwerwiegenden Symptome der Erkrankung. Der Endpunkt „Zeit bis zur 10%igen Verschlechterung im 6MWT“ wurde seitens des G-BA als patientenrelevant für Patienten mit nmDMD eingestuft (3).

Laut FDA stellen TFTs und das NSAA nützliche Instrumente dar, um die motorische Funktion bei gehfähigen Kindern im Alter zwischen vier und 7 Jahren zu erfassen (29). TFTs werden auch in aktuellen Studien häufig eingesetzt, da sie laut EMA relevante Informationen liefern (2,56). TFTs wurden während des oben angesprochenen Meetings (TREAT-NMD) als valides und bei DMD klinisch sinnvolles Instrument mit guter Reliabilität eingestuft (Tabelle 4-55) (24). Patienten, die  $\geq 12$  sec benötigten, um 10 m zu gehen oder zu laufen, hatten ein etwa 90%iges Risiko, ihre Gehfähigkeit im Folgejahr zu verlieren (21,57). Eine Zeit unter 6 sec im 10 m-Gehtest war mit einem Erhalt der Gehfähigkeit über die nächsten 12 Monate assoziiert, während Patienten, die  $>10$ -12 sec benötigten, ein hohes Risiko hatten, ihre Gehfähigkeit in den nächsten 12 Monaten zu verlieren (19). Basierend auf den Eingangsdaten der Studie 007 und unter Einsatz verteilungsbasierter Methoden ( $SD * \sqrt{1-r}$  bzw.  $1/3 SD$ ) wurden für die TFTs MIDs geschätzt: Aufstehen aus dem Liegen 3.7 bzw. 3.6 sec, 4 Stufen hinauf: 2.1 bzw. 2.2 sec, 4 Stufen hinunter: 2.1 bzw. 2.2 sec, 10 m gehen/laufen: 2.3 bzw. 1.4 sec (18). 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 (21,58).

Das NSAA ist ein oft eingesetztes Instrument, um die Krankheitsprogression bei DMD zu erfassen und wird zunehmend in klinischen Studien bei DMD als Endpunkt eingesetzt (24,51,59,60). Laut EMA kann das NSAA bei gehfähigen Jungen mit DMD eingesetzt werden (2). Das NSAA wurde während des oben angesprochenen Meetings (TREAT-NMD) als valides und bei DMD klinisch sinnvolles Instrument mit guter Reliabilität eingestuft (Tabelle 4-55) (24). In einer Studie bei 191 gehfähigen Jungen mit DMD erwies sich das NSAA mittels Rasch-Analyse als reliables und valides Instrument. Eine Ausnahme bildete die

Fähigkeit, den Kopf im Liegen zu heben, und es zeigten sich Abhängigkeiten zwischen wenigen Tests (50). Das NSAA ist ein bei DMD-Patienten nützliches Instrument zur Erfassung der körperlichen Fähigkeiten. Zwischen dem NSAA und dem 6MWT konnte über zwei Jahre eine lineare Beziehung gezeigt werden, was die klinische Bedeutung beider Tests unterstützt. Das NSAA zusammen mit dem 6MWT liefern Information über verschiedene Aspekte der motorischen Funktion (54). Das NSAA erfasst verglichen mit dem 6MWT ein breiteres Spektrum an Funktionen, die tägliche Alltagsaktivitäten widerspiegeln (27). In einer longitudinalen Studie wurden 198 Jungen mit DMD (4-18 Jahre alt) eingeschlossen, die stabil Kortikosteroide erhielten (täglich oder intermittierend) und deren Fähigkeiten regelmäßig mittels NSAA überwacht wurden. Es wurden die Rasch-transformierten Daten des NSAA analysiert (Skala von 0-100). Mittels NSAA konnten Unterschiede zwischen den beiden Kortikosteroid-Regimen detektiert werden und die Wirksamkeit änderte sich mit dem Alter. Die Fähigkeiten nahmen in einer zeitlichen Reihenfolge beginnend mit der Fähigkeit, auf einem Bein zu hüpfen, bis zur Fähigkeit, zu stehen, ab (Abbildung 64). In einer großen Studie, in die 513 gehfähige Jungen mit DMD eingeschlossen wurden, nahm der lineare NSAA-Score bei Kindern im Alter von 7 Jahren und älter im Mittel um 8 Punkte pro Jahr ab, während der lineare Score bei jüngeren Patienten im Mittel um 4.0 Punkte zunahm (51).

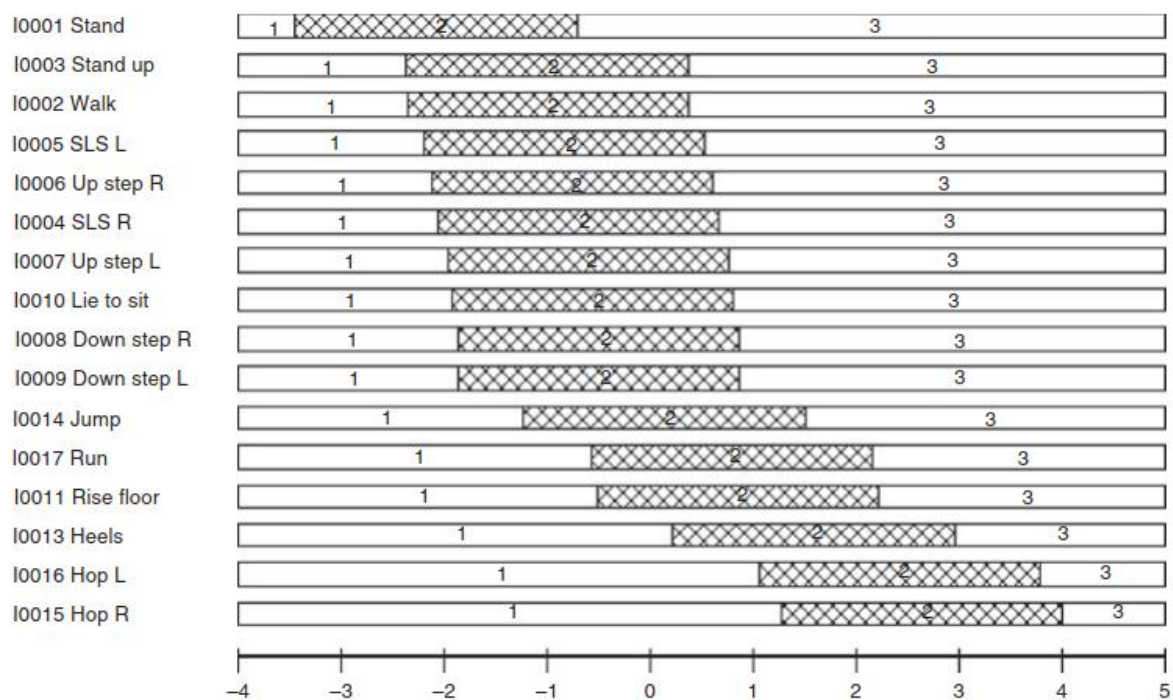


Abbildung 64: Lineare NSAA-Skala bei DMD (25)

Die x-Achse zeigt die körperlichen Fähigkeiten, wobei die Fähigkeiten von links nach rechts zunehmen. Die y-Achse zeigt die Response-Kategorien: 1 = unfähig, 2 = fähig mit Modifikationen, 3 = fähig.

Der G-BA schätzt die in der Ataluren-Studie 007 erfasste Änderung der Alltagsaktivitäten gemessen mittels ‚Step Activity Monitoring‘ als für die Nutzenbewertung relevant ein, da die Partizipation im Alltag für die Betroffenen von großer Bedeutung ist (3). Auch die EMA hebt die Bedeutung des Erhalts von Alltagsaktivitäten hervor (2). Die hohe Relevanz des Erhalts

von Alltagsaktivitäten für DMD-Patienten und deren Familien wurde auch während des TREAT-NMD-Meetings herausgestellt (24). Für das in der Studie 020 eingesetzte Instrument zur Erfassung der Alltagsaktivitäten/Krankheitssymptome liegen jedoch noch keine Daten zur Validität oder zu MIDs vor.

Entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der Verringerung von Nebenwirkungen oder einer *Verbesserung der Lebensqualität*“ (36). Das in der Studie 020 eingesetzte PODCI ist ein Instrument zur Einschätzung der körperlichen Fähigkeiten von Kindern mit orthopädischen Behinderungen aus Sicht der Patienten oder deren Betreuern (20). Das PedsQL-Instrument und PODCI werden häufig eingesetzt, um die gesundheitsbezogene Lebensqualität bei Kindern zu ermitteln. Die Eignung beider Instrumente wurde bei 52 gehfähigen DMD-Patienten (4-15 Jahre) und 36 gesunden Kindern vergleichbaren Alters überprüft (durch die Eltern berichtet). Die Domänen Transfer/Basismobilität und Sport/körperliche Funktion des PODCI (sowie körperliche Funktion im PedsQL) zeigten signifikante Assoziationen mit dem Alter (und entsprechend der Krankheitsprogression) sowie klinischen Endpunkten (z.B. TFTs) bei DMD-Patienten (26). In einer Studie bei 24 Patienten mit DMD im Alter zwischen 4 und 12 Jahren zeigte die 6MWD eine hohe Korrelation mit der PODCI-Domäne Transfer/Basismobilität ( $r=0.79$ ) (20). Die EMA stuft bestimmte Domänen des PODCI als nützliche Endpunkte bei gehfähigen DMD-Patienten ein (2). Das PODCI wurde während des oben angesprochenen Meetings (TREAT-NMD) als valides und bei DMD klinisch sinnvolles Instrument mit guter Reliabilität eingestuft. Allerdings ist unklar, ob das PODCI auf eine Behandlung anspricht (Tabelle 4-55). Das PODCI ist sensitiver und enger mit der Krankheitsprogression assoziiert als das in der Studie 007 eingesetzte Instrument PedsQL (24). Basierend auf den Daten von 24 gehfähigen Kindern mit DMD (4-12 Jahre) wurde mittels einer verteilungsbasierten Methode (1/3 SD) eine MCID von 4.5 Punkten vorgeschlagen (20). Zusammenfassend ist jedoch festzuhalten, dass die Entwicklung von Lebensqualitäts-Instrumenten für pädiatrische DMD Patienten erst am Anfang steht.

Arzneimittelsicherheit: entsprechend § 2 Satz 3 AM-NutzenV ist der Nutzen eines Arzneimittels „der patientenrelevante therapeutische Effekt insbesondere hinsichtlich der Verbesserung des Gesundheitszustandes, der Verkürzung der Krankheitsdauer, der Verlängerung des Überlebens, der *Verringerung von Nebenwirkungen* oder einer *Verbesserung der Lebensqualität*“ (36).

#### **4.4.2.4 Rationale der Analysen der Subpopulationen nach BL-6MWD (<300, ≥300 bis <400, ≥400 m)**

Vor Beginn des klinischen Studienprogramms von Ataluren war der natürliche Verlauf der Änderung der 6MWD bei DMD-Patienten weitgehend unbekannt. Bei der Planung beider randomisierter Ataluren-Studien standen der Intention, eine möglichst homogene Patienten-



population zu rekrutieren, selbstverständlich die regulatorischen Anforderungen gegenüber, Patienten in möglichst allen Krankheitsphasen einzuschließen.

Aus diesen Gründen wurde in die erste RCT mit Ataluren (Studie 007) ein relativ breites und folglich heterogenes Patientenkollektiv eingeschlossen. Nachfolgende und parallele longitudinale Studien und eine Analyse des Placeboarms der Ataluren-Studie 007 zeigten, dass Patienten mit relativ guter initialer Gehfähigkeit mittelfristig stabil bleiben, d.h. ein Behandlungseffekt ist während einer ethisch vertretbaren, einjährigen Studiendauer einer Placebo-kontrollierten Studie kaum nachweisbar (8,19). Andererseits ist ein kritischer Anteil von intaktem Muskelgewebe erforderlich, um die Wirksamkeit einer Dystrophin-induzierenden Therapie zeigen zu können (19). Dementsprechend können Patienten, deren Gehfähigkeit schon deutlich eingeschränkt ist, nicht mehr vollständig von einer Dystrophin-regenerierenden Therapie profitieren, sodass auch bei diesem Kollektiv je nach Testsystem ein Behandlungseffekt nur bedingt oder nicht nachweisbar ist. Zudem erschwert die Heterogenität des Krankheitsverlaufes bei diesen Patienten den Nachweis eines Effekts (8). Zu einem ähnlichen Schluss kommt auch die FDA in einem Richtlinien-Entwurf zur Arzneimittelentwicklung bei DMD und verwandten Dystrophinopathien. Bei Arzneimitteln, die die Krankheitsprogression verzögern aber wahrscheinlich keinen Effekt auf bereits bestehende Muskelschäden haben, kann eine „prognostische Anreicherung“ hilfreich sein (z.B. Selektion von Patienten mit Prädiktoren einer klinischen Progression während der geplanten Studie). Die FDA weist auf „floor“ und „ceiling“ Effekte bei Funktionstests hin, d.h. es sind keine Effekte messbar, weil der Patient entweder im Studienverlauf die Fähigkeit verliert, an den Tests teilzunehmen oder diese zu beenden, oder die körperliche Leistungsfähigkeit bleibt während der Dauer der Studie noch vollständig erhalten (29).

Bei DMD kommt erschwerend hinzu, dass die Gehstrecke im 6MWT bis zum Alter von 12 Jahren zunimmt, insbesondere im Alter zwischen fünf und 8 Jahren (30). Andere Autoren berichten ein Cut-off von 7 Jahren, ab dem sich die Gehstrecke im 6MWT oder der NSAA-Score verschlechtern (57,60). Bei Kindern mit DMD finden der initiale Muskelabbau und die zunehmenden körperlichen Einschränkungen vor dem Hintergrund von Wachstum und Entwicklung statt (30,57). Bei Kindern mit DMD nahmen die Gehstrecken im 6MWT bis zu einem Alter von etwa 7 Jahren über ein Jahr proportional zu und diese entsprachen etwa 80% des normativen Wertes gesunder Kinder (57), was bei den betroffenen Patienten jedoch nicht mit einer verbesserten Muskelfunktion gleichzusetzen ist (60). Dies bedeutet, dass bei diesen Patienten normale Wachstums- und Entwicklungsprozesse nur schwer von einem Behandlungseffekt unterschieden werden können. Bei älteren Kindern nahm die normative 6MWD aufgrund der fortschreitenden Erkrankung wie erwartet über ein Jahr ab (57). Beim Design zukünftiger klinischer Studien sollte daher berücksichtigt werden, dass die Erkrankung in verschiedenen Gruppen in Abhängigkeit vom Alter, der Entwicklung und dem Stadium der Erkrankung unterschiedlich stark voranschreitet, was zu klinisch relevanten Unterschieden der Effektgrößen führen kann (56).

Die FDA favorisiert Wirksamkeitsstudien bei Patienten mit verschiedenen Krankheitsstadien und Phänotypen. Die Arzneimittelentwicklung kann allerdings auf bestimmte Krankheits-

Subgruppen ausgerichtet sein, wenn dies wissenschaftlich begründet ist. Selbst wenn nur wenige Daten für verschiedenen Krankheits-Subgruppen vorliegen, können diese eine breitere Indikation stützen. Ausgenommen hiervon sind begründete Bedenken, dass der Effekt auf eine bestimmte Gruppe begrenzt ist, oder dass das Risiko bei anderen Gruppen unakzeptabel hoch ist (29). Nach der EMA-Richtlinie sollte, wenn möglich auch eine breite Patientenpopulation in konfirmatorische Studien eingeschlossen werden. Auch die EMA empfiehlt prä-spezifizierte Subgruppen-Analysen z.B. nach funktionellem Status (2). Diese Empfehlungen, die zum Zeitpunkt der Planung der Studien noch nicht vorlagen, entsprechen grundsätzlich dem Vorgehen bei den Ataluren-Studien, d.h. in beide Studien wurden gehfähige Patienten mit einem relativ breiten Krankheitsspektrum eingeschlossen. Die Studie 007 hatte etwas weiter gefasste Einschlusskriterien als die Studie 020 (Alter:  $\geq 5$  vs.  $\geq 7$ -16 Jahre, BL-6MWD  $\geq 75$  vs.  $\geq 150$  m, BL-6MWD  $\leq 80\%$  des normativen Wertes nur in Studie 020). Dennoch zeigte sich, dass die Heterogenität der Patientenpopulation in Bezug auf die Endwicklung der Mobilität und Funktionalität als zu groß angenommen werden muss. Die Heterogenität der Studie 020 war trotz geänderter Einschlusskriterien derjenigen der Studie 007 vergleichbar. Die untere BL-6MWD-Grenze war in der Studie 020 höher als in der Studie 007, war aber im Rückblick noch zu niedrig, da ein Teil der Patienten mit schon weit fortgeschrittener Muskelschädigung nicht mehr auf das primäre Testsystem ansprechen konnte. Das Einschlusskriterium der normativen 6MWD (Prozent der vorausgesagten 6MWD) war in der Studie 020 grundsätzlich gerechtfertigt, da dadurch normale Wachstums- und Entwicklungsprozesse berücksichtigt werden können (57). Die gewählte 6MWD-Obergrenze ( $\leq 80\%$  des vorausgesagten Wertes) erscheint im Nachhinein dennoch zu hoch, da sich die Gehfähigkeit von Patienten mit noch relativ gutem Gehvermögen in der Regel auch ohne Behandlung innerhalb von 48 Wochen nicht nennenswert verschlechtert, was die Ergebnisse der Subgruppen in den Studien 007 und 020 und der entsprechenden Meta-Analyse eindrucksvoll belegen. Bei Kindern unter 7 Jahren entspricht eine Gehstrecke von etwa 80% des vorausgesagten Wertes einer absoluten Gehstrecke von 428 m im 6MWT (19). Eine Gehstrecke von 414.2, 385.4 bzw. 359.6 m im 6MWT entsprach bei Jungen mit DMD im Alter von 6.5 bis  $< 7.5$ , 7.5 bis  $< 8.5$  bzw. 8.5 bis  $< 9.5$  Jahren 74.3%, 65.7% bzw. 59.2% des vorausgesagten Wertes nach Geiger (61). Eine absolute vs. relative 6MWD-Obergrenze (z.B. 400 m) wäre unter Umständen wünschenswert gewesen, hätte aber vielleicht dazu geführt, dass die Patienten absichtlich langsamer gehen, um sich für die Studie zu qualifizieren.

Bei der Planung der Ataluren-Studien lagen Hinweise vor, die eine Stratifizierung nach BL-6MWD von  $< 350$  vs.  $\geq 350$  m nahelegten. Unabhängige Studien sowie die Ataluren-Studien selbst liefern Evidenzen, dass diese Unterteilung zu grob ist, und die tatsächlichen Krankheitsstadien nicht adäquat widerspiegelt. Im Protokoll der Studie 020 waren prospektiv Subgruppen-Analysen nach BL 6MWD definiert ( $< 300$ ,  $\geq 300$  bis  $< 400$ ,  $\geq 400$  m bzw.  $< 350$ ,  $\geq 350$  m), die auf Erkenntnissen der Arbeitsgruppe von Sweeney et al. beruhen (62).

Sweeney et al. (62) studierten unter Anwendung von Magnetresonanztomographie in einer natürlichen Verlaufsstudie den Zusammenhang zwischen der Gehfähigkeit im 6MWT und der Degeneration des gesunden Muskelgewebes (Ersatz durch fibrotisches Gewebe und Fettgewebe) in DMD Patienten. Auf Grund der zur Zeit vorliegenden Daten der

Magnetresonanztomographie in klinischen Studien zu DMD sieht die FDA einen vielversprechenden und vermehrten Einsatz dieser Technologie in zukünftigen Studien (63). Sweeney et al. (2014) konnten schon in frühen Krankheitsstadien eine Schädigung des gesunden Muskelgewebes in DMD Patienten nachweisen, während sich die funktionale Performance der Patienten in diesem Stadium noch verbesserte oder zumindest stabil blieb (Abbildung 65). Eine Degeneration des gesunden Muskelgewebes korreliert in frühen Krankheitsstadien somit nicht direkt mit der 6MWD, da der Einfluss der Gewebeschäden auf die Gehfähigkeit erst verzögert auftritt. Während der Ersatz des durch die Erkrankung geschädigten Muskelgewebes durch Fettgewebe im Krankheitsverlauf stetig zunimmt (Abbildung 66), ist der vollständige Verlust der Gehfähigkeit im späten Krankheitsverlauf ein kurzfristiger Prozess, der offensichtlich durch Unterschreiten eines kritischen Schwellenwertes von gesundem Muskelgewebe ausgelöst wird. Ab einem in der Magnetresonanztomographie messbaren Gehalt an Fettgewebe in der Muskulatur (*Musculus vastus lateralis*) von ca. 80% verlieren die Patienten kurzfristig ihre Gehfähigkeit (Abbildung 67). Eine proportionale Korrelation der Abnahme des gesunden Muskelgewebes mit der Abnahme der 6MWD wurde von Sweeney nur in einem Band der BL 6MWD von ca. 300-400 m – also in der mittleren Phase der Krankheitsprogression - nachgewiesen (Abbildung 68).

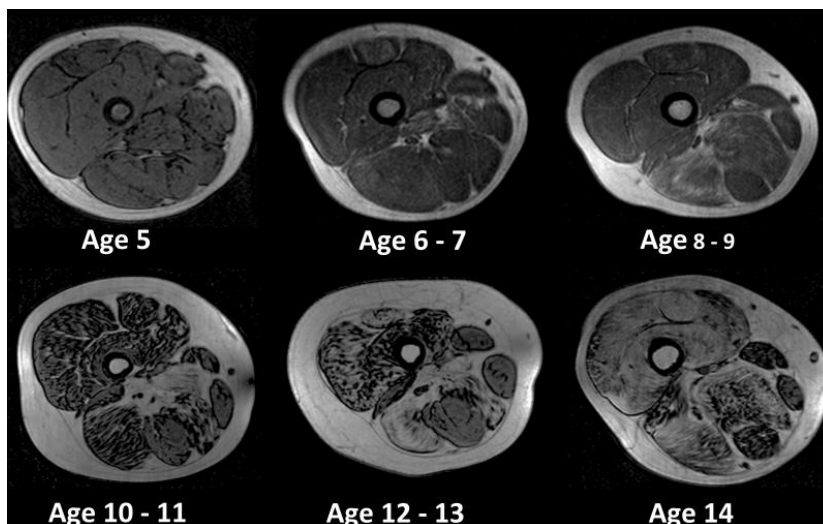


Abbildung 65: Magnetresonanztomographie des Oberschenkels von Patienten mit DMD im Alter von 5-14 Jahren (62)

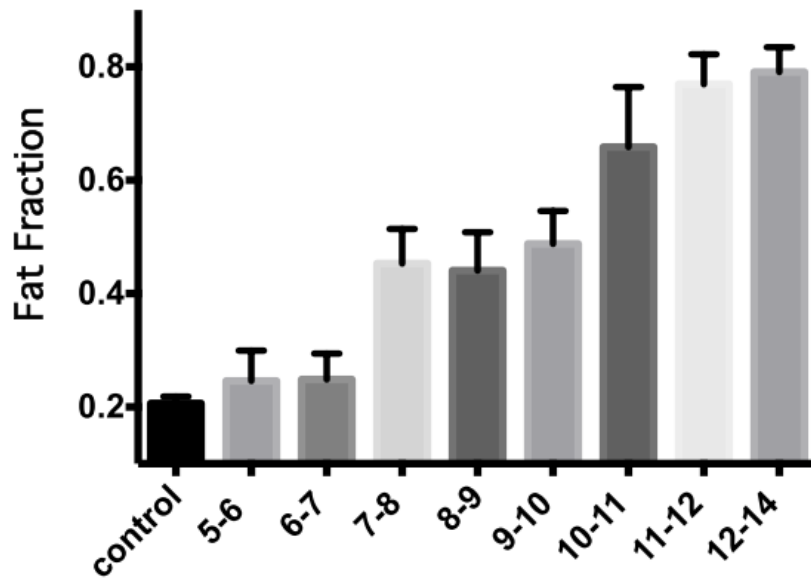


Abbildung 66: Fettgehalt des *M. vastus lateralis* von Patienten mit DMD im Alter von 5-14 Jahren (62)

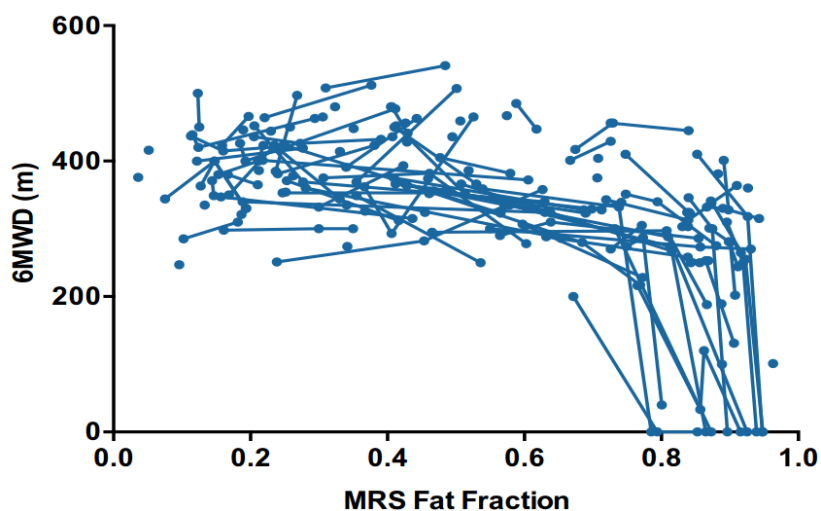


Abbildung 67: Beziehung zwischen Fettfraktion im *M. vastus lateralis* und dem Verlust der Gehfähigkeit (62)

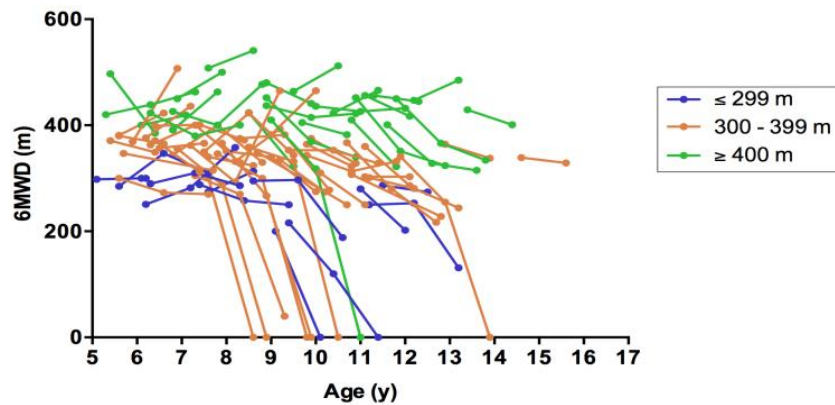


Abbildung 68: Beziehung zwischen Alter und Verlust der Gehstrecke im 6MWT (62)

Da Dystrophin nur eine intakte Muskulatur vor Schäden schützen kann und deren Funktion stabilisiert, ist zu erwarten, dass eine früh eingeleitete Dystrophin-regenerierende Therapie die intakte Muskulatur und Muskelfunktion erhält, sowie die Krankheitsprogression stabilisiert oder zumindest verzögert (53). Die Ergebnisse von Sweeney et al. (62) prognostizieren aber auch, dass der Therapieerfolg einer im frühen Krankheitsstadium eingeleiteten Dystrophin-regenerierenden Intervention im 6MWT nur nach längerer Beobachtungsdauer erkennbar wird, da ein Einfluss auf die Gehfähigkeit erst nach weiter fortgeschrittenen Muskelgewebsschäden erkennbar wird.

Aufgrund zunehmender Erkenntnis über den Krankheitsverlaufs, wird auch in anderen klinischen Studien zu DMD der Versuch unternommen, homogenere Patientenkollektive (u.a. engere 6MWD-Kriterien) einzuschliessen (8). Dies zeigt die nachfolgende Tabelle 4-56.

Tabelle 4-56: Übersicht der Einschlusskriterien von Phase 2/3-Studien bei DMD-Patienten

Studie	Zeitpunkt	BL-6MWD	Alter	Stabile Kortikosteroid-Vorbehandlung
Idebenone, Placebo, RCT, Phase 2 (64)	2005-2007	Keine	8-16 Jahre	Stabile Dosis, wenn vorhanden
L-Glutamin, RCT, Placebo, Phase 2 (65)	2006-2008	Fähig, >170 m zu gehen	Keine	Nicht definiert
Ataluren-Studie 007 Placebo, RCT, Phase 2b (14)	2008-2009	≥75 m	≥5 Jahre	Nicht definiert
GSK2402968, Placebo, RCT, Phase 3 (66)	2010-2013	≥75 m	≥5 Jahre	Ja
GSK2402968, Placebo, RCT, Phase 2 (67)	2011-2013	≥75 m (aus dem Liegen aufstehen: ≤15 sec ohne Hilfen)	≥5 Jahre	Ja
Ataluren-Studie 020 Placebo, RCT, Phase 3 (14)	2013-2015	≥150 m und ≤80% des vorhergesagten Wertes	≥7 und ≤16 Jahre	Ja
Eteplirsin, Placebo, RCT, Phase 2 (68)	2011-2012	200-400 m ±10%	7-13 Jahre	Ja
Metformin, L-Citrullin, Placebo, RCT, Phase 3 (69)	2013, laufend	≥150 m	7-10 Jahre	Ja
Eteplirsin, unbehandelte Kontrolle, Phase 3 (70)	2014, laufend	>300 m, primäre Analyse: 300-450 m	7-16 Jahre	Ja
PF-06252616, Placebo, RCT, Phase 2 (71)	2014, laufend	4 Stufen hinauf: ≥0.33 bis ≤1.6 Stufen/sec	6 bis <10 Jahre	Ja
SRP-4045, SRP-4053, Placebo, RCT, Phase 3 (72)	2016, laufend	≥300 bis ≤450 m	7-16 Jahre	Ja

Aus den Einschlusskriterien der angeführten Studien ist erkennbar, dass sich in klinischen Studien ein Band für die Baseline-Gehstrecke mit einer unteren Grenze von 300 m und einer oberen Grenze von 400 m entwickelt. Nur in diesem Band sind Therapieeffekte von Dystrophin-restaurierenden Therapien effizient messbar. Die Daten des Placeboarms der Ataluren-Studie 020 bestätigen dies eindrucksvoll. Bei Patienten, die zu Studienbeginn 400 m oder mehr im 6MWT zurücklegen konnten, änderten sich über 48 Wochen weder die Gehstrecke noch die Ergebnisse der TFTs nennenswert, während es bei Patienten mit einer Gehstrecke unter 300 m über 48 Wochen konsistent über verschiedene Testsysteme zu deutlich nachweisbaren Verschlechterungen kam. Der Großteil der Patienten, die ihre Gehfähigkeit verloren (14 Patienten der Placebo-Gruppe gesamt), hatte zu Studienbeginn eine BL-Gehstrecke von unter 300 m. Die 300-400 m Subpopulation nimmt eine Sonderstellung ein (8,14), bei der ein Behandlungseffekt gut messbar ist, da sich die Gehfähigkeit über ein Jahr kontinuierlich verschlechtert. Andererseits werden die Ergebnisse nicht durch einen raschen Verlust der körperlichen Funktionen überlagert (späte gehfähige Krankheitsphase), welche durch eine Therapie, die die Muskelfunktion stabilisiert, wahrscheinlich nur bedingt

oder bei bereits weit fortgeschrittener Schädigung der Skelettmuskulatur nicht mehr aufzuhalten ist. Bemerkenswert ist, dass sich die Unterschiede zwischen den 6MWD-Subpopulationen auch in anderen Endpunkten, die unterschiedliche Muskelfunktionen und Fähigkeiten erfassen (z.B. 6MWT primär Ausdauer (59); TFTs primär Spitzenaktivität), sowie dem PODCI (PRO) übereinstimmend nachweisen lassen. Daher bestätigen und reproduzieren die Ergebnisse der Ataluren-Studien die Ergebnisse anderer Arbeitsgruppen (Tabelle 4-56).

Auf Basis der oben dargestellten Fakten konnte bei einem Teil der in die Ataluren-Studien eingeschlossenen Patienten über 48 Wochen mit den eingesetzten Instrumenten kein therapeutischer Effekt nachgewiesen werden, da es sich entweder um stabile Patienten mit relativ guter Ausgangsfunktion handelte (6MWD  $\geq 400$  m) oder um Patienten in einem fortgeschrittenen Stadium der Erkrankung (6MWD  $< 300$  m), deren Zustand sich über 48 Wochen erheblich und unabhängig von der Behandlung verschlechterte. Bei der ersten Gruppe wäre ein Behandlungseffekt erst nach Langzeittherapie und bei der zweiten Gruppe bei sehr großen Fallzahlen messbar, vorausgesetzt es wäre noch ausreichend intaktes Gewebe der Skelettmuskulatur vorhanden. Letzteres scheint bei Patienten mit einer 6MWD unter 300 m häufig nicht der Fall zu sein, bei denen oft  $\geq 80\%$  der gesunden Muskelmasse durch Fettgewebe ersetzt sind (62). Sehr lange Placebo-kontrollierte Studien sind bei einer sehr seltenen und besonders schweren Erkrankung ethisch nicht vertretbar bzw. nicht realisierbar, insbesondere, wenn ein Arzneimittel bereits eine Zulassung hat. Wie oben dargelegt, stellen die 6MWD-Subpopulationen aufgrund der über verschiedene Tests zur Muskelfunktion konsistenten Unterschiede klar voneinander abgrenzbare Patienten-Subpopulationen bzw. Krankheitsstadien innerhalb des Krankheitskontinuums dar. Aus diesem Grund ist es gerechtfertigt, den Therapieeffekt von Ataluren in der Patientenpopulation mit BL-Werten von  $\geq 300$  bis  $< 400$  m im 6MWT zu untersuchen, bei denen ein Therapieeffekt innerhalb einer ethisch vertretbaren Studiendauer von einem Jahr nachweisbar ist (34). Dieses Vorgehen, d.h. Einschluss eines relativ breiten Patientenkollektivs und prä-spezifizierte Subgruppen-Analysen, steht grundsätzlich im Einklang mit den Empfehlungen der EMA und FDA (2,29). Wie bereits oben dargelegt wird diese Schlussfolgerung auch durch das IQWiG gestützt. Ist eine Heterogenität bezüglich des Effekts in den verschiedenen Patientengruppen vorhanden, ist der über alle Patienten hinweg geschätzte Gesamteffekt nicht sinnvoll interpretierbar (7). Aufgrund der präsentierten Daten und aufgrund des Wirkmechanismus von Ataluren erscheint es plausibel, dass der Effekt einer Muskel-stabilisierenden Therapie mit Ataluren nicht auf die Subgruppe mit einer 6MWD von  $\geq 300$  bis  $< 400$  m beschränkt sein kann. In dieser Gruppe sind jedoch die Testsysteme besonders sensitiv, sodass im Rahmen einer klinischen Studie ein Effekt innerhalb eines Jahres messbar ist. Voraussetzung ist, dass der Abbau der Skelettmuskulatur nicht bereits so weit fortgeschritten ist, dass die Patienten von einer Therapie nicht mehr profitieren können, und nach heutigem Kenntnisstand ein hohes Risiko haben, kurzfristig Rollstuhl-pflichtig zu werden.

#### 4.4.2.5 Meta-Analysen der RCTs PTC124-GD-007-DMD und PTC124-GD-020-DMD – RCT

##### 4.4.2.5.1 Vergleichbarkeit der Studien 070 und 020

Entsprechend IQWiG „Allgemeine Methoden – Version 4.2“ ist es vor einer Meta-Analyse erforderlich, zu prüfen, ob die Studien hinsichtlich der Fragestellung vergleichbar sind (7). Das primäre Ziel beider Studien war der Effekt von Ataluren auf die Gehfähigkeit und in beiden Studien war die Änderung der 6MWD von BL bis Woche 48 der primäre Endpunkt (8,35). Bei beiden Studien handelte es sich um Placebo-kontrollierte, randomisierte Doppelblindstudien der Evidenzstufe 1b über 48 Wochen. In beiden Studien wurde nach Alter (<9 vs. ≥9 Jahre) und BL-6MWD (<350 vs. ≥350) stratifiziert. In der Studie 020 wurde zusätzlich nach der Dauer der Kortikosteroid-Vorbehandlung stratifiziert, während in der Studie 007 danach stratifiziert wurde, ob die Patienten zu BL Kortikosteroide einnahmen oder nicht. Die Ataluren-Dosierung (10, 10, 20 mg/kg KG pro Tag) war ebenfalls in beiden Studien identisch (ein zusätzlicher Ataluren-Arm in der Studie 007, in dem eine höher Dosierung eingesetzt wurde, wird hier nicht weiter berücksichtigt) (8,35).

In beiden Studien war eine genetisch gesicherte Diagnose einer nmDMD Voraussetzung für den Einschluss. Andere Haupteinschlusskriterien unterschieden sich zwischen den Studien etwas:

Studie 007: ≥5 Jahre, Gehstrecke im 6MWT ≥75 m

Studie 020: ≥7-16 Jahre, Gehstrecke im 6MWT ≥150 m, ≤80% der normativen 6MWD, stabile Kortikosteroid-Vorbehandlung (8,35).

Tabelle 4-57 zeigt die in die Meta-Analyse eingeschlossen patientenrelevanten Endpunkte. Die Endpunktoperationalisierungen des primären sowie der relevanten sekundären Endpunkte waren zwischen den Studien vergleichbar. In der Studie 020 wurde die Zeit, um aus dem Liegen, aufzustehen, im Rahmen des NSAA erfasst.

Tabelle 4-57: Vergleich der in Studie 007 und 020 erhobenen patientenrelevanten Endpunkte

Endpunkt	Studie 007	Studie 020
<i>Primärer Wirksamkeitsendpunkt</i>		
Änderung der 6MWD von BL bis Woche 48	√	√
<i>Sekundäre Wirksamkeitsendpunkte</i>		
Zeit bis zu einer mindestens 10%igen Verschlechterung der 6MWD	√	√
Zeit bis zu einer mindestens 10%igen Verbesserung der 6MWD	√	-



Endpunkt	Studie 007	Studie 020
Anteil Patienten mit einer mindestens 10%igen Verschlechterung der 6MWD	√	(√)
Anteil Patienten mit einer mindestens 10%igen Verbesserung der 6MWD	√	-
TFTs, Zeit, um		
10 m zu gehen/laufen	√	√
4 Stufen hinaufzugehen	√	√
4 Stufen hinabzugehen	√	√
aus Rückenlage aufzustehen	√	√ <sup>1</sup>
Änderung der Alltagsaktivitäten gemessen mittels ‚Step Activity Monitoring‘	√	-
Änderung in der Häufigkeit von Stürzen pro Tag (Patiententagebuch)	√	-
Änderung der Häufigkeit einer Rollstuhlnutzung (Patiententagebuch)	√	-
PedsQL physikalische, emotionale, soziale, schulische Subskalen, Fatigue-Skala	√	-
<i>Explorative Wirksamkeitsendpunkte</i>		
Änderung des NSAA von BL bis Woche 48	-	√
Aktivitäten des täglichen Lebens und des Krankheitsstatus mittels Survey (PRO)	-	√
PODCI-Domänen Transfer/Basismobilität und Sport/körperliche Funktion	-	√
<i>Nebenwirkungen</i>	√	√

<sup>1</sup> Die Zeit, die benötigt wurde, um aus der Rücklage aufzustehen, wurde als Teil des NSAA erfasst.

Im SAP der Studie 007 war eine MMRM-Analyse der Kovarianz (RANCOVA) des primären Endpunkts spezifiziert. In das Modell wurden die Behandlung, die BL-6MWD, das Alter (<9 vs. ≥9 Jahre), die Kortikosteroid-Behandlung (ja/nein), die Visite und die Interaktion zwischen Behandlung und Visite eingeschlossen. Ein ANCOVA mit LOCF-Imputation sowie die Primäranalyse mit multipler Imputation waren im SAP als Sensitivitätsanalysen spezifiziert. In einer korrigierten ITT-Analyse (kITT) wurden *post-hoc* bei zwei Patienten die

BL-6MWD-Werte durch die Screening-Werte ersetzt, da diese zu BL verletzungsbedingt ein deutlich schlechteres Gehvermögen besaßen als zur Screening-Visite. Es handelte sich um einen Patienten der Placebo-Gruppe und einen Patienten der hier nicht relevanten Hochdosis-Ataluren-Gruppe. In das primäre MMRM-Modell wurde nachträglich (post-hoc) ein BL-Visite-Interaktionsterm aufgenommen, da die Beziehung zwischen BL-6MWD und post-BL-6MWD mit der Zeit variieren kann. Die untransformierten Daten wurden analysiert und ein Permutationstest wurde durchgeführt (9). Die Gründe für die Änderung des statistischen Modells der Studie 007 wurden in dem vorherigen Nutzendossier detaillierter dargelegt (38). Für die Primäranalyse der Studie 020 wurde ein ANCOVA mit multipler Imputation eingesetzt. Das ANCOVA-Modell schloss die Stratifizierungsfaktoren (Alter  $<9$  vs.  $\geq 9$  Jahre, Dauer der Kortikosteroid-Vorbehandlung  $\geq 6$  bis  $<12$  vs.  $\geq 12$  Monate, BL-6MWD  $\geq 350$  vs.  $<350$  m) und BL-6MWD ein (siehe Kapitel 4.3.1.2) (9).

Folgende Patienten-Populationen wurden in den in diesem Dossier präsentierten Meta-Analysen ausgewertet:

- kITT-Population der Studie 007 (BL-6MWD-Wert eines Patienten der Placebo-Gruppe durch den Screening-Wert ersetzt)
- ITT-Population der Studie 007
- ITT-Population der Studie 020
- kITT/ITT (Studie 007/020) BL-6MWD-Subpopulationen:  $<300$ ,  $\geq 300$  bis  $<400$  und  $\geq 400$  m;  $<350$  und  $\geq 350$  m

Da kein Patient aus der kITT-Analyse der Studie 007 ausgeschlossen wurde, entspricht diese hinsichtlich der Patienten-Charakteristika der ITT-Population. Die demographischen und anderen Charakteristika der Analyse-Populationen (ITT Studie 007, ADP-ITT Studie 007, ITT Studie 020) sowie der ITT-Subpopulation mit einer Gehstrecke im 6MWT zwischen  $\geq 300$  und  $<400$  m der Studien 007 und 020 sind in Tabelle 4-58 bzw. Tabelle 4-59 dargestellt (für den 6MWT und die TFTs sind die Werte der kITT Population der Studie 007 dargestellt). Die Daten der ADP-Subpopulation mit einer Gehstrecke im 6MWT zwischen  $\geq 300$  und  $<400$  m zu Studienbeginn (Studie 007) sind in Tabelle 4-59 der Vollständigkeit halber dargestellt. Da nur wenige Patienten in diese Analyse eingeschlossen wurden, wird nur die größere kITT-Population der Studie 007 bei der Analyse der 6MWD-Subpopulationen berücksichtigt.

Hinsichtlich des mittleren Alters ergaben sich keine relevanten Unterschiede zwischen den Populationen und Gruppen (Studie 007: ITT und ADP-ITT 007, Studie 020: ITT). Allerdings war der Anteil der Patienten, die zu Studienbeginn 9 Jahre oder älter waren, in der Studie 020 (53.5%/50.0% Placebo/Ataluren) und insbesondere in der ADP-Population der Studie 007 (61.3%/59.4%) höher als in der ITT-Population der Studie 007 (44%/44%). Hierbei ist zu beachten, dass in die Studie 007 Patienten ab einem Alter von fünf Jahren eingeschlossen wurden, während in die Studie 020 und die ADP-ITT-Population der Studie 007, Patienten ab

einem Alter von 7 Jahren berücksichtigt wurden. Die mittlere BL-6MWD war in der Studie 020 (ITT 364/365 m Placebo/Ataluren) geringfügig höher als in der ITT-Population der Studie 007 (361/350 m), während diese in der ADP-Population der Studie 007 (339/341 m) niedriger war. Trotz der Intention, die Vergleichbarkeit der Population zu erhöhen, war die ITT-Population der Studie 007 der ITT-Population der Studie 020 insgesamt ähnlicher als die ADP-ITT-Population (Studie 007), die den Haupteinschlusskriterien der Studie 020 entsprach. Die Streubreite der 6MWD war in allen Populationen hoch und zwischen den ITT-Populationen vergleichbar. Der Anteil der Patienten mit einer 6MWD von 350 m oder höher war in der Studie 007 (ITT) deutlich niedriger als in der Studie 020 (40%/44% vs. 64%/64%; 54.8%/56.3% ADP-ITT Studie 070). Hinsichtlich der 6MWD-Subpopulationen (<300, ≥300 bis <400 m, ≥400 m) ergaben sich zwischen den Studiengruppen nur relativ moderate Unterschiede. Der Anteil der Patienten mit einer Gehstrecke zwischen ≥300 und <400 m war in beiden Studien in vergleichbarer Größenordnung (38.6%/38.6% kITT Studie 007, 38.7%/43.8% ADP-ITT Studie 007, 45.6%/41.2% ITT Studie 020). Die Definition einer 6MWD-Obergrenze (≤80% des vorausgesagten Wertes) hat daher nicht dazu geführt, dass weniger Patienten mit noch relativ guter Gehfähigkeit in die Studie 020 eingeschlossen wurden. Hinsichtlich der TFTs ergaben sich nur geringe Unterschiede fraglicher klinischer Relevanz. Insgesamt zeigt der Vergleich der Patientenpopulation der Studie 007, in die ein breiteres Patientenkollektiv eingeschlossen wurde, und der Studie 020, dass das Ziel einer prognostischen Anreicherung in der Studie 020 insgesamt nicht erreicht wurde. Hinsichtlich der BL 6MWD erwies sich die ITT-Population der Studie 020 als ebenso heterogen wie die ITT-Population der Studie 007.

Ein Unterschied zwischen beiden Studien bestand darin, dass die Patienten der Studie 007 im Gegensatz zur Studie 020 nicht mit Kortikosteroiden vorbehandelt sein mussten. Allerdings nahmen etwa 70% der Teilnehmer vor und während der Studie 007 (ITT) Kortikosteroide ein. *Per definitionem* nahmen alle Patienten der Studie 020 Kortikosteroide ein. In der Studie 007 nahmen die Mehrzahl der Patienten täglich Kortikosteroide ein. Die Dosen schwankten bei diesen Patienten zwischen 7.5 und 33 mg für Deflazacort, 10 und 30 mg für Prednisolon und 10 und 25 mg für Prednison. Die Empfehlungen zur Standardisierung der Kortikosteroid-Therapie waren in beiden Studien insgesamt ähnlich. Änderungen des Typs oder des Regimes sollten in beiden Studien wenn möglich vermieden werden.

Tabelle 4-58: Demographische und andere Charakteristika der Patienten – Vergleich der RCTs 007 und 020

Charakteristik	Studie 007 (ITT)		Studie 007 (ADP-ITT)		Studie 020 (ITT)	
	Placebo (n=57)	Ataluren (n=57)	Placebo (n=31)	Ataluren (n=32)	Placebo (n=114)	Ataluren (n=114)
Alter (Jahre) M ( $\pm$ SD) Minimum, Maximum	8.3 (2.33) 5, 15	8.8 (2.91) 5, 20	9.3 (2.07) 7, 15	9.1 (2.01) 7, 15	9.0 (1.66) 7, 14	8.9 (1.77) 7, 14
Altersgruppe, n (%) <9 Jahre $\geq$ 9 Jahre	32 (56) 25 (44)	32 (56) 25 (44)	12 (38.7) 19 (61.3)	13 (40.6) 19 (59.4)	53 (46.5) 61 (53.5)	57 (50.0) 57 (50.0)
Kaukasier oder weiß, n (%)	54 (94.7)	53 (93.0)	29 (93.5)	31 (96.9)	85 (74.6)	88 (77.2)
Größe (cm) M ( $\pm$ SD)	123.4 (11.77)	124.5 (15.30)	126.17 (9.413)	123.72 (9.884)	126.26 (10.434)	125.99 (10.648)
Gewicht (kg) M ( $\pm$ SD)	28.6 (9.12)	31.2 (12.11)	32.52 (9.278)	31.68 (10.474)	30.68 (10.464)	31.35 (10.779)
BMI (kg/m <sup>2</sup> ) M ( $\pm$ SD)	18 (3.7)	19 (3.5)	20.12 (3.858)	20.21 (4.075)	18.85 (4.083)	19.31 (4.374)
BL-6MWD (m) M ( $\pm$ SD) Minimum, Maximum	361.12 (87.524) 158.5, 533.3	350.00 (97.553) 75.0, 525.0	339.13 (84.619) 158.5, 459.0	341.01 (84.788) 159.5, 488.0	363.54 (81.266) 142.5, 526.0	365.19 (72.610) 166.8, 511.0
BL-6MWD, n (%) <350 m $\geq$ 350 m	34 (60) 23 (40)	32 (56) 25 (44)	14 (45.2) 17 (54.8)	14 (43.8) 18 (56.3)	41 (36.0) 73 (64.0)	41 (36.0) 73 (64.0)
BL-6MWD, n (%) <300 m $\geq$ 300 bis <400 m $\geq$ 400 m	13 (22.8) 22 (38.6) 22 (38.6)	15 (26.3) 22 (38.6) 20 (35.1)	9 (29.0) 12 (38.7) 10 (32.3)	8 (25.0) 14 (43.8) 10 (31.3)	21 (18.4) 52 (45.6) 41 (36.0)	24 (21.1) 47 (41.2) 43 (37.7)

Charakteristik	Studie 007 (ITT)		Studie 007 (ADP-ITT)		Studie 020 (ITT)	
	Placebo (n=57)	Ataluren (n=57)	Placebo (n=31)	Ataluren (n=32)	Placebo (n=114)	Ataluren (n=114)
4 Stufen hinauf (sec) M (±SD) Minimum, Maximum	6.01 (5.671) 1.4, 30.0	6.94 (6.474) 1.4, 30.0	6.95 (6.472) 1.5, 30.0	6.28 (4.884) 1.6, 22.5	(n=111) 6.69 (7.279) 1.1, 50.0	(n=111) 5.97 (5.367) 1.3, 35.9
4 Stufen hinab (sec) M (±SD) Minimum, Maximum	5.49 (5.755) 0.0, 30.0	6.08 (5.985) 1.5, 30.0	6.26 (6.702) 0.0, 30.0	6.16 (6.004) 1.5, 30.0	(n=108) 4.85 (4.972) 1.1, 36.9	(n=111) 5.06 (5.179) 1.4, 34.0
10 m rennen/gehen (sec) M (±SD) Minimum, Maximum	6.72 (2.675) 3.6, 15.3	7.45 (4.373) 3.6, 30.0	7.07 (2.627) 3.6, 11.3	7.77 (4.685) 3.6, 30.0	6.80 (2.919) 3.2, 17.0	6.63 (3.073) 3.0, 25.0
Kortikosteroid-Vorbe- handlung, n (%) 6 bis <12 Monate ≥12 Monate	Insgesamt: 40 (70.2%)	Insgesamt: 41 (71.9%)	5 (16.1) 26 (83.9)	7 (21.9) 25 (78.1)	18 (15.8) 96 (84.2)	19 (16.7) 95 (83.3)
Quellen:	Studienreport 007: Tabelle 10, Tabelle 11, Tabelle 14.1.2.1, Tabelle 14.1.3.2, Tabelle 14.2.1.3.1s, Tabelle 14.2.3.1.15, Tabelle 14.2.3.3.17 (9) ISE-Tabellen 2.1.2-4 bis 2.1.2-6 (15)		ISE Tabelle 2.1.1-10 (15)		Studienreport 020: Tabelle 10, 14.1.3.1.1 (8)	

Tabelle 4-59: Demographische und andere Charakteristika der Patienten (Subgruppe BL-6MWD  $\geq 300$  bis  $< 400$  m) – Vergleich der RCTs 007 und 020

Charakteristik	Studie 007 BL-6MWD $\geq 300$ bis $< 400$ m (ITT)		Studie 007 BL-6MWD $\geq 300$ bis $< 400$ m (ADP-ITT)		Studie 020 BL-6MWD $\geq 300$ bis $< 400$ m (ITT)	
	Placebo (n=22)	Ataluren (n=22)	Placebo (n=12)	Ataluren (n=14)	Placebo (n=52)	Ataluren (n=47)
Alter (Jahre) M ( $\pm$ SD) Minimum, Maximum	7.5 (2.22) 5, 15	8.4 (3.43) 5, 20	8.9 (2.11) 7, 15	8.9 (2.06) 7, 15	9.0 (1.49) 7, 12	8.9 (1.81) 7, 14
Altersgruppe, n (%) <9 Jahre $\geq 9$ Jahre	16 (72.7) 6 (27.3)	12 (54.5) 10 (45.5)	6 (50.0) 6 (50.0)	5 (35.7) 9 (64.3)	22 (42.3) 30 (57.7)	23 (48.9) 24 (51.1)
Kaukasier oder weiß, n (%)	19 (86.4)	20 (90.9)	10 (83.3)	14 (100.0)	38 (73.1)	37 (78.7)
Größe (cm) M ( $\pm$ SD)	119.16 (9.217)	121.83 (13.606)	124.88 (7.754)	122.70 (8.275)	126.73 (9.502)	123.76 (10.750)
Gewicht (kg) M ( $\pm$ SD)	25.61 (8.233)	28.89 (12.310)	30.02 (8.739)	28.34 (7.225)	30.71 (10.191)	30.43 (11.174)
BMI (kg/m <sup>2</sup> ) M ( $\pm$ SD)	17.71 (3.536)	18.81 (2.990)	19.03 (4.069)	18.63 (3.054)	18.78 (4.122)	19.41 (4.844)
BL-6MWD (m) M ( $\pm$ SD) Minimum, Maximum	342.34 (26.094) 300.0, 389.0	351.01 (29.160) 301.6, 397.0	341.13 (30.756) 300.0, 384.2	348.64 (30.018) 301.6, 394.1	354.52 (29.712) 300.7, 399.5	356.69 (26.672) 305.0, 398.0
BL-6MWD, n (%) <350 m $\geq 350$ m	10 (45.5) 12 (54.5)	10 (45.5) 12 (54.5)	5 (41.7) 7 (58.3)	6 (42.9) 8 (57.1)	Nicht verfügbar	Nicht verfügbar
4 Stufen hinauf (sec) M ( $\pm$ SD) Minimum, Maximum	5.13 (3.658) 1.5, 17.6	5.27 (2.855) 2.6, 14.9	5.77 (4.534) 1.5, 17.6	4.77 (1.479) 2.6, 7.6	6.10 (3.072) 1.7, 18.5	5.16 (1.959) 1.6, 11.0

Charakteristik	Studie 007 BL-6MWD ≥300 bis <400 m (ITT)		Studie 007 BL-6MWD ≥300 bis <400 m (ADP-ITT)		Studie 020 BL-6MWD ≥300 bis <400 m (ITT)	
	Placebo (n=22)	Ataluren (n=22)	Placebo (n=12)	Ataluren (n=14)	Placebo (n=52)	Ataluren (n=47)
4 Stufen hinab (sec) M (±SD) Minimum, Maximum	4.67 (3.774) 2.1, 19.5	4.62 (2.496) 2.2, 13.4	5.34 (4.970) 2.1, 19.5	4.34 (1.804) 2.2, 8.7	(n=51) 4.63 (2.753) 1.1, 19.7	4.03 (1.645) 1.7, 10.0
10 m rennen/ gehen (sec) M (±SD) Minimum, Maximum	6.54 (2.147) 3.8, 11.3	7.32 (5.276) 4.2, 30.0	6.79 (2.281) 3.8, 11.3	8.21 (6.418) 4.3, 30.0	6.74 (1.717) 3.7, 12.0	6.27 (1.639) 3.1, 10.9
Kortikosteroid-Vorbe- handlung, n (%) 6 bis <12 Monate ≥12 Monate	2 (9.1) 13 (59.1)	4 (18.2) 14 (63.6)	2 (16.7) 10 (83.3)	3 (21.4) 11 (78.6)	7 (13.5) 45 (86.5)	7 (14.9) 40 (85.1)
Quellen	ISE Tabelle 2.1.1-5		ISE Tabelle 2.1.1-13		Studienbericht 020: Tabelle 14.1.3.1.3 (8)	

*Vergleich der Ergebnisse der Studien 007 und 020*

Wie aus Tabelle 4-60 ersichtlich war der Behandlungseffekt von Ataluren hinsichtlich der Änderung der 6MWD (primärer Endpunkt) bis Woche 48 mit circa 30 m verglichen mit Placebo in der Studie 007 deutlicher als in der Studie 020 (12.98 m), der Effekt war in seiner Ausrichtung in beiden Studien gleichgerichtet. Der Unterschied war in beiden Primäranalysen nicht signifikant, während in der prä-spezifizierten ANCOVA Sensitivitätsanalyse mit LOCF-Imputation und der *post-hoc* kITT-Analyse (augmentiertes MMRM Modell) der Studie 007 das Signifikanzniveau erreicht wurde.

Tabelle 4-60: Ergebnisse des primären Endpunkts in Studie 007 und 020 (8,9)

Statistisches Modell	Studie 007		Studie 020	
	$\Delta$ LSM $\pm$ SE vs. Placebo (95% KI) m	p-Wert	$\Delta$ LSM vs. Placebo (95% KI) m	p-Wert
MMRM der Kovarianz, RANCOVA (ITT) <sup>1</sup> Prä-spezifizierte Primär-Analyse, nicht transformierte Daten	26.44 $\pm$ 15.53 (-4.21, 57.09)	0.0905 (nominal) 0.1592 (Dunnett) (rank-transformiert: 0.1490, 0.2539)	-	
MMRM der Kovarianz, RANCOVA (kITT) <sup>3</sup> Post-hoc-Analyse	31.7 (5.1, 58.3)	0.0197 (nominal) 0.0367 (Dunnett)	-	
ANCOVA (ITT, multiple Imputation) <sup>4</sup> Prä-spezifizierte Primär-Analyse	-	-	12.98 $\pm$ 10.415 (-7.44, 33.39)	0.213
MMRM (ITT) <sup>5</sup> Prä-spezifizierte Sensitivitäts-analyse			13.49 $\pm$ 10.374 (-6.96, 33.93)	0.195
ANCOVA mit LOCF-Imputation (ITT) <sup>2</sup> Prä-spezifizierte Sensitivitäts-analyse, nicht transformierte Daten	28.3570 $\pm$ 14.0056 (0.71, 56.01)	0.0445 (nominal) 0.0808 (Dunnett)	-	
Analyse mit multipler Imputation (ITT) <sup>1</sup> Prä-spezifizierte Sensitivitäts-analyse	25.66 $\pm$ 15.558 (-4.836, 56.153)	0.0991	-	

<sup>1</sup> Faktoren: BL-6MWD (Kovariate), Behandlungsarm, Visite, Visite-Behandlungsarm, Altersgruppe (<9 vs.  $\geq$ 9 Jahre), Kortikosteroid-Anwendung zu BL (ja/nein); unstrukturierte Varianz-/Kovarianz-Matrix (Tabelle 23, Tabelle 14.2.1.8.7B Studienbericht 007)

<sup>2</sup> Faktoren: Behandlung, Alter und Kortikosteroid Stratifikationsfaktoren, BL-6MWD (Kovariate) (Tabelle 14.2.2.15 Studienbericht 007)



<sup>3</sup> Faktoren: BL-6MWD (Kovariate), Behandlungsarm, Visite, Visite-Behandlungsarm, BL-6MWD-Visite, Altersgruppe (<9 vs. ≥9 Jahre), Kortikosteroid-Anwendung zu BL (ja/nein); unstrukturierte Varianz-/Kovarianz-Matrix (Tabelle 27 Studienbericht 007)

<sup>4</sup> Faktoren: Stratifikationsfaktoren (Alter <9 vs. ≥9 Jahre BL-6MWD-Gruppe <350 vs. ≥350 m, Dauer der Kortikosteroid-Vorbehandlung (≥6 bis <12 vs. ≥12 Monate), Behandlung und BL-6MWD (Tabelle 14.2.1.3.1 Studienbericht 020)

<sup>5</sup> Faktoren: Stratifikationsfaktoren (Alter <9 vs. ≥9 Jahre BL-6MWD-Gruppe <350 vs. ≥350 m, Dauer der Kortikosteroid-Vorbehandlung (≥6 bis <12 vs. ≥12 Monate), Behandlung, BL-6MWD, Visite, Interaktion zwischen Behandlung und Visite, Visite und BL-6MWD und Visite und jeder der Stratifikationsfaktoren (Tabelle 14.2.1.4 Studienbericht 020)

Tabelle 4-61, Tabelle 4-62 und Tabelle 4-63 geben einen Überblick der Ergebnisse des primären und relevanter sekundärer Endpunkte der Analyse-Populationen der Studie 007 und 020. Hinsichtlich der patientenrelevanten Endpunkte 6MWD und TFTs ergaben sich in der Gesamtpopulation übereinstimmend positive Trends, die teilweise das Signifikanzniveau erreichten. Gründe für die Unterschiede zwischen den Studien werden unten diskutiert. Für den primären Endpunkt, sowie die sekundären Endpunkte waren die Behandlungseffekte in der Subpopulation mit einer 6MWD von ≥300 bis <400 m in beiden Studien übereinstimmend deutlicher ausgeprägt als in der Gesamtpopulation. Bei Patienten mit einer Gehstrecke unter 350 m zeigten sich bei einigen Endpunkten signifikante Effekte zugunsten von Ataluren. In keinen der beiden Studien zeigten die Lebensqualitätsinstrumente einen signifikanten Einfluss Unterschied zwischen den Behandlungsarmen. Hinsichtlich der Alltagsaktivitäten ermittelt mittels Pedometer in Studie 007 und eines Surveys (PRO) in Studie 020 ergaben sich zwischen den Gruppen keine signifikanten Unterschiede (Daten nicht gezeigt; siehe Tabelle 4-32 für Studie 020 und früheres Dossier zur Nutzenbewertung für Studie 007).

Tabelle 4-61: Ergebnisse des primären und relevanter sekundärer Endpunkte in der Gesamtpopulation aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo (8,9)

Endpunkt: Änderung bis Woche 48 Ataluren vs. Placebo	Studie 007 (kITT) <sup>1</sup> n=57 Placebo n=57 Ataluren	Studie 007 (ITT) <sup>1</sup> n=57 Placebo n=57 Ataluren	Studie 020 (ITT) n=114 Placebo n=114 Ataluren
<i>Primärer Endpunkt und supportive Analyse</i>			
6MWD (m) $\Delta$ LSM 95% KI p-Wert	31.7 5.1, 58.3 <b>p=0.0197</b>	26.4 -4.2, 57.1 p=0.0905	12.98 -7.44, 33.39 p=0.213
Zeit bis 10%ige Verschlechterung der 6MWD	HR 0.508 0.273, 0.946 <b>p=0.0326</b>	HR 0.520 0.280, 0.966 <b>p=0.0386</b>	HR 0.75 0.51, 1.12 p=0.1603
<i>Änderung der proximalen Muskelfunktion</i>			
10 m gehen/rennen (sec) $\Delta$ LSM 95% KI p-Wert	-1.37 -3.48, 0.75 p=0.2045	Rang-transformiert -15.02 -92.55, 62.51 p=0.7025	-1.07 -2.40, 0.27 p=0.117
4 Stufen hinauf (sec) $\Delta$ LSM 95% KI p-Wert	-2.57 -4.82, -0.32 <b>p=0.0252</b>	Rang-transformiert -68.92 -150.7, 12.88 p=0.0981	-1.43 -2.90, 0.05 p=0.058
4 Stufen hinunter (sec) $\Delta$ LSM 95% KI p-Wert	-1.74 -4.20, 0.71 p=0.1626	Rang-transformiert -38.16 -133.5, 57.21 p=4307	-1.97 -3.52, -0.43 <b>p=0.012</b>
Aus dem Liegen aufstehen $\Delta$ LSM 95% KI p-Wert	-0.06 -2.35, 2.23 p=0.9580	Rang-transformiert -13.32 -107.0, 44.40 p=0.4153	-0.36 -1.96, 1.24 p=0.661
NSAA $\Delta$ LSM 95% KI p-Wert (transformiert)	Nicht erhoben	Nicht erhoben	1.51 -1.16, 4.17 p=0.268
NSAA $\Delta$ LSM 95% KI p-Wert	Nicht erhoben	Nicht erhoben	0.80 -0.23, 1.82 p=0.128

<b>Endpunkt: Änderung bis Woche 48 Ataluren vs. Placebo</b>	<b>Studie 007 (kITT)<sup>1</sup> n=57 Placebo n=57 Ataluren</b>	<b>Studie 007 (ITT)<sup>1</sup> n=57 Placebo n=57 Ataluren</b>	<b>Studie 020 (ITT) n=114 Placebo n=114 Ataluren</b>
<i>Lebensqualität</i>			
PedsQL physisch Δ LSM 95% KI p-Wert	3.43 -4.37, 11.24 p=0.3865	3.56 -4.31, 11.42 p=0.3733	Nicht erhoben
PedsQL emotional Δ LSM 95% KI p-Wert	-3.66 -10.92, 3.60 p=0.3207 p=0.5024	-4.20 -11.5, 3.05 p=0.2543 p=0.4107	Nicht erhoben
PedsQL sozial Δ LSM 95% KI p-Wert	-2.25 -9.41, 4.91 p=0.5359	-2.38 -9.52, 4.76 p=0.5117	Nicht erhoben
PedsQL schulisch Δ LSM 95% KI p-Wert	2.56 -4.41, 9.53 p=0.4695	2.54 -4.42, 9.50 p=0.4727	Nicht erhoben
PODCI Transfer/ Basis- mobilität Δ LSM 95% KI p-Wert	Nicht erhoben	Nicht erhoben	1.64 -2.11, 5.39 p=0.391
PODCI Sport/physikalisch Δ LSM 95% KI p-Wert	Nicht erhoben	Nicht erhoben	2.15 -1.75, 6.05 p=0.281

<sup>1</sup> p-Werte: 1. Wert nominal, 2. Wert Dunnets

Quelle: Siehe Legende Tabelle 4-62

Tabelle 4-62: Ergebnisse des primären und relevanter sekundärer Endpunkte in den 6MWD-Subpopulationen (8,9)

Endpunkt: Änderung bis Woche 48 Ataluren vs. Placebo	Studie 007 (6MWD <300 m kITT) n=13 Placebo n=15 Ataluren	Studie 020 (6MWD <300 m ITT) n=21 Placebo n=24 Ataluren	Studie 007 (6MWD ≥300 bis <400 m kITT) n=22 Placebo n=22 Ataluren	Studie 020 (6MWD ≥300 bis <400 m ITT) n=52 Placebo n=47 Ataluren	Studie 007 (6MWD ≥400 m kITT) n=22 Placebo n=20 Ataluren	Studie 020 (6MWD ≥400 m ITT) n=41 Placebo n=43 Ataluren
<i>Primärer Endpunkt und supportive Analyse</i>						
6MWD (m) Δ LSM 95% KI p-Wert p-Wert Interaktion	25.1	-7.71 -54.93, 39.51 p=0.749	48.7	42.89 11.75, 74.03 <b>p=0.007</b>	18.9	-9.51 -43.19, 24.18 p=0.580 <b>p=0.050</b>
Zeit bis 10%ige Verschlechterung der 6MWD	HR 1.08	HR 0.48 0.24, 0.93 <b>p=0.0306</b>	HR 0.28	HR 0.79 0.44, 1.41 p=0.4181	HR 0.31	HR 1.52 0.59, 3.91 p=0.3872
<i>Änderung der proximalen Muskelfunktion</i>						
10 m gehen/rennen (sec) Δ LSM 95% KI p-Wert p-Wert Interaktion	-1.2	-2.75 -5.69, 0.18 p=0.066	-3.2	-1.84 -3.79, 0.12 p=0.066	-0.2	0.21 -1.91, 2.32 p=0.848 p=0.208
4 Stufen hinauf (sec) Δ LSM 95% KI p-Wert p-Wert Interaktion	-4.4	-0.47 -3.94, 3.00 p=0.790	-3.4	-3.46 -5.71, -1.21 <b>p=0.003</b>	-0.5	0.17 -2.26, 2.60 p=0.893 <b>p=0.080</b>

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

Endpunkt: Änderung bis Woche 48 Ataluren vs. Placebo	Studie 007 (6MWD <300 m kITT) n=13 Placebo n=15 Ataluren	Studie 020 (6MWD <300 m ITT) n=21 Placebo n=24 Ataluren	Studie 007 (6MWD ≥300 bis <400 m kITT) n=22 Placebo n=22 Ataluren	Studie 020 (6MWD ≥300 bis <400 m ITT) n=52 Placebo n=47 Ataluren	Studie 007 (6MWD ≥400 m kITT) n=22 Placebo n=20 Ataluren	Studie 020 (6MWD ≥400 m ITT) n=41 Placebo n=43 Ataluren
4 Stufen hinunter (sec) Δ LSM 95% KI p-Wert p-Wert Interaktion	-0.7	-0.97 -4.55, 2.61 p=0.595	-4.1	-4.36 -6.64, -2.07 <b>&lt;0.001</b>	-0.3	-0.13 -2.58, 2.32 p=0.917 <b>p=0.037</b>
NSAA Δ LSM 95% KI p-Wert (transformiert) p-Wert Interaktion	Nicht erhoben	0.65 -5.59, 6.89 p=0.837	Nicht erhoben	4.26 0.17, 8.35 <b>p=0.041</b>	Nicht erhoben	-1.06 -5.46, 3.34 p=0.637 p=0.207
NSAA Δ LSM 95% KI p-Wert p-Wert Interaktion	Nicht erhoben	0.37 -2.02, 2.77 p=0.760	Nicht erhoben	1.70 0.11, 3.30 <b>p=0.037</b>	Nicht erhoben	-0.11 -1.83, 1.60 p=0.896 p=0.291
<i>Lebensqualität</i>						
PODCI Transfer/ Basis-mobilität Δ LSM 95% KI p-Wert p-Wert Interaktion	Nicht erhoben	-0.86 -9.12, 7.39 p=0.837	Nicht erhoben	3.54 -2.03, 9.11 p=0.213	Nicht erhoben	2.06 -3.90, 8.03 p=0.498 p=0.615

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

<b>Endpunkt: Änderung bis Woche 48 Ataluren vs. Placebo</b>	<b>Studie 007 (6MWD &lt;300 m kITT) n=13 Placebo n=15 Ataluren</b>	<b>Studie 020 (6MWD &lt;300 m ITT) n=21 Placebo n=24 Ataluren</b>	<b>Studie 007 (6MWD ≥300 bis &lt;400 m kITT) n=22 Placebo n=22 Ataluren</b>	<b>Studie 020 (6MWD ≥300 bis &lt;400 m ITT) n=52 Placebo n=47 Ataluren</b>	<b>Studie 007 (6MWD ≥400 m kITT) n=22 Placebo n=20 Ataluren</b>	<b>Studie 020 (6MWD ≥400 m ITT) n=41 Placebo n=43 Ataluren</b>
PODCI Sport/physikalisch Δ LSM 95% KI p-Wert p-Wert Interaktion	Nicht erhoben	2.83 -6.06, 11.71 p=0.533	Nicht erhoben	4.49 -1.49, 10.47 p=0.141	Nicht erhoben	-0.89 -7.32, 5.55 p=0.797 p=0.464

Quelle (Studie 007 Modul 4): Tabelle 4-40 (10%ige Verschlechterung 6MWD ITT, kITT), Tabelle 4-44 (TFTs kITT)  
 Quelle (Studienbericht 007): Tabelle 23 (6MWD ITT), Tabelle 27 (6MWD kITT), Tabelle 14.2.3.1.14, Tabelle 14.2.3.2.2.2S, Tabelle 14.2.3.3.2.2S, Tabelle 14.2.3.1.2.2S, Tabelle 14.2.3.4.2.2S (TFTs ITT), Tabelle 31 (Aktivität ITT), Tabelle 14.2.7.2B, Tabelle 14.2.7.10B, Tabelle 14.2.7.14B (PedsQL ITT), Tabelle 14.2.7.2F, Tabelle 14.2.7.6F, Tabelle 14.2.7.10F, Tabelle 14.2.7.14F (PedsQL kITT); ISE: Tabelle 3.2.2 (10%ige Verschlechterung Subgruppe), Tabelle 3.1.2 (6MWD, TFTs)

Quelle (Studienbericht 020): Tabelle 14.2.1.3.1 (6MWD ITT), Tabelle 14.2.1.3.3.2 (6MWD Subgruppe), Tabelle 14.2.1.5.1 (10%ige Verschlechterung ITT), Tabelle 14.2.1.5.3 (10%ige Verschlechterung Subgruppe), Tabelle 14.2.2.2.2 (10 m gehen/rennen ITT), Tabelle 14.2.2.2.8 (10 m gehen/rennen Subgruppe), Tabelle 14.2.2.1.2 (aus dem Liegen Aufstehen ITT), Tabelle 14.2.2.3.2 (4 Stufen hinauf ITT), Tabelle 14.2.2.3.8 (4 Stufen hinauf Subgruppe), Tabelle 14.2.2.4.2 (4 Stufen hinunter ITT), Tabelle 14.2.2.4.8 (4 Stufen hinunter Subgruppe), Tabelle 14.2.2.1.2 (Aus dem Liegen aufstehen), Tabelle 14.2.3.4 (NSAA transformiert ITT), Tabelle 14.2.3.6 (NSAA Subgruppe), Tabelle 14.2.3.8 (NSAA Subgruppe), Tabelle 14.2.3.2 (NSAA ITT), Tabelle 14.2.4.1.2 (PODCI Transfer/Basismobilität ITT), Tabelle 14.2.4.1.4 (PODCI Subgruppe), Tabelle 14.2.4.2.2 (PODCI ITT), Tabelle 14.2.4.2.4 (PODCI Subgruppe)

Tabelle 4-63: Ergebnisse des primären und relevanter sekundärer Endpunkte in den 6MWD-Subpopulationen (8,9)

Endpunkt: Änderung bis Woche 48 Ataluren vs. Placebo	Studie 007 (6MWD <350 m kITT) n=35 Placebo n=32 Ataluren	Studie 020 (6MWD <350 ITT) n=41 Placebo n=41 Ataluren	Studie 007 (6MWD ≥350 m kITT) n=22 Placebo n=25 Ataluren	Studie 020 (6MWD ≥350 ITT) n=73 Placebo n=73 Ataluren
<i>Primärer Endpunkt und supportive Analyse</i>				
6MWD (m) Δ LSM 95% KI p-Wert p-Wert Interaktion	59.81 18.03, 101.59 <b>p=0.0053</b>	21.89 -12.32, 56.11 p=0.210	13.39 -21.20, 47.98 p=0.4458 p=0.3824	7.98 -17.53, 33.48 p=0.540 p=0.482
Zeit bis 10%ige Verschlechterung der 6MWD	HR 0.49	HR 0.57 0.34, 0.94 <b>p=0.0293</b>	HR 0.55	HR 1.15 0.62, 2.16 p=0.6542
<i>Änderung der proximalen Muskelfunktion</i>				
10 m gehen/rennen (sec) Δ LSM 95% KI p-Wert p-Wert Interaktion	-3.4	-2.43 -4.65, -0.20 <b>p=0.033</b>	-0.4	-0.30 -1.96, 1.36 p=0.721 <i>p=0.137</i>
4 Stufen hinauf (sec) Δ LSM 95% KI p-Wert p-Wert Interaktion	-6.2	-2.98 -5.46, -0.49 <b>p=0.019</b>	-0.1	-0.57 -2.39, 1.25 p=0.541 <i>p=0.131</i>
4 Stufen hinunter (sec) Δ LSM 95% KI p-Wert p-Wert Interaktion	-4.9	-3.61 -6.24, -0.98 p=0.007	0.1	-1.07 -2.96, 0.82 p=0.268 <i>p=0.135</i>
NSAA Δ LSM 95% KI p-Wert (transformiert) p-Wert Interaktion	Nicht erhoben	3.35 -1.14, 7.85 p=0.143	Nicht erhoben	0.45 -2.87, 3.76 p=0.792 p=0.305
NSAA Δ LSM 95% KI p-Wert p-Wert Interaktion	Nicht erhoben	1.46 -0.26, 3.18 p=0.097	Nicht erhoben	0.41 -0.87, 1.70 p=0.527 p=0.336

<b>Endpunkt: Änderung bis Woche 48 Ataluren vs. Placebo</b>	<b>Studie 007 (6MWD &lt;350 m kITT) n=35 Placebo n=32 Ataluren</b>	<b>Studie 020 (6MWD &lt;350 ITT) n=41 Placebo n=41 Ataluren</b>	<b>Studie 007 (6MWD ≥350 m kITT) n=22 Placebo n=25 Ataluren</b>	<b>Studie 020 (6MWD ≥350 ITT) n=73 Placebo n=73 Ataluren</b>
<i>Lebensqualität</i>				
PODCI Transfer/ Basismobilität Δ LSM 95% KI p-Wert p-Wert Interaktion	Nicht erhoben	1.89 -4.41, 8.18 p=0.557	Nicht erhoben	1.51 -3.19, 6.20 p=0.529 p=0.062
PODCI Sport/physi- kalisches Δ LSM 95% KI p-Wert p-Wert Interaktion	Nicht erhoben	6.27 -0.25, 12.78 p=0.059	Nicht erhoben	-0.17 -5.04, 4.71 p=0.946 p=0.127

Quelle (Studienbericht 070): Tabelle 14.2.1.28E, Tabelle 14.2.1.28F

Quelle (Studienbericht 020): Tabelle 14.2.1.3.3.1 (6MWD), Tabelle 14.2.1.5.4 (Verslechterung im 6MWD), Tabelle 14.2.2.2.10 (10 m gehen/laufen), Tabelle 14.2.2.3.10 (4 Stufen hinauf), Tabelle 14.2.2.4.10 (4 Stufen hinunter), Tabelle 14.2.3.10 (NSAA nicht transformiert), Tabelle 14.2.3.12 (NSAA transformiert), Tabelle 14.2.4.1.6, Tabelle 14.2.4.2.6 (PODCI)

Quelle (ISE, Studie 007): Tabelle 3.1.2

Die meisten TEAEs waren in beiden Studien milder oder mittelschwerer Ausprägung. Die Inzidenz der TEAEs mit möglichem oder wahrscheinlichem Zusammenhang mit der Behandlung und der schweren TEAEs war in der Studie 007 sowohl in der Placebo- als auch der Ataluren-Gruppe höher als in der Studie 020. In beiden Studien waren die Prüfarzte angehalten, eine Grad 3 (schwere) Progression der Erkrankung zu berichten, wenn ein Patient seine Gehfähigkeit verlor. In der Studie 007 wurde bei 10.5% der Patienten der Placebo-Gruppe und bei 7.0% der Patienten der Ataluren-Gruppe eine Grad 3-Progression der Erkrankung dokumentiert. Auch in der Studie 020 wurde bei Patienten der Placebo-Gruppe etwas häufiger eine schwere Progression der Erkrankung (Grad 3) berichtet als bei Patienten der Ataluren-Gruppe (5.2% vs. 1.7%). Hinsichtlich der Inzidenz der schwerwiegenden TEAEs ergaben sich zwischen den Studien keine relevanten Unterschiede und nur insgesamt zwei Patienten brachen die Behandlung aufgrund eines TEAEs ab. Hinsichtlich des Musters der häufigsten TEAEs ergaben sich zwischen den Studien keine offensichtlichen, klinisch relevanten Unterschiede.



Tabelle 4-64: Vergleich der TEAEs in Studie 070 und 020 (8,9)

TEAE	Studie 007		Studie 020	
	Placebo (n=57)	Ataluren (n=57)	Placebo	Ataluren
Patienten mit $\geq 1$ UE	56 (98.2)	55 (96.5)	101 (87.8)	103 (89.6)
Zusammenhang mit der Behandlung				
Möglich	20 (35.1)	25 (43.9)	18 (15.7)	27 (23.5)
Wahrscheinlich	6 (10.5)	5 (8.8)	6 (5.2)	12 (10.4)
Grad 3 (schweres) UE	9 (15.8)	8 (14.0)	9 (7.8)	7 (6.1)
Grad 4 (Lebensbedrohliches) UE	0	0	0	0
Abbruch aufgrund UE	0	0	1 (0.9)	1 (0.9)
Schwerwiegendes UE	3 (5.3)	2 (5.3)	4 (3.5)	4 (3.5)
Häufigste UEs ( $\geq 20\%$ in einer Gruppe)				
Erbrechen	22 (38.6)	32 (56.1)	21 (18.3)	26 (22.6)
Kopfschmerz	14 (24.6)	22 (38.6)	21 (18.3)	21 (18.3)
Diarrhoe	14 (24.6)	11 (19.3)	10 (8.7)	20 (17.4)
Nasopharyngitis	13 (22.8)	13 (22.8)	22 (19.1)	24 (20.9)
Fieber	12 (21.1)	14 (24.6)	12 (10.4)	16 (13.9)

Quelle (Studienbericht 070): Tabelle 39, Tabelle 40, Tabelle 41

Quelle (Studienbericht 020): Tabelle 20, Tabelle 21

#### 4.4.2.5.2 Vergleichbarkeit der Analyse-Populationen der Meta-Analysen

Für die Darstellung der Ergebnisse der Meta-Analyse wurden folgende Dokumente herangezogen:

- CTD-Modul 2.7.3 inklusive Appendix (ISE) (14,15)
- CTD-Modul 2.7.4 (16).

Die „as-treated“ Population (Sicherheitsanalyse) der Meta-Analysen umschloss 172 Patienten pro Gruppe (16). Die Gesamtpopulation der Meta-Analysen (kITT Studie 007 und ITT Studie 020) umfasste 342 randomisierte Patienten, von denen 334 (97.7%) die Behandlung beendeten. Die 6MWD-Subpopulationen (Placebo/Ataluren; kITT/ITT) schlossen 34/39 (<300 m), 74/69 ( $\geq 300$  bis <400 m) und 63/63 ( $\geq 400$  m) bzw. 64/66 (<350 m) und 107/105 ( $\geq 350$  m) randomisierte Patienten ein. Nur einzelne Patienten brachen die Studien vorzeitig ab

(14). Tabelle 4-65 zeigt die Disposition der Patienten in Analyse-Populationen der Meta-Analyse.

Tabelle 4-65: Disposition der Patienten in den Analyse-Populationen der Meta-Analysen

Kategorie	007 kITT/020 ITT		007 kITT/020 ITT Subgruppe BL- 6MWD $\geq$ 300 bis <400 m	
	Placebo	Ataluren	Placebo	Ataluren
Randomisierte Patienten	171 (100.0)	171 (100.0)	74 (100.0)	69 (100.0)
Patienten in Woche 48	167 (98.8)	167 (97.7)	72 (97.3)	67 (97.1)
Vorzeitiger Abbruch	3 (1.8)	4 (2.3)	2 (2.7)	2 (2.9)

Quelle (ISE): Tabelle 1.1.1, Tabelle 1.1.1-2, Tabelle 1.1.7 (15)

Tabelle 4-67 zeigt die demographischen und andere BL-Charakteristika der in die Meta-Analyse eingeschlossenen Patienten. Dargestellt sind die Daten der Primärpopulation (Studie 007 kITT, Studie 020 ITT), und der Subgruppe mit einer BL-6MWD von  $\geq$ 300 bis <400 m (007 kITT/020 ITT). Die demographischen Charakteristika waren zwischen den Gruppen insgesamt vergleichbar (14,15). Das Alter schwankte zwischen fünf und 20 Jahren und war in der Placebo- und Ataluren-Gruppe ähnlich. Insgesamt 81.3% bzw. 82.5% der Patienten der Placebo- bzw. Ataluren-Gruppe waren Kaukasier. Der Anteil der Patienten, die Kortikosteroide einnahmen, war zwischen den Gruppen ausgeglichen. Etwas mehr Patienten der Placebo- als der Ataluren-Gruppe nahmen Herz-/Kreislaufmedikamente ein (18.0% vs. 12.8%). Etwas mehr Patienten der Placebo- als der Ataluren-Gruppe nahmen Herz-Kreislauf-Medikamente ein (Tabelle 4-66) (16).

Tabelle 4-66: Begleitbehandlung in den Studien 070 und 020 – Gepoolte Analyse („as-treated“) (16)

Begleitbehandlung, n (%)	Placebo (n=172)	Ataluren (n=172)
Systemische Kortikosteroide	126 (73.3)	120 (69.8)
Deflazacort	58 (33.7)	52 (30.2)
Prednisolon	31 (18.0)	31 (18.0)
Prednison	41 (23.8)	39 (22.7)
Hydrokortison	0 (0.0)	1 (0.6)
Kardiale Medikation	31 (18.0)	22 (12.8)
ACE-Inhibitoren	26 (15.1)	19 (11.0)
Angiotensin II-Antagonisten	4 (2.3)	4 (2.3)
Beta-Blocker	4 (2.3)	2 (1.2)

Arzneimittel wurden als Begleitmedikation eingestuft, wenn sie zu irgendeinem Zeitpunkt während der Behandlung eingenommen wurde.

Summary of Clinical Safety: Tabelle 19 und 20 (16)

Hinsichtlich der demographischen und anderer Charakteristika zu Beginn der Studie zeigten sich in der Gesamt-Population und der Subgruppe mit einer BL-6MWD von  $\geq 300$  bis  $< 400$  m keine auffälligen Unterschiede zwischen der Placebo- und der Ataluren-Gruppe. Die mittlere BL-6MWD betrug  $362.23 \pm 83.219$  m (Maximum 533.3 m) in der Placebo-Gruppe und  $360.13 \pm 81.798$  m (Maximum 525.25 m) in der Ataluren-Gruppe. Die Mediane von 368.00 bzw. 373.25 m zeigen, dass ein nicht unerheblicher Teil der Patienten beider Gruppen zu Beginn der Studien noch ein vergleichsweise gutes Gehvermögen besaß. Die demographischen Charakteristika unterschieden sich nicht wesentlich zwischen der Gesamt-Population und den Patienten mit einer BL-6MWD von  $\geq 300$  bis  $< 400$  m. Die Dauer der Kortikosteroid-Vorbehandlung (in Prozent Patienten) war in beiden Populationen ähnlich.

Die Patientencharakteristika unterschieden sich erwartungsgemäß zwischen den 6MWD-Gruppen ( $< 300$ ,  $\geq 300$  bis  $< 400$ ,  $\geq 400$  m; Tabelle 4-68). Die Patienten, die initial weniger als 300 m im 6MWT zurücklegen konnten, waren etwas älter als diejenigen der anderen Subpopulationen und waren entsprechend etwas größer und hatten ein etwas höheres Körpergewicht. Der Anteil der Patienten, die 9 Jahre oder älter waren, nahm von der Gruppe mit dem schlechtesten Gehvermögen zu der Gruppe mit dem besten Gehvermögen sukzessive ab. Das etwas höhere Alter in der erstgenannten Gruppe ( $< 300$  m) spiegelt vermutlich das Fortschreiten der Erkrankung mit zunehmendem Alter wider. Das mittlere Alter der Patienten mit einer BL-6MWD von  $\geq 300$  bis  $< 400$  m und  $\geq 400$  m war insgesamt vergleichbar. Die schon eingeschränkte Muskelfunktion (Gehfähigkeit) in der Gruppe mit einer 6MWD unter 300 m zeigte sich auch bei den TFTs, wie dies auch schon in der Studie 020 beobachtet wurde. Die proximale Muskelfunktion gemessen mittels TFTs nahm sukzessive von denjenigen mit bereits eingeschränkten bis zu denjenigen mit noch relativ guten körperlichen Fähigkeiten zu. Der Anteil der Patienten mit einer Gehstrecke unter 300 m, die über

≥12 Monate mit Kortikosteroiden vorbehandelt worden waren, war in der Placebo-Gruppe höher als in der Ataluren-Gruppe (85.3% vs. 69.2%). Da die körperlichen Funktionen (6MWD und TFTs) zwischen der Placebo- und Ataluren-Gruppe insgesamt vergleichbar waren, erscheint es unwahrscheinlich, dass dieser Befund durch eine zwischen den Gruppen unterschiedliche Schwere der Erkrankung bedingt sein könnte. In der Gruppe mit der besten Gehfähigkeit war der Anteil der Patienten, die über ≥12 Monate mit Kortikosteroiden vorbehandelt worden waren, tendenziell niedriger als bei denjenigen mit der mittleren Gehstrecke. Die Unterschiede erscheinen nicht so gravierend, dass sie die konsistenten und deutlich unterschiedlichen Behandlungseffekte zwischen den Subpopulationen allein erklären könnten.

Tabelle 4-67: Demographische und andere Charakteristika der Patienten – Meta-Analysen

Charakteristik	Meta-Analyse: 007 kITT/020 ITT		Meta-Analyse: BL-6MWD $\geq$ 300 bis <400 m 007 kITT/020 ITT	
	Placebo (n=171)	Ataluren (n=171)	Placebo (n=74)	Ataluren (n=69)
Alter (Jahre) M ( $\pm$ SD) Minimum, Maximum	8.8 (1.93) 5, 15	8.9 (2.21) 5, 20	8.5 (1.84) 5,15	8.7 (2.43) 5, 20
Altersgruppe, n (%) <9 Jahre $\geq$ 9 Jahre	85 (49.7) 86 (50.3)	89 (52.0) 82 (48.0)	38 (51.4) 36 (48.6)	35 (50.7) 34 (49.3)
Alter bei Diagnosestellung (Jahre), M ( $\pm$ SD)	4.2 (2.43)	4.0 (2.02)	4.3 (2.58)	4.0 (2.12)
Zeit seit Diagnosestellung (Jahre), M ( $\pm$ SD)	4.53 (2.500)	4.87 (2.777)	4.28 (2.558)	4.72 (3.112)
Kaukasier oder weiß, n (%)	139 (81.3)	141 (82.5)	57 (77.0)	57 (82.6)
Größe (cm) M ( $\pm$ SD)	125.32 (10.946)	125.50 (12.368)	124.48 (9.982)	123.15 (11.669)
Gewicht (kg) M ( $\pm$ SD)	29.99 (10.058)	31.29 (11.205)	29.20 (9.877)	29.94 (11.480)
BMI (kg/m <sup>2</sup> ) M ( $\pm$ SD)	18.71 (3.942)	19.36 (4.103)	18.46 (3.963)	19.22 (4.326)
BL-6MWD (m) M ( $\pm$ SD) Minimum, Maximum	362.23 (83.219) 142.5, 533.3	360.13 (81.798) 75.0, 525.0	350.90 (29.052) 300.0, 399.5	354.88 (27.403) 301.6, 398.0
BL-6MWD, n (%) <350 m $\geq$ 350 m	64 (37.4) 107 (62.6)	66 (38.6) 105 (61.4)	30 (40.5) 44 (59.5)	27 (39.1) 42 (60.9)

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

Charakteristik	Meta-Analyse: 007 kITT/020 ITT		Meta-Analyse: BL-6MWD $\geq 300$ bis $< 400$ m 007 kITT/020 ITT	
	Placebo (n=171)	Ataluren (n=171)	Placebo (n=74)	Ataluren (n=69)
BL-6MWD, n (%)				
<300 m	34 (19.9)	39 (22.8)	0 (0.0)	0 (0.0)
$\geq 300$ bis $< 400$ m	74 (43.3)	69 (40.4)	74 (100.0)	69 (100.0)
$\geq 400$ m	63 (36.8)	63 (36.8)	0 (0.0)	0 (0.0)
4 Stufen hinauf (sec)	(n=168)	(n=168)	(n=73)	
M ( $\pm$ SD)	6.25 (5.645)	6.26 (5.600)	5.81 (3.264)	5.19 (2.262)
Minimum, Maximum	1.1, 30.0	1.3, 30.0	1.5, 18.5	1.6, 14.9
4 Stufen hinab (sec)	(n=165)	(n=168)	(n=73)	
M ( $\pm$ SD)	5.02 (4.973)	5.38 (5.340)	4.64 (3.069)	4.22 (1.957)
Minimum, Maximum	0.0, 30.0	1.4, 30.0	1.1, 19.7	1.7, 13.4
10 m rennen/ gehen (sec)	(n=171)	(n=171)		
M ( $\pm$ SD)	6.77 (2.832)	6.90 (3.567)	6.68 (1.843)	6.61 (3.264)
Minimum, Maximum	3.2, 17.0	3.0, 30.0	3.7, 12.0	3.1, 30.0
Kortikosteroid-Vorbe- handlung, n (%)				
6 bis $< 12$ Monate	24 (14.0)	28 (16.4)	9 (12.2)	11 (15.9)
$\geq 12$ Monate	130 (76.0)	127 (74.3)	58 (78.4)	54 (78.3)
Quelle	ISE Tabelle 2.1.1, 2.1.2 (15)		ISE Tabelle 2.1.1-2, Tabelle 2.1.2-2 (15)	

Tabelle 4-68: Demographische und andere Charakteristika der Patienten – Meta-Analysen der 6MWD-Subpopulationen

Charakteristik	Meta-Analyse: BL-6MWD <300 m 007 kITT/020 ITT		Meta-Analyse: BL-6MWD ≥300 bis <400 m 007 kITT/020 ITT		Meta-Analyse: BL-6MWD ≥400 m 007 kITT/020 ITT	
	Placebo (n=34)	Ataluren (n=39)	Placebo (n=74)	Ataluren (n=69)	Placebo (n=63)	Ataluren (n=63)
Alter (Jahre) M (±SD) Minimum, Maximum	9.9 (2.07) 5, 15	9.6 (2.11) 5, 15	8.5 (1.84) 5, 15	8.7 (2.43) 5, 20	8.4 (1.74) 5, 13	8.5 (1.92) 5, 15
Altersgruppe, n (%) <9 Jahre ≥9 Jahre	9 (26.5) 25 (73.5)	14 (35.9) 25 (64.1)	38 (51.4) 36 (48.6)	35 (50.7) 34 (49.3)	28 (60.3) 5 (39.7)	40 (63.5) 23 (36.5)
Alter bei Diagnosestellung (Jahre), M (±SD)	4.4 (2.28)	4.6 (1.77)	4.3 (2.58)	4.0 (2.12)	4.1 (2.36)	3.7 (2.01)
Zeit seit Diagnosestellung (Jahre), M (±SD)	5.53 (2.380)	4.99 (2.616)	4.28 (2.558)	4.72 (3.112)	4.28 (2.394)	4.95 (2.506)
Kaukasier oder weiß, n (%)	30 (88.2)	31 (79.5)	57 (77.0)	57 (82.6)	52 (82.5)	53 (84.1)
Größe (cm) M (±SD)	131.05 (11.200)	130.04 (11.943)	124.48 (9.982)	123.15 (11.669)	123.21 (11.002)	125.27 (12.787)
Gewicht (kg) M (±SD)	34.98 (12.509)	35.95 (11.304)	29.20 (9.877)	29.94 (11.480)	28.22 (7.851)	29.89 (10.186)
BMI (kg/m <sup>2</sup> ) M (±SD)	19.93 (5.154)	20.90 (4.700)	18.46 (3.963)	19.22 (4.326)	18.33 (2.991)	18.57 (3.155)
BL-6MWD (m) M (±SD)	235.97 (48.791)	243.08 (52.520)	350.90 (29.052)	354.88 (27.403)	443.67 (33.064)	438.33 (31.124)
BL-6MWD, n (%) <350 m ≥350 m	34 (100.0) 0	39 (100.0) 0	30 (40.5) 44 (59.5)	27 (39.1) 42 (60.9)	0 63 (100.0)	0 63 (100.0)

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

Charakteristik	Meta-Analyse: BL-6MWD <300 m 007 kITT/020 ITT		Meta-Analyse: BL-6MWD ≥300 bis <400 m 007 kITT/020 ITT		Meta-Analyse: BL-6MWD ≥400 m 007 kITT/020 ITT	
	Placebo (n=34)	Ataluren (n=39)	Placebo (n=74)	Ataluren (n=69)	Placebo (n=63)	Ataluren (n=63)
4 Stufen hinauf (sec) M (±SD) Minimum, Maximum	13.72 (7.975) 4.3, 30.0	13.91 (7.511) 5.4, 30.0	5.81 (3.264) 1.5, 18.5	5.19 (2.262) 1.6, 14.9	2.96 (1.282) 1.1, 7.4	3.07 (1.209) 1.3, 6.3
4 Stufen hinab (sec) M (±SD) Minimum, Maximum	10.98 (8.313) 0.0, 30.0	11.98 (8.284) 4.5, 30.0	4.64 (3.069) 1.1, 19.7	4.22 (1.957) 1.7, 13.4	2.72 (1.075) 1.2, 5.6	2.89 (0.895) 1.4, 5.4
10 m rennen/ gehen (sec) M (±SD) Minimum, Maximum	10.86 (2.529) 6.5, 17.0	10.83 (3.588) 6.6, 25.0	6.68 (1.843) 3.7, 12.0	6.61 (3.264) 3.1, 30.0	4.68 (0.986) 3.2, 8.5	4.79 (0.983) 3.0, 7.4
Kortikosteroid-Vorbe- handlung, n (%) 6 bis <12 Monate ≥12 Monate	1 (2.9) 29 (85.3)	6 (15.4) 27 (69.2)	9 (12.2) 58 (78.4)	11 (15.9) 54 (78.3)	14 (22.2) 43 (68.3)	11 (17.5) 46 (73.0)
Quelle	ISE Tabelle 2.1.1-1, 2.1.2-1 (15)		ISE Tabelle 2.1.1-2, Tabelle 2.1.2-2 (15)		ISE Tabelle 2.1.1-3, Tabelle 2.1.2-3 (15)	



Zusammenfassend werden die Meta-Analyse aus folgenden Gründen als gerechtfertigt angesehen:

- Beide Studien waren Placebo-kontrollierte randomisierte Studien ähnlichen Designs jeweils über 48 Wochen. In beiden Studien wurde die zugelassene Ataluren-Dosierung geprüft.
- Es wurde keine kontrollierte Ataluren-Studie aus den Meta-Analysen ausgeschlossen.
- Die Fragestellung und der primäre Endpunkt waren in beiden Studien identisch. Die in den Meta-Analysen ausgewerteten patientenrelevanten Endpunkte sind allgemein als valide Instrumente in DMD-Studien anerkannt (24).
- Trotz enger gefasster Einschlusskriterien in der Studie 020 waren die Patienten-Charakteristika beider Studien insgesamt vergleichbar.
- Die Patientencharakteristika der beiden Behandlungsgruppen waren in den Meta-Analysen jeweils vergleichbar.
- Die Heterogenitätstests (siehe unten) lieferten keine Hinweise auf eine bedeutende Heterogenität der Behandlungseffekte zwischen den Studien.
- Aufgrund der vergleichsweise höheren Fallzahlen liefern die Meta-Analysen für eine seltene und klinisch sehr heterogene Erkrankung eine höhere Ergebnissicherheit (Evidenzstufe 1a) der Effektschätzung.

#### 4.4.2.5.3 Ergebnisse der Meta-Analysen

##### *Gehstrecke im 6MWT – Meta-Analysen*

In Tabelle 4-69 sind die Ergebnisse der Meta-Analysen für den primären Endpunkt (Änderung der Gehstrecke im 6MWT) der beiden Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD in der Gesamtpopulation zusammengefasst. Die Meta-Analysen wurden mit den Effektschätzern unterschiedlicher Analysepopulationen (kITT, ITT) und unterschiedlicher Auswertungsmethoden (prädefiniertes MMRM Modell, ANCOVA Modell, augmentiertes MMRM Modell) für die Studie PTC124-GD-007-DMD berechnet, um die Sensitivität auf den Gesamteffektschätzer zu adressieren. Die unterschiedlichen Analysepopulationen und Analysemethoden für die Studie PTC124-GD-007-DMD, die in Kapitel 4.3.1.3.1.1 näher beschrieben sind, ergeben in den Meta-Analysen mit der Studie PTC124-GD-020-DMD jeweils signifikante Gesamteffektschätzer. In keiner der Meta-Analysen ergab sich ein Hinweis auf Heterogenität der Effektschätzer zwischen den Studien.

Tabelle 4-69: Ergebnisse der Meta-Analysen für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

6MWT	Studien- bezeichnung	$\Delta$ LSM	SE	95% KI Start	95% KI Ende	Hetero- genität p-Wert	Signifi- kanz p-Wert
007: kITT, prädefinierte MMRM / 020: ITT, ANCOVA							
	MA	18.21		1.55	34.86	0.385	0.032
	007	28.63	14.70	-0.19	57.44		
	020	12.98	10.42	-7.43	33.39		
<i>Sensitivitätsanalysen</i>							
007: ITT, prädefinierte MMRM / 020: ITT, ANCOVA							
	MA	17.16		0.20	34.11	0.472	0.047
	007	26.44	15.53	-4.00	56.88		
	020	12.98	10.42	-7.43	33.39		
007: ITT, ANCOVA <sup>1</sup> / 020: ITT, ANCOVA							
	MA	18.46		2.08	34.84	0.378	0.027
	007	28.36	14.01	0.91	55.81		
	020	12.98	10.42	-7.43	33.39		
007: kITT, augmentiertes MMRM <sup>2</sup> / 020: ITT ANCOVA							
	MA	20.40		2.45	38.36	0.271	0.026
	007	31.71	13.46	5.33	58.09		
	020	12.98	10.42	-7.43	33.39		

<sup>1</sup>: als Sensitivitätsanalyse in der Studie 007 definiert (9)

<sup>2</sup>: zulassungsbegründende augmentierte MMRM Analyse (post-hoc) (9)

MA: Meta-Analyse;

kITT: korrigierte ITT Population (9)

*Gehstrecke im 6MWT – Meta-Analysen der Subgruppen (<300 m,  $\geq$ 300 m bis <400 m,  $\geq$ 400 m)*

Eine tabellarische Zusammenfassung der Subgruppenanalysen für den Endpunkt 6MWT findet sich in Tabelle 4-70 und Tabelle 4-71. Die Durchführung der Subgruppenanalysen ist in Kapitel 4.3.1.3.2.1 beschrieben. Der bewertungsrelevante Therapieeffekt für diesen Endpunkt ergibt sich aus dem Gesamteffektschätzer für die Subgruppe 300-400 m, da die Assay-Sensitivität in den anderen Subgruppen unter den konkreten Studienbedingungen für diesen Endpunkt reduziert ist (s. Kapitel 4.4.2.4).

Der in der Meta-Analyse sehr homogene Therapieeffekt von 45.33 m [95% KI: 20.16; 70.49; p=0.00042] im 6MWT stellt eine gegenüber der Vergleichstherapie bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens im Sinne von § 2 Absatz 3 AM-NutzenV dar, insbesondere eine Abschwächung der Verschlechterung schwerwiegender Symptome. Eine Verschlechterung der 6MWD um 30 Meter im Populationsmittel ist prognostisch für die Progression der Erkrankung und für den zukünftigen Verlust der Gehfähigkeit (46). Der Gesamtschätzer ist dabei noch unter einem Signifikanzniveau von 0.1% signifikant, und ist in dieser Größenordnung in Bezug auf das Signifikanzniveau mit einem dramatischen Effekt vergleichbar. Der Therapieeffekt im 6MWT ist daher als beträchtlicher Zusatznutzen zu quantifizieren.

Tabelle 4-70: Ergebnisse der Meta-Analysen der Subgruppen (&lt;300 m, ≥300 m bis &lt;400 m, ≥400 m) für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

Subgruppe	Studien- bezeichnung	n (Ataluren)	n (Placebo)	Δ LSM	SE	KI Start	KI Ende	Heterogenität zw. Studien		p-Wert	Test auf Subgruppen- unterschiede	
								p-Wert	I <sup>2</sup>		p-Wert	I <sup>2</sup>
300-400m	MA	69	74	45.33		20.16	70.49	0.79	0.00	0.0004		
	007	22	22	49.92	21.80	7.19	92.65					
	020	47	52	42.89	15.89	11.75	74.03					
<300m	MA	39	34	4.63		-30.99	40.24	0.44	0.00	0.7990		
	007	15	13	20.90	27.68	-33.35	75.15					
	020	24	21	-7.70	24.09	-54.92	39.52					
>400m	MA	63	63	0.96		-25.67	27.59	0.32	0.00	0.9438		
	007	20	22	18.39	22.19	-25.10	61.88					
	020	43	41	-9.50	17.19	-43.19	24.19					
Total		171	171	19.76		-1.62	41.13	0.14	39.27	0.0701	<b>0.037</b>	69.56

MA: Meta-Analyse

Tabelle 4-71: Ergebnisse der gepoolten Subgruppen für 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

Sub-gruppe	Studien-bezeichnung	n (Ataluren)	n (Placebo)	$\Delta$ LSM (m)	SE (m)	KI Start (m)	KI Ende (m)	p-Wert
300-400 m	MA	69	74	45.33	12.84	20.16	70.49	0.00042
<300m + >400m	MA	102	97	2.28	10.88	-19.05	23.60	0.83438

Die Patientenrelevanz der Therapieeffekte für den Endpunkt 6MWT wurde mittels einer Effektgrößenabschätzung (Hedges' g) bestimmt. Die Effektgröße des Therapieeffekts in der 300-400 m Subgruppe liegt dabei deutlich über dem Schwellenwert für einen patientenrelevanten Effekt (s. Tabelle 4-81).

#### *Zeit bis zur 10%igen Verschlechterung im 6MWT – Meta-Analyse*

Die Meta-Analyse der Ereigniszeitanalysen für die ITT Population beider eingeschlossenen Studien zeigte für den Endpunktparameter, die Hazard Ratio für die Zeit bis zur 10%igen Verschlechterung im 6MWT, einen signifikanten Behandlungsunterschied (HR: 0.67; 95% KI: 0.48–0.94). Zwischen den Effektschätzern der Einzelstudien bestand eine unbedeutende Heterogenität ( $I^2=0.0\%$ ;  $p=0.32$ ) (siehe Tabelle 4-72).

Da der Endpunkt die irreversible, deterministische Verschlechterung der schwerwiegenden Krankheitssymptomatik erfasst, und da das obere Konfidenzintervall der Hazard Ratio in der Meta-Analyse unter dem Schwellenwert von 0.95 liegt, wird der Zusatznutzen für diesen Endpunkt als beträchtlich eingestuft.

Tabelle 4-72: Ergebnisse der Meta-Analysen für Zeit bis zur 10%igen Verschlechterung im 6MWT aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

10%ige Verschlechterung	Studien-bezeichnung	HR	95% KI Start	95% KI Ende	Heterogenität p-Wert	Signifikanz p-Wert
007: ITT / 020: ITT						
	MA	0.67	0.48	0.94	0.32	0.021
	007	0.52	0.28	0.97		
	020	0.75	0.50	1.12		

*Zeit, um 4 Treppenstufen hinaufzusteigen – Meta-Analysen*

Für den Endpunkt ‚Zeit, um 4 Treppenstufen hinaufzusteigen‘ zeigten die Meta-Analysen unter Einschluss der gesamten Studienpopulationen für die unterschiedlichen Analysepopulationen und Auswertungsmethoden der Studie 007 jeweils signifikante Behandlungseffekte. Auch für diesen Endpunkt ergab sich keine Hinweise auf Heterogenität der Effekte zwischen den Studien (s. Tabelle 4-73).

Tabelle 4-73: Ergebnisse der Meta-Analysen für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

<b>4 Treppenstufen hinaufsteigen</b>	<b>Studienbezeichnung</b>	<b>Δ LSM</b>	<b>SE</b>	<b>95% KI Start</b>	<b>95% KI Ende</b>	<b>Heterogenität p-Wert</b>	<b>Signifikanz p-Wert</b>
<b>007: kITT, prädefinierte MMRM / 020: ITT ANCOVA</b>							
	MA	-1.74		-2.97	-0.51	0.45	0.0056
	007	-2.46	1.14	-4.70	-0.23		
	020	-1.43	0.75	-2.90	0.05		
<i>Sensitivitätsanalysen:</i>							
<b>007: ITT, prädefinierte MMRM / 020: ITT ANCOVA</b>							
	MA	-1.71		-2.95	-0.47	0.49	0.0069
	007	-2.39	1.17	-4.68	-0.10		
	020	-1.43	0.75	-2.90	0.05		
<b>007: cITT, augmentiertes MMRM / 020: ITT ANCOVA</b>							
	MA	-1.77		-3.00	-0.54	0.40	0.0049
	007	-2.57	1.15	-4.82	-0.32		
	020	-1.43	0.75	-2.90	0.05		

*Zeit, um 4 Treppenstufen hinaufzusteigen – Meta-Analysen der Subgruppen (<300 m, ≥300 m bis <400 m, ≥400 m)*

Eine tabellarische Zusammenfassung der Subgruppenanalysen für diesen Endpunkt findet sich in Tabelle 4-74 und Tabelle 4-75. Die Durchführung der Subgruppenanalysen ist in Kapitel 4.3.1.3.2.1 beschrieben. Der bewertungsrelevante Therapieeffekt für diesen Endpunkt ergibt sich aus dem Gesamteffektschätzer für die Subgruppe 300-400 m, da die Assay-Sensitivität in den anderen Subgruppen unter den konkreten Studienbedingungen für diesen Endpunkt nicht gegeben ist (s. Kapitel 4.4.2.4).

Der in der Meta-Analyse sehr homogene Therapieeffekt von -3.47 sec [95% KI: -5.35; -1.59; p=0.0003] stellt einen zumindest geringen Zusatznutzen von Ataluren gegenüber der Vergleichstherapie dar.

Tabelle 4-74: Ergebnisse der Meta-Analysen der Subgruppen (&lt;300 m, ≥300 m bis &lt;400 m, ≥400 m) für Zeit, um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

Subgruppe	Studien- bezeichnung	n (Ataluren)	n (Placebo)	Δ LSM	SE	KI Start	KI Ende	Heterogenität zw. Studien		p-Wert	Test auf Subgruppen- unterschiede	
								p-Wert	I <sup>2</sup>		p-Wert	I <sup>2</sup>
300-400m	MA	69	74	-3.47		-5.35	-1.59	0.96	0.00	0.0003		
	007	22	22	-3.40	1.74	-6.81	0.02					
	020	47	52	-3.50	1.15	-5.75	-1.25					
<300m	MA	39	34	-2.30		-6.22	1.61	0.15	51.01	0.2483		
	007	15	13	-4.51	2.18	-8.79	-0.24					
	020	24	21	-0.50	1.77	-3.97	2.97					
>400m	MA	63	63	-0.05		-2.04	1.95	0.73	0.00	0.9632		
	007	20	22	-0.55	1.77	-4.03	2.93					
	020	43	41	0.20	1.24	-2.23	2.63					
Total		171	171	-1.92		-3.53	-0.31	0.15	38.87	0.0195	<b>0.049</b>	66.86

MA: Meta-Analyse



Tabelle 4-75: Ergebnisse der gepoolten Subgruppen für Zeit um 4 Treppenstufen hinaufzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

<b>Sub- gruppe</b>	<b>Studien- bezeich- nung</b>	<b>n (Ataluren)</b>	<b>n (Placebo)</b>	<b>Δ LSM</b>	<b>SE</b>	<b>KI Start</b>	<b>KI Ende</b>	<b>p-Wert</b>
300-400 m	MA	69	74	-3.47	0.96	-5.35	-1.59	0.00030
<300m + >400m	MA	102	97	-0.86	0.91	-2.65	0.93	0.34638

Die Patientenrelevanz der Therapieeffekte für den Endpunkt Zeit, um 4 Treppenstufen hinaufzusteigen wurde mittels einer Effektgrößenabschätzung (Hedges'  $g$ ) bestimmt. Die Effektgröße des Therapieeffekts in der 300-400 m Subgruppe liegt dabei deutlich über dem Schwellenwert für einen patientenrelevanten Effekt (s. Tabelle 4-81).

#### *Zeit, um 4 Treppenstufen hinabzusteigen– Meta-Analysen*

Für den Endpunkt ‚Zeit, um 4 Treppenstufen hinabzusteigen‘ ergaben die Meta-Analysen unter Einschluss der gesamten Studienpopulationen für die unterschiedlichen Analysepopulationen und Auswertungsmethoden der Studie 007 jeweils signifikante Behandlungseffekte. Auch für diesen Endpunkt ergab sich kein Hinweis auf eine Heterogenität der Effekte zwischen den Studien (s. Tabelle 4-76).

Tabelle 4-76: Ergebnisse der Meta-Analysen für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

4 Treppenstufen hinabsteigen	Studienbezeichnung	$\Delta$ LSM	SE	95% KI Start	95% KI Ende	Heterogenität p-Wert	Signifikanz p-Wert
007: kITT, prädefinierte MMRM / 020: ITT, ANCOVA							
	MA	-1.89		-3.20	-0.59	0.85	0.004
	007	-1.70	1.24	-4.12	0.73		
	020	-1.98	0.79	-3.52	-0.43		
<i>Sensitivitätsanalysen:</i>							
007: ITT, prädefinierte MMRM / 020: ITT, ANCOVA							
	MA	-1.88		-3.19	-0.57	0.81	0.01
	007	-1.63	1.26	-4.10	0.85		
	020	-1.98	0.79	-3.52	-0.43		
007: cITT, augmentiertes MMRM / 020: ITT, ANCOVA							
	MA	-1.91		-3.22	-0.60	0.88	0.004
	007	-1.75	1.25	-4.20	0.71		
	020	-1.98	0.79	-3.52	-0.43		

*Zeit, um 4 Treppenstufen hinabzusteigen – Meta-Analysen der Subgruppen (<300 m,  $\geq$ 300 m bis <400 m,  $\geq$ 400 m)*

Eine tabellarische Zusammenfassung der Subgruppenanalysen für diesen Endpunkt findet sich in Tabelle 4-77 und Tabelle 4-78. Die Durchführung der Subgruppenanalysen ist in Kapitel 4.3.1.3.2.1 beschrieben. Der bewertungsrelevante Therapieeffekt für diesen Endpunkt ergibt sich aus dem Gesamteffektschätzer für die Subgruppe 300-400 m, da die Assay-Sensitivität in den anderen Subgruppen unter den konkreten Studienbedingungen für diesen Endpunkt nicht gegeben ist (s. Kapitel 4.4.2.4).

Der in der Meta-Analyse sehr homogene Therapieeffekt von -4.31 sec [95% KI: -6.25; -2.37; p=0.00001] stellt einen zumindest geringen Zusatznutzen von Ataluren gegenüber der Vergleichstherapie dar.

Tabelle 4-77: Ergebnisse der Meta-Analysen der Subgruppen (&lt;300 m, ≥300 m bis &lt;400 m, ≥400 m) für Zeit, um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

Subgruppe	Studien- bezeichnung	n (Ataluren)	n (Placebo)	Δ LSM	SE	KI Start	KI Ende	Heterogenität zw. Studien		p-Wert	Test auf Subgruppen- unterschiede	
								p-Wert	I <sup>2</sup>		p-Wert	I <sup>2</sup>
300-400m	MA	69	74	-4.31		-6.25	-2.37	0.89	0.00	0.0000		
	007	22	22	-4.08	1.87	-7.76	-0.41					
	020	47	52	-4.40	1.17	-6.68	-2.12					
<300m	MA	39	34	-0.90		-3.72	1.93	0.93	0.00	0.5340		
	007	15	13	-0.73	2.35	-5.33	3.87					
	020	24	21	-1.00	1.83	-4.58	2.58					
>400m	MA	63	63	-0.18		-2.23	1.87	0.91	0.00	0.8643		
	007	20	22	-0.36	1.91	-4.11	3.38					
	020	43	41	-0.10	1.25	-2.55	2.35					
Total		171	171	-1.93		-3.70	-0.15	0.11	45.08	0.0337	<b>0.011</b>	77.93

MA: Meta-Analyse

Tabelle 4-78: Ergebnisse der gepoolten Subgruppen für Zeit um 4 Treppenstufen hinabzusteigen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

<b>Sub- gruppe</b>	<b>Studien- bezeich- nung</b>	<b>n (Ataluren)</b>	<b>n (Placebo)</b>	<b>Δ LSM</b>	<b>SE</b>	<b>KI Start</b>	<b>KI Ende</b>	<b>p-Wert</b>
300-400 m	MA	69	74	-4.31	0.99	-6.25	-2.37	0.00001
<300m + >400m	MA	102	97	-0.43	0.85	-2.09	1.23	0.61168

Die Patientenrelevanz der Therapieeffekte für den Endpunkt Zeit, um 4 Treppenstufen hinabzusteigen wurde mittels einer Effektgrößenabschätzung (Hedges' g) bestimmt. Die Effektgröße des Therapieeffekts in der 300-400 m Subgruppe liegt dabei deutlich über dem Schwellenwert für einen patientenrelevanten Effekt (s. Tabelle 4-81).

#### *Zeit, um 10 m zu gehen oder zu laufen– Meta-Analysen*

Für den Endpunkt ‚Zeit, um 10 m zu gehen oder zu laufen‘ ergab die Meta-Analyse unter Einschluss der gesamten Studienpopulationen für die als primär definierte Analysepopulation und Auswertungsmethode der Studie 007 (kITT/prädefiniertes MMRM Model) ein signifikanter Behandlungseffekt ( $p=0.043$ ). Unter Einschluss der ITT Population (ein durch eine Verletzung verzerrter Baseline-Wert eines Patienten in der Placebo-Gruppe) und unter Verwendung des prädefinierten Analysemodells für die Studie 007 lag der p-Wert für den Gesamteffektschätzer knapp über dem Signifikanzniveau von 0.05. Unter Einschluss der kITT Population und unter Verwendung des augmentierten MMRM Modells zeigte sich eine signifikante Überlegenheit von Ataluren gegenüber Placebo ( $p=0.046$ ) (s. Tabelle 4-79).

Tabelle 4-79: Ergebnisse der Meta-Analysen für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

10m Gehen / Laufen	Studien- bezeich- nung	$\Delta$ LSM	SE	95% KI Start	95% KI Ende	Hetero- genität p-Wert	Signifi- kanz p-Wert
007: kITT, prädefinierte MMRM / 020: ITT, ANCOVA							
	MA	-1.16		-2.28	-0.03	0.81	0.043
	007	-1.37	1.05	-3.43	0.69		
	020	-1.07	0.68	-2.40	0.27		
<i>Sensitivitätsanalysen:</i>							
007: ITT, prädefinierte MMRM / 020: ITT, ANCOVA							
	MA	-1.08		-2.22	0.06	0.97	0.063
	007	-1.12	1.12	-3.31	1.07		
	020	-1.07	0.68	-2.40	0.27		
007: cITT, augmentiertes MMRM / 020: ITT, ANCOVA							
	MA	-1.15		-2.28	-0.02	0.81	0.046
	007	-1.37	1.08	-3.48	0.75		
	020	-1.07	0.68	-2.40	0.27		

*Zeit, um 10 m zu gehen oder zu laufen – Meta-Analysen der Subgruppen (<300 m,  $\geq$ 300 m bis <400 m,  $\geq$ 400 m)*

Eine tabellarische Zusammenfassung der Subgruppenanalysen für diesen Endpunkt findet sich in Tabelle 4-80. Die Durchführung der Subgruppenanalysen ist in Kapitel 4.3.1.3.2.1 beschrieben. Da der globale Interaktionstest für die Subgruppen keinen signifikanten p-Wert ergab, wurden die Effekte der Subgruppen nicht weiter berücksichtigt. Der bewertungsrelevante Therapieeffekt für diesen Endpunkt ergibt sich aus dem Gesamteffektschätzer für die Meta-Analyse unter Einschluss der gesamten Studienpopulationen (Abbildung 17). Das Ausmaß des Zusatznutzens ist für diesen Endpunkt nicht quantifizierbar.

Tabelle 4-80: Ergebnisse der Meta-Analysen der Subgruppen (&lt;300 m, ≥300 m bis &lt;400 m, ≥400 m) für Zeit, um 10 m zu gehen oder zu laufen aus PTC124-GD-007-DMD und PTC124-GD-020-DMD; Ataluren versus Placebo

Subgruppe	Studien- bezeichnung	n (Ataluren)	n (Placebo)	Δ LSM	SE	KI Start	KI Ende	Heterogenität zw. Studien		p-Wert	Test auf Subgruppen- unterschiede	
								p-Wert	I <sup>2</sup>		p-Wert	I <sup>2</sup>
300-400m	MA	69	74	-2.18		-3.82	-0.54	0.48	0.00	0.0093		
	007	22	22	-3.09	1.54	-6.10	-0.07					
	020	47	52	-1.80	1.00	-3.76	0.16					
<300m	MA	39	34	-2.18		-4.50	0.14	0.50	0.00	0.0651		
	007	15	13	-1.15	1.93	-4.93	2.63					
	020	24	21	-2.80	1.50	-5.73	0.13					
>400m	MA	63	63	0.09		-1.65	1.83	0.85	0.00	0.9210		
	007	20	22	-0.15	1.57	-3.23	2.93					
	020	43	41	0.20	1.08	-1.91	2.31					
Total		171	171	-1.34		-2.41	-0.27	0.41	1.49	0.0144	<b>0.129</b>	51.17

MA: Meta-Analyse

*Patientenrelevanz der Therapieeffekte für Morbiditätsendpunkte (Hedges' g):*

Um die Patientenrelevanz der Therapieeffekte aus den Meta-Analysen für die Subgruppen 300-400 m zu bestimmen, wurden die Therapieunterschiede zwischen Ataluren und Placebo als standardisierte Mittelwertsdifferenzen (Hedges' g) dargestellt. Für die Endpunkte 6MWT, Zeit, um 4 Treppen hinaufsteigen, und Zeit, um 4 Treppen hinabsteigen ergaben sich deutliche, patientenrelevante Effekte, da das Konfidenzintervall der Therapieeffekte über dem Schwellenwert von 0.2 lag (s. Tabelle 4-81).

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

Tabelle 4-81: Patientenrelevanz der Therapieeffekte für Morbiditätspunkte (Hedges' g)

Subgruppe	Studienbezeichnung	n (Ataluren)	n (Placebo)	$\Delta$ LSM	SE	KI Start	KI Ende	Heterogenität zw. Studien		Signifikanz p-Wert
								p-Wert	I <sup>2</sup>	
6MWT 300-400m	MA	69	74	<b>0.581</b>		<b>0.246</b>	<b>0.917</b>	0.710	0.000	0.0007
	007	22	22	0.678	0.311	0.069	1.287			
	020	47	52	0.539	0.205	0.138	0.941			
4 Treppen hinaufsteigen 300-400m	MA	69	74	<b>-0.598</b>		<b>-0.934</b>	<b>-0.263</b>	0.935	0.000	0.0005
	007	22	22	-0.578	0.308	-1.182	0.027			
	020	47	52	-0.608	0.206	-1.012	-0.204			
4 Treppen hinabsteigen 300-400m	MA	69	74	<b>-0.720</b>		<b>-1.060</b>	<b>-0.381</b>	0.771	0.000	<0.0001
	007	22	22	-0.646	0.310	-1.253	-0.038			
	020	47	52	-0.754	0.209	-1.163	-0.345			
10 m Gehen/ Laufen 300-400m	MA	69	74	-0.430		-0.763	-0.098	0.527	0.000	0.0111
	007	22	22	-0.594	0.309	-1.199	0.011			
	020	47	52	-0.360	0.203	-0.757	0.038			



*Verträglichkeit – Meta-Analysen*

Die Meta-Analysen der unerwünschten Ereignisse aus PTC124-GD-007-DMD und PTC124-GD-020-DMD bzw. der unerwünschten Ereignisse in den jeweiligen SOC's sind im Kapitel 4.3.1.3.1.6 dargestellt. Eine tabellarische Zusammenfassung der Ergebnisse aus den Meta-Analysen der beiden RCTs unter Einschluss der Safety-Population („as treated“) ist in Tabelle 4-82 zu finden. Für kein unerwünschtes Ereignis/SOC konnte in den Meta-Analysen ein signifikanter Unterschied zwischen den beiden Behandlungsarmen festgestellt werden.

Für die Endpunktkategorie Nebenwirkungen ergibt sich somit kein Schaden für eine Therapie mit Ataluren im Vergleich zu Placebo. Die Verträglichkeit des Wirkstoffs zeigt sich in der Zusammenfassung der Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD als außergewöhnlich gut.

Tabelle 4-82: Verträglichkeit – Ergebnisse der Meta-Analysen für PTC124-GD-007-DMD und PTC124-GD-020-DMD

SOC	UE (PT)	Studien- bezeichnung	Events Ataluren	N Total Ataluren	Events Placebo	N Total Placebo	RR (Ataluren/ Placebo)	95% KI Start	95% KI Ende	Meta-Analysen	
										Heterogenität p-Wert	Signifikanz p-Wert
<i>Nicht schwerwiegende (nicht schwere) unerwünschte Ereignisse</i>											
Erkrankungen des Gastrointestinaltrakts	MA		94	172	85	172	1.114	0.9219	1.345	0.803	0.2643
	007		42	57	37	57	1.135	0.8876	1.452		
	020		52	115	48	115	1.083	0.8065	1.455		
Erbrechen	MA		58	172	43	172	1.369	0.9984	1.877	0.622	0.0512
	007		32	57	22	57	1.455	0.9752	2.170		
	020		26	115	21	115	1.238	0.7407	2.069		
Diarrhöe	MA		31	172	24	172	1.250	0.4987	3.133	0.066	0.6341
	7		11	57	14	57	0.786	0.3905	1.581		
	020		20	115	10	115	2.000	0.9796	4.083		
Abdominalschmerz	MA		14	172	9	172	1.557	0.6933	3.497	0.787	0.2835
	007		7	57	4	57	1.750	0.5419	5.652		
	020		7	115	5	115	1.400	0.4576	4.283		
Flatulenz	MA		7	172	5	172	1.385	0.4538	4.229	0.733	0.5670

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

SOC	UE (PT)	Studien- bezeichnung	Events Ataluren	N Total Ataluren	Events Placebo	N Total Placebo	RR (Ataluren/ Placebo)	95% KI Start	95% KI Ende	Meta-Analysen	
										Heterogenität p-Wert	Signifikanz p-Wert
		007	5	57	4	57	1.250	0.3537	4.418		
		020	2	115	1	115	2.000	0.1839	21.751		
Abdominalbeschwerden		MA	3	172	4	172	0.873	0.0150	50.727	0.0496	0.9479
		007	0	57	4	57	0.111	0.0061	2.017		
		020	3	115	0	115	7.000	0.3656	134.008		
Nausea		MA	15	172	14	172	1.074	0.5377	2.146	0.850	0.8395
		007	8	57	7	57	1.143	0.4439	2.942		
		020	7	115	7	115	1.000	0.3623	2.760		
Erkrankungen des Nervensystems		MA	53	172	40	172	1.336	0.9442	1.891	0.592	0.1019
		007	25	57	17	57	1.471	0.8966	2.412		
		020	28	115	23	115	1.217	0.7479	1.982		
Kopfschmerzen		MA	43	172	35	172	1.248	0.8008	1.945	0.257	0.3276
		007	22	57	14	57	1.571	0.8971	2.753		
		020	21	115	21	115	1.000	0.5788	1.728		
Erkrankungen der Niere und der		MA	27	172	18	172	1.500	0.8582	2.620	0.785	0.1547
		007	10	57	6	57	1.667	0.6490	4.280		

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

SOC	UE (PT)	Studien- bezeichnung	Events Ataluren	N Total Ataluren	Events Placebo	N Total Placebo	RR (Ataluren/ Placebo)	95% KI Start	95% KI Ende	Meta-Analysen	
										Heterogenität p-Wert	Signifikanz p-Wert
<b>Harnwege</b>		020	17	115	12	115	1.417	0.7089	2.831		
	Enuresis	MA	6	172	1	172	3.672	0.5817	23.178	0.421	0.1665
		007	4	57	0	57	9.000	0.4957	163.401		
		020	2	115	1	115	2.000	0.1839	21.751		
<b>Allgemeine Erkrankungen und Beschwerden am Verabreichungsort</b>		MA	52	172	53	172	0.989	0.7215	1.357	0.556	0.9471
		007	23	57	21	57	1.095	0.6887	1.742		
		020	29	115	32	115	0.906	0.5890	1.394		
	Fieber	MA	30	172	24	172	1.244	0.7639	2.027	0.788	0.3799
		007	14	57	12	57	1.167	0.5922	2.298		
		020	16	115	12	115	1.333	0.6605	2.692		
	Progression einer Erkrankung	MA	13	172	20	172	0.650	0.3341	1.264	0.961	0.2044
		007	4	57	6	57	0.667	0.1987	2.237		
		020	9	115	14	115	0.643	0.2899	1.426		
<b>Infektionen und parasitäre Erkrankungen</b>		MA	101	172	93	172	1.050	0.7314	1.507	0.042	0.7923
		007	38	57	43	57	0.884	0.6980	1.119		
		020	63	115	50	115	1.260	0.9653	1.645		

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

SOC	UE (PT)	Studien- bezeichnung	Events Ataluren	N Total Ataluren	Events Placebo	N Total Placebo	RR (Ataluren/ Placebo)	95% KI Start	95% KI Ende	Meta-Analysen	
										Heterogenität p-Wert	Signifikanz p-Wert
Gastroenteritis	MA		14	172	9	172	1.549	0.6808	3.526	0.335	0.2967
	007		9	57	4	57	2.250	0.7348	6.890		
	020		5	115	5	115	1.000	0.2975	3.361		
Rhinitis	MA		14	172	6	172	2.316	0.9077	5.908	0.683	0.0789
	007		6	57	2	57	3.000	0.6319	14.243		
	020		8	115	4	115	2.000	0.6195	6.457		
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	MA		63	172	60	172	1.053	0.7988	1.389	0.872	0.7126
	007		28	57	26	57	1.077	0.7310	1.587		
	020		35	115	34	115	1.029	0.6936	1.528		
Sturz	MA		32	172	27	172	1.179	0.7380	1.884	0.445	0.4907
	007		11	57	7	57	1.571	0.6560	3.764		
	020		21	115	20	115	1.050	0.6027	1.829		
Kontusion	MA		9	172	7	172	1.285	0.4774	3.459	0.334	0.6196
	007		6	57	3	57	2.000	0.5256	7.611		
	020		3	115	4	115	0.750	0.1717	3.277		
Stoffwechsel- und	MA		23	172	17	172	1.327	0.7317	2.406	0.338	0.3518

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

SOC	UE (PT)	Studien- bezeichnung	Events Ataluren	N Total Ataluren	Events Placebo	N Total Placebo	RR (Ataluren/ Placebo)	95% KI Start	95% KI Ende	Meta-Analysen	
										Heterogenität p-Wert	Signifikanz p-Wert
<b>Ernährungsstörungen</b>		007	7	57	3	57	2.333	0.6348	8.576		
		020	16	115	14	115	1.143	0.5853	2.231		
Appetit vermindert		MA	8	172	4	172	1.988	0.6069	6.509	0.675	0.2564
		007	5	57	2	57	2.500	0.5057	12.360		
		020	3	115	2	115	1.500	0.2554	8.810		
<b>Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen</b>		MA	57	172	51	172	1.128	0.8261	1.541	0.391	0.4482
		007	25	57	19	57	1.316	0.8222	2.106		
		020	32	115	32	115	1.000	0.6595	1.516		
Rückenschmerzen		MA	20	172	13	172	1.539	0.7907	2.996	0.696	0.2045
		007	9	57	5	57	1.800	0.6428	5.041		
		020	11	115	8	115	1.375	0.5742	3.293		
<b>Schwere unerwünschte Ereignisse</b>											
schwere UEs Grad >=3		MA	15	172	18	172	0.836	0.4381	1.595	0.840	0.5869
		007	8	57	9	57	0.889	0.3692	2.140		
		020	7	115	9	115	0.778	0.2998	2.018		

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

SOC	UE (PT)	Studien- bezeichnung	Events Ataluren	N Total Ataluren	Events Placebo	N Total Placebo	RR (Ataluren/ Placebo)	95% KI Start	95% KI Ende	Meta-Analysen	
										Heterogenität p-Wert	Signifikanz p-Wert
<i>Schwerwiegende unerwünschte Ereignisse</i>											
	SUE	MA	6	172	9	172	0.680	0.2413	1.918	0.392	0.4662
		007	2	57	5	57	0.400	0.0809	1.978		
		020	4	115	4	115	1.000	0.2563	3.902		
<i>Studienabbruch auf Grund von unerwünschten Ereignissen</i>											
	Studienabbruch auf Grund von UE	MA	1	172	1	172	1.000	0.0633	15.796	1.000	1.0000
		007	0	57	0	57	0.000	0.0000	0.000		
		020	1	115	1	115	1.000	0.0633	15.796		

#### 4.4.2.5.4 Zusammenfassung der Ergebnisse der Meta-Analysen

##### *Wirksamkeitsergebnisse in der Gesamtpopulation*

Bei der DMD handelt es sich um eine sehr schwerwiegende, progredient verlaufende Erkrankung. Bei den in die Bewertung eingeschlossenen Studien handelte es sich um randomisierte placebokontrollierte Studien qualitativ hoher Ergebnissicherheit (Evidenzstufe 1b). Die Meta-Analysen ist der Evidenzstufe 1a zuzuordnen. Die Fragestellung und der primäre Endpunkt (Änderung der 6MWD) waren in beiden Studien identisch. 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. Die beiden in die Meta-Analyse eingeschlossenen Studien waren hinsichtlich aller Wirksamkeitsendpunkte in der Gesamtpopulation homogen ( $p > 0.2$ ), sodass es gerechtfertigt war die Ergebnisse der Studien in eine Meta-Analyse zusammenzuführen.

Die Behandlungseffekte waren in den Meta-Analysen über fünf unterschiedliche Endpunkte übereinstimmend signifikant zugunsten einer Behandlung mit Ataluren. Die Heterogenität der eingeschlossenen Patientenkollektive und die fehlende Assay-Sensitivität in bestimmten Subpopulationen wurde durch die größere Fallzahl der Meta-Analysen ausgeglichen. Da es sich um zwei homogene Studien hoher qualitativer Ergebnissicherheit handelt und die Unterschiede konsistent signifikant waren, kann entsprechend IQWiG – Allgemeine Methoden (7) ein Beleg abgeleitet werden, der der Evidenzstufe 1a entspricht.

Für den Endpunkt der 10%igen Verschlechterung der 6MWD ergab sich für die Gesamtpopulation der in die Meta-Analyse eingeschlossenen Studien ein beträchtlicher Zusatznutzen (Tabelle 4-83), da das Konfidenzintervall der Hazard Ratio unter dem Schwellenwert von 0.95 lag (HR: 0.67 [95% KI: 0.48; 0.94],  $p=0.021$ ) (Abbildung 7). Der Endpunkt ist patientenrelevant und reflektiert die irreversible, deterministische Verschlechterung der schwerwiegenden Krankheitssymptomatik.

Für den Endpunkt Zeit um 10 m zu gehen oder zu laufen zeigten sich in der Subgruppenanalyse keine signifikante Interaktion (Abbildung 63). Aus diesem Grund wird für diesen Endpunkt der Gesamteffektschätzer der Meta-Analyse unter Einschluss der Gesamtpopulationen für die Zusatznutzenbewertung herangezogen (Abbildung 15). Da für diesen Endpunkt kein formal validiertes Responsekriterium existiert, ist der Zusatznutzen als nicht quantifizierbar auszuweisen (Tabelle 4-83).

Für die drei Endpunkte Veränderung der Gehstrecke im 6MWT, Zeit um 4 Treppenstufen hinaufzusteigen, und Zeit um 4 Treppenstufen hinabzusteigen wurde der Zusatznutzen aus den Subgruppenanalysen abgeleitet.



### *Wirksamkeitsergebnisse der Subgruppen-Analysen*

Wie in Kapitel 4.4.2.4 dargelegt, sind Therapieeffekte bei den Subpopulationen mit einer Gehstrecke im BL-6MWT unter 300 m bzw. von  $\geq 400$  m unter den gewählten Studienbedingungen kaum bzw. nicht nachweisbar, wodurch der Behandlungseffekt von Ataluren in der Gesamtpopulation deutlich unterschätzt wird. Die homogenen Ergebnisse der beiden Studien und der Meta-Analysen bestätigen dies eindrücklich: Patienten mit einer Gehstrecke von  $\geq 400$  m bleiben über ein Jahr weitgehend klinisch stabil und bei Patienten mit einer Gehstrecke unter 300 m ist ein Großteil des Muskelgewebes irreversibel geschädigt, und die Erkrankung schreitet in der Regel so rasch voran, dass eine Dystrophin-restaurierende Therapie nur noch wenig Erfolg zeigen kann.

Die für die Nutzenbewertung relevanten Subgruppen-Analysen (BL-6MWD  $< 300$  m,  $\geq 300$  m bis  $< 400$  m,  $\geq 400$  m) werden zusammenfassend aus folgenden Gründen als gerechtfertigt und biologisch plausibel eingestuft:

- Die EMA als auch die FDA empfiehlt den Einschluss nicht zu restriktiver Patientenkollektive mit *a priori* definierten Subgruppen-Analysen (2,29). Laut EMA sollten in Anbetracht des progressiven Charakters der Erkrankung, verschiedene Cut-Offs auf geeigneten Skalen eingesetzt werden, um die Assay-Sensitivität zu gewährleisten (2).
- Die Subgruppen-Analyse war im SAP der Studie 020 prospektiv geplant (34). Es wurde bereits bei der Finalisierung des SAPs vermutet, dass sich die Effektgrößen abhängig von der Gehstrecke im 6MWT zur BL unterscheiden könnten wie dies bereits in der Studie 007 bei den 6MWD-Gruppen ( $< 350$  vs.  $\geq 350$  m) der Fall gewesen war (38).
- Die Subpopulation (6MWD  $\geq 300$  bis  $< 400$  m) war mit 43.3% (Placebo; n=74) und 40.4% (Ataluren; n=69) der in die Meta-Analyse eingeschlossenen Patienten vergleichsweise groß. Die deutlichen und konsistenten Unterschiede zwischen den Subpopulationen können nicht durch andere nicht mit der BL-6MWD in Zusammenhang stehende Patientencharakteristika hinreichend plausibel erklärt werden (15).
- Entsprechend der IQWiG – Allgemeine Methoden sollte der Effekt in der Gesamtpopulation bei Vorliegen eines Belegs eines unterschiedlichen Effektes in Subgruppen nicht herangezogen werden (7).
- Der positive Effekt von Ataluren auf die 6MWD und die TFTs bei Patienten mit einer 6MWD von  $\geq 300$  bis  $< 400$  m wurde in beiden RCTs sowie den Meta-Analyse nachgewiesen. Der Behandlungseffekt von Ataluren auf die Änderung der 6MWD war in beiden Studien auffallend homogen und klinisch bedeutsam (48.7 m Studie 007; 42.89 m Studie 020; ITT) (15). Die Patientenrelevanz des Behandlungseffekts von Ataluren wurde in dieser Subpopulation nachgewiesen (Hedges' g).
- Die signifikante Subgruppeninteraktion war nicht nur auf den 6MWT beschränkt, sondern der deutliche Therapieeffekt bei Patienten mit einer 6MWD von  $\geq 300$  m bis  $< 400$  m zeigte sich konsistent und über beide Studien homogen auch in den Zeit-Funktionstests

zur Muskelfunktion, was nach Sun et al. (2010) für tatsächliche Effektunterschiede mit hoher Ergebnissicherheit spricht (39). In den Meta-Analysen war die Atalurenbehandlung in dieser Subpopulation hinsichtlich der Änderung der 6MWD sowie der analysierten TFTs der Placebobehandlung hoch signifikant überlegen. Bei Patienten mit einer 6MWD unter 300 m bzw. von  $\geq 400$  m zeigte sich demgegenüber in den Meta-Analysen durchgängig in allen Tests keine signifikanten Effekt von Ataluren, allerdings positive Trends bei Patienten in einem fortgeschrittenen Stadium der Erkrankung (15).

- Das CHMP (2014) hebt in einem Richtlinienentwurf zu Subgruppen in konfirmatorischen Studien die Relevanz der biologischen Plausibilität hervor (40). Wie oben dargelegt sind die zwischen den Subpopulationen unterschiedlichen Effekte biologisch plausibel: Bei klinisch stabilen Patienten (6MWD  $\geq 400$  m) kann ein Behandlungseffekt innerhalb eines Jahres noch nicht nachgewiesen werden. Die Ergebnisse der Einzelstudien und der Meta-Analysen belegen dies eindrücklich. Bei Patienten in einem bereits weit fortgeschrittenen Stadium der Erkrankung (6MWD  $< 300$  m) ist ein Großteil der gesunden Muskelmasse durch Fettgewebe ersetzt (62), sodass eine Dystrophin-regenerierende Therapie bei diesen Patienten hinsichtlich des Erhalts der Gehfähigkeit bzw. Ausdauer nur noch wenig Aussicht auf Erfolg hat. Der Verlust der Gehfähigkeit bei einem Teil dieser Patienten führt zu hoher Variabilität, was den Nachweis eines Effekts bei kleiner Fallzahl erschweren bzw. verhindern kann.

Nach heutigem Erkenntnisstand ist davon auszugehen, dass der deutliche Behandlungseffekt von Ataluren bei den Patienten mit einer Gehstrecke von  $\geq 300$  bis  $< 400$  m verglichen mit den Patienten am oberen bzw. unteren Spektrum der Erkrankung eine höhere Sensitivität gegenüber den Testsystemen widerspiegelt und nicht einen Behandlungseffekt, der auf eine Subpopulation beschränkt ist. Auf Grund des Wirkmechanismus von Ataluren ist eine Wirksamkeit über das gesamte Krankheitsspektrum zu erwarten. Wie Studien zum natürlichen Verlauf zeigen (siehe z.B. (45)), verschlechtert sich die Gehfähigkeit bei Jungen in einem frühen Stadium der Erkrankung, die noch eine relativ lange Strecke im 6MWT zurücklegen können, langsam über Jahre. Daher ist ein Effekt einer muskelstabilisierenden Therapie bei diesen Patienten erst nach längerer Behandlung nachweisbar. Eine derartig lange Placebo-kontrollierte Studie wäre insbesondere angesichts der besonderen Schwere der Erkrankung ethisch nicht vertretbar. Es ist auch fraglich, ob eine hinreichend große Zahl Patienten angesichts einer mit Ataluren zugelassenen Therapie bereit wäre, an einer solchen Studie teilzunehmen.

Bei Patienten mit einer BL-Gehstrecke unter 300 m im 6MWT verhindert die durch das rasche Fortschreiten der Erkrankung bedingte hohe Variabilität den Nachweis eines signifikanten Behandlungseffekts bzw. die Muskeldystrophie ist bereits so weit fortgeschritten, dass die Patienten nicht mehr vollständig auf die Therapie ansprechen können. Bei diesen Patienten ist auch zu berücksichtigen, dass die körperlichen Funktionen in einer zeitlichen Reihenfolge verloren gehen (56), d.h. bei einigen Patienten beeinflusst eine Therapie die Fähigkeit, längere Strecken zu gehen, möglicherweise nicht mehr, könnte aber dennoch Effekte auf andere für die Patienten und deren Angehörige wichtige

Alltagstätigkeiten haben wie z.B. Treppen steigen, oder kurze Strecken zu laufen. Dies ist biologisch plausibel, da diese Tätigkeiten weniger Ausdauer erfordert als der 6MWT.

In der Zusammenschau erscheint es sinnvoll davon auszugehen, dass die deutlichen Therapieeffekte in der Subpopulation mit einer Gehstrecke zwischen  $\geq 300$  m und  $< 400$  m auch in der Population mit Gehstrecken  $> 400$  m nach längeren Beobachtungsdauern zu erwarten sind. Auch andere Autoren oder Institutionen gehen davon aus, dass eine Therapie der DMD möglichst früh begonnen werden sollte, um die initiale Schädigung des Muskelgewebes und die darauffolgende progressive Behinderung aufhalten zu können. Nach Merlini und Sabatelli (2015) ist der Abbau des Dystrophins bereits bei sehr jungen Kindern mit DMD biologisch und klinisch nachweisbar und es besteht kein Grund, eine Behandlung bis zum Alter von 7 Jahren hinauszuzögern, wenn die Muskelschwäche und der Verlust an gesundem Muskelgewebe bereits in einem fortgeschrittenen Stadium sind (49). Entsprechend dem Ataluren-Bericht des National Institute for Clinical Excellence (NICE) sollten Kinder mit DMD früh behandelt werden, wenn sie noch klinisch stabil sind, um den Eintritt in die „Decline“-Phase und den darauf folgenden Verlust der Gehfähigkeit zu verzögern (53). Auch die EMA kommt in der Richtlinie zu dem Schluss, dass Therapien wie z.B. Exon-Skipping, die eine gewisse Muskelqualität erfordern, bei jüngeren Kindern einen größeren Effekt auf die Muskelstärke und -funktion haben könnten als bei älteren Kindern in einem fortgeschrittenen Stadium der Muskeldystrophie (2).

Für den Endpunkt Veränderung der Gehstrecke im 6MWT ergibt sich aus dem Gesamteffektschätzer für die Subgruppe  $\geq 300$  m bis  $< 400$  m ein beträchtlicher Zusatznutzen ( $\Delta$  LSM: 45.33 m [95% KI: 20.16; 70.49];  $p=0.0004$ ) (Abbildung 52, Tabelle 4-83). Die Behandlung mit Ataluren führt, verglichen mit Placebo, zu einer Abschwächung schwerwiegender (bzw. schwerer) Symptome (bzw. Folgekomplikationen), in diesem Falle einer für die Patientenpopulation klinisch bedeutsamen geringeren Abnahme der Gehstrecke, die ein direktes Maß für die Krankheitsprogression und den zukünftigen Verlust der Gehfähigkeit darstellt (46). Der Therapieeffekt entspricht mit einem p-Wert von 0.0004 einer bisher nicht erreichten deutlichen Verbesserung des therapielevanten Nutzens, insbesondere eine Abschwächung der Verschlechterung schwerwiegender Symptome.

Für die Endpunkte Zeit, um 4 Treppenstufen hinaufzusteigen, und Zeit, um 4 Treppenstufen hinabzusteigen ergibt sich aus den Gesamteffektschätzern für die Subgruppe  $\geq 300$  m bis  $< 400$  m jeweils zumindest ein geringer Zusatznutzen (Tabelle 4-83). Die Therapieeffekte in den Meta-Analysen waren homogen und numerisch und sehr groß (Abbildung 57 und Abbildung 62). Für diese Endpunkte sind jedoch noch keine Responsekriterien auf Patientenebene validiert. Eine Darstellung der Therapieunterschiede zwischen den Behandlungsgruppen mittels der Effektgröße Hedges'g zeigte jedoch, dass die Effekte deutlich über dem Schwellenwert von patientenrelevanten Effekten lagen (Tabelle 4-81).

Die Sicherheitsdaten aus den Meta-Analysen der Studie 007 und der Studie 020 liefern keine Hinweise auf ein relevantes Schadenspotential von Ataluren welches sich negativ auf die Bewertung des Ausmaßes des Zusatznutzens auswirken könnte (Tabelle 4-82, Tabelle 4-83). Zudem waren die unter Ataluren berichteten UEs mehrheitlich milder bis mittelschwerer

Ausprägung und keines der schwerwiegenden UEs stand mit der Ataluren-Behandlung in Zusammenhang.

Tabelle 4-83: Ausmaß des Zusatznutzens auf Endpunktebene aus PTC124-GD-007-DMD und PTC124-GD-020-DMD (Meta-Analysen); Ataluren versus Placebo

Endpunktkategorie	Endpunkt	Effektschätzer Ataluren vs. Placebo (Woche 48)	Einstufung des Zusatznutzens
Mortalität			Endpunkt nicht durch das zugelassene Anwendungsgebiet abgedeckt
Morbidität	Veränderung der Gehstrecke im 6MWT	$\Delta$ LSM: 45.33 m [95% KI: 20.16; 70.49] p=0.0004 Hedges' g: 0.581 [95% KI: 0.246; 0.917] Meta-Analyse der Subpopulationen: $\geq$ 300 m bis <400 m BL-Gehstrecke kITT (007)/ITT (020)	Beträchtlich
	Zeit bis zur 10%igen Verschlechterung der 6MWD	HR: 0.67 [95% KI: 0.48; 0.94] p=0.021 Meta-Analyse der Gesamtpopulationen: kITT (007)/ITT (020)	Beträchtlich
	Zeit, um 4 Treppenstufen hinaufzusteigen	$\Delta$ LSM: -3.47 sec [95% KI: -5.35; -1.59] p=0.0003 Hedges' g: -0.598 [95% KI: -0.934; -0.263] Meta-Analyse der Subpopulationen $\geq$ 300 m bis <400 m BL-Gehstrecke kITT (007)/ITT (020)	Gering
	Zeit, um 4 Treppenstufen hinabzusteigen	$\Delta$ LSM: -4.31 sec [95% KI: -6.25; -2.37] p=0.00001 Hedges' g: -0.720 [95% KI: -1.060; -0.381] Meta-Analyse der Subpopulationen $\geq$ 300 m bis <400 m BL-Gehstrecke kITT (007)/ITT (020)	Gering
	Zeit, um 10 m zu gehen oder zu rennen	$\Delta$ LSM: -1.16 sec [95% KI: -2.28; -0.03] p=0.043 Meta-Analyse der Gesamtpopulationen: kITT (007)/ITT (020)	Nicht quantifizierbar
	NSAA Gesamtscore, transformiert	$\Delta$ LSM: 1.51 [95% KI: -1.16; 4.17] p=0.268	Kein Zusatznutzen belegt

Endpunktkategorie	Endpunkt	Effektschätzer Ataluren vs. Placebo (Woche 48)	Einstufung des Zusatznutzens
		Studie 020; ITT	
Gesundheitsbezogene Lebensqualität	PODCI Transfer/Basismobilität	$\Delta$ LSM: 1.64 [95% KI: -2.11; 5.39] p=0.391 Studie 020; ITT	Kein Zusatznutzen belegt
	PODCI Sport/körperliche Funktion:	$\Delta$ LSM: 2.15 [95% KI: -1.75; 6.05] p=0.281 Studie 020; ITT	Kein Zusatznutzen belegt
Verträglichkeit	TEAEs inklusive schwerwiegender (bzw. schwerer) TEAEs	Keine statistisch signifikanten Unterschiede  Meta-Analyse der Gesamtpopulationen: As treated (007) / as treated (020)	Kein Schadenspotential

In der **Gesamtschau** ergibt sich für Ataluren zur Behandlung der Duchenne-Muskeldystrophie infolge einer Nonsense-Mutation im Dystrophin-Gen bei gehfähigen Patienten im Alter ab 5 Jahren ein **beträchtlicher Zusatznutzen**.

#### 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 4-84: 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.

<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.

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 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.

#### 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).*

Studienbezeichnung	Quelle	Referenz
PTC124-GD-007-	Bushby K, Finkel R, Wong B, Barohn R, Campbell C, Comi GP, et al. Ataluren treatment of patients with nonsense mutation	(73)

<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.

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



Studienbezeichnung	Quelle	Referenz
DMD oder 007	dystrophinopathy. Muscle Nerve. 2014 Oct;50(4):477–87. (im Dossier verwendete Studienbezeichnung PTC124-GD-007-DMD oder 007)	
	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). (im Dossier verwendete Studienbezeichnung PTC124-GD-007-DMD oder 007)	(9)

PTC124-GD-020-DMD oder 020	Elfring G, Valencia X, Ong T. A Phase 3 Efficacy and Safety Study of Ataluren (PTC124) in Patients with Nonsense Mutation Dystrophinopathy. Sponsor: PTC Therapeutics, South Plainfield, USA, Protokoll-Nr. PTC124-GD-020-DMD, Statistischer Analyseplan Version 1.0, 19. August 2015. (im Dossier verwendete Studienbezeichnung PTC124-GD-020-DMD oder 020)	(34)
	McDonald C. A Phase 3 Efficacy and Safety Study of Ataluren (PTC124) in Patients with Nonsense Mutation Dystrophinopathy. Sponsor: PTC Therapeutics, South Plainfield, USA, Protokoll-Nr. PTC124-GD-020-DMD, Klinischer Studienbericht, 9. Dezember 2015. (im Dossier verwendete Studienbezeichnung PTC124-GD-020-DMD oder 020)	(8)
	McDonald C, Reha A, Elfring G, Spiegel R. A Phase 3 Efficacy and Safety Study of Ataluren (PTC124) in Patients with Nonsense Mutation Dystrophinopathy. Sponsor: PTC Therapeutics, South Plainfield, USA, Protokoll-Nr. PTC124-GD-020-DMD, Studienprotokoll Version Nr. 2.0, 4. März 2014. (im Dossier verwendete Studienbezeichnung PTC124-GD-020-DMD oder 020)	(44)
	PTC Therapeutics International Limited. Module 2.7.3 Summary of Clinical Efficacy. Ataluren (PTC124) – Addendum. 2015. (im Dossier verwendete Bezeichnung: Modul 2.7.3)	(14)
	PTC Therapeutics International Limited. Module 2.7.3 Summary of Clinical Efficacy. Ataluren (PTC124) – Addendum. Appendix (Integrated Summary of Efficacy). 2015. (im Dossier verwendete Bezeichnung: ISE)	(15)

	<p>PTC Therapeutics International Limited. Phase 3 Study of Ataluren in Patients With Nonsense Mutation Duchenne Muscular Dystrophy (ACT DMD). Studienregistereintrag NCT01826487. [Internet]. 2014 [zitiert 9. März 2016]. Verfügbar unter:  <a href="https://clinicaltrials.gov/ct2/show/NCT01826487?term=ataluren&amp;rank=3">https://clinicaltrials.gov/ct2/show/NCT01826487?term=ataluren&amp;rank=3</a>  (im Dossier verwendete Studienbezeichnung PTC124-GD-020-DMD oder 020)</p>	(42)
	<p>PTC Therapeutics International Limited. A Phase 3 Efficacy and Safety Study of Ataluren (PTC124) in Patients with Nonsense Mutation Dystrophinopathy. Studienregistereintrag EudraCT Nr. 2012-004527-20. [Internet]. 2015 [zitiert 9. März 2016]. Verfügbar unter: <a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=+A+Phase+3+Efficacy+and+Safety+Study+of+Ataluren+%28PTC124%29+in+Patients+with+Nonsense+Mutation+Dystrophinopathy">https://www.clinicaltrialsregister.eu/ctr-search/search?query=+A+Phase+3+Efficacy+and+Safety+Study+of+Ataluren+%28PTC124%29+in+Patients+with+Nonsense+Mutation+Dystrophinopathy</a>  (im Dossier verwendete Studienbezeichnung PTC124-GD-020-DMD oder 020)</p>	(43)

#### 4.7 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge), 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. European Medicines Agency (EMA). Assessment report for initial marketing authorisation application. Translarna (ataluren). Procedure No. EMEA/H/C/002720. Applicant: PTC Therapeutics Limited [Internet]. 2014 [zitiert 14. Februar 2016]. Verfügbar unter: [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)
2. European Medicines Agency (EMA). Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. EMA/CHMP/236981/2011, Corr. 1 [Internet]. 2015 [zitiert 11. Februar 2016]. Verfügbar unter: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2015/12/WC500199239.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/12/WC500199239.pdf)
3. Gemeinsamer Bundesausschuss. Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V. Bewertung von Arzneimitteln für seltene Leiden nach § 35a Absatz 1 Satz 10 i.V.m. 5. Kapitel § 12 Nr. 1 Satz 2 VerfO. Wirkstoff: Ataluren. 2 März 2015 [Internet]. 2015. Verfügbar unter: [https://www.g-ba.de/downloads/92-975-726/2015-02-27\\_Nutzenbewertung-G-BA\\_Ataluren.pdf](https://www.g-ba.de/downloads/92-975-726/2015-02-27_Nutzenbewertung-G-BA_Ataluren.pdf)

4. Bundesministerium für Gesundheit. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Ataluren. BAnz AT. 2015;B3:1–4.
5. European Medicines Agency (EMA), Committee for Orphan Medicinal Products. 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]. 2015 [zitiert 14. Februar 2016]. Verfügbar unter: [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)
6. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses. [Internet]. 2015 [zitiert 14. Februar 2016]. Verfügbar unter: [https://www.g-ba.de/downloads/62-492-1002/VerfO\\_2014-12-18\\_iK-2015-04-16.pdf](https://www.g-ba.de/downloads/62-492-1002/VerfO_2014-12-18_iK-2015-04-16.pdf)
7. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Allgemeine Methoden. Version 4.2 [Internet]. 2015 [zitiert 10. Februar 2016]. Verfügbar unter: [https://www.iqwig.de/download/IQWiG\\_Methoden\\_Version\\_4-2.pdf](https://www.iqwig.de/download/IQWiG_Methoden_Version_4-2.pdf)
8. McDonald C. A Phase 3 Efficacy and Safety Study of Ataluren (PTC124) in Patients with Nonsense Mutation Dystrophinopathy. Sponsor: PTC Therapeutics, South Plainfield, USA, Protokoll-Nr. PTC124-GD-020-DMD, Klinischer Studienbericht, 9. Dezember 2015.
9. Wong B, Barth J. 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, Klinischer Studienbericht. 2011.
10. PTC Therapeutics International Limited. SAS output. Analysis of 4 stair climb using predefined MMRM model based on cITT. 2016.
11. PTC Therapeutics International Limited. SAS output. Analysis of 4 stair descend using predefined MMRM model based on cITT. 2016.
12. PTC Therapeutics International Limited. SAS output. Analysis of 6MWD using predefined MMRM model based on cITT. 2016.
13. PTC Therapeutics International Limited. SAS output. Analysis of 10m run/walk using predefined MMRM model based on cITT. 2016.
14. PTC Therapeutics International Limited. MODULE 2.7.3 SUMMARY OF CLINICAL EFFICACY. ATALUREN (PTC124). 2015.
15. PTC Therapeutics International Limited. Module 2.7.3 Summary of Clinical Efficacy. Ataluren (PTC124). Appendix (Integrated Summary of Efficacy). 2015.
16. PTC Therapeutics International Limited. MODULE 2.7.4 SUMMARY OF CLINICAL SAFETY. ATALUREN (PTC124). 2015.

17. 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 [zitiert 5. Februar 2016]. Verfügbar unter: [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)
18. McDonald CM, Henricson EK, Abresch RT, Florence J, Eagle M, Gappmaier E, u. a. 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*. September 2013;48(3):357–68.
19. McDonald CM, Henricson EK, Abresch RT, Florence JM, Eagle M, Gappmaier E, u. a. 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*. September 2013;48(3):343–56.
20. Henricson E, Abresch R, Han JJ, Nicorici A, Goude Keller E, de Bie E, u. a. 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.
21. McDonald CM, Abresch RT, Carter GT, Fowler WM, Johnson ER, Kilmer DD, u. a. Profiles of neuromuscular diseases. Duchenne muscular dystrophy. *Am J Phys Med Rehabil Assoc Acad Physiatr*. Oktober 1995;74(5 Suppl):S70-92.
22. Mendell JR, Moxley RT, Griggs RC, Brooke MH, Fenichel GM, Miller JP, u. a. Randomized, double-blind six-month trial of prednisone in Duchenne’s muscular dystrophy. *N Engl J Med*. 15. Juni 1989;320(24):1592–7.
23. Griggs RC, Moxley RT, Mendell JR, Fenichel GM, Brooke MH, Pestronk A, u. a. Prednisone in Duchenne dystrophy. A randomized, controlled trial defining the time course and dose response. *Clinical Investigation of Duchenne Dystrophy Group. Arch Neurol*. April 1991;48(4):383–8.
24. Lynn S, Aartsma-Rus A, Bushby K, Furlong P, Goemans N, De Luca A, u. a. Measuring clinical effectiveness of medicinal products for the treatment of Duchenne muscular dystrophy. *Neuromuscul Disord NMD*. Januar 2015;25(1):96–105.
25. Mayhew AG, Cano SJ, Scott E, Eagle M, Bushby K, Manzur A, u. a. Detecting meaningful change using the North Star Ambulatory Assessment in Duchenne muscular dystrophy. *Dev Med Child Neurol*. November 2013;55(11):1046–52.
26. McDonald CM, McDonald DA, Bagley A, Sienko Thomas S, Buckon CE, Henricson E, u. a. Relationship between clinical outcome measures and parent proxy reports of health-related quality of life in ambulatory children with Duchenne muscular dystrophy. *J Child Neurol*. September 2010;25(9):1130–44.
27. Messina S, Vita GL, Sframeli M, Mondello S, Mazzone E, D’Amico A, u. a. Health-related quality of life and functional changes in DMD: a 12-month longitudinal cohort

- study. *Neuromuscul Disord* [Internet]. 2016; Verfügbar unter: <http://dx.doi.org/doi:10.1016/j.nmd.2016.01.003>
28. Wei Y, Speechley K, Campbell C. Health-Related Quality of Life in Children with DuchenneMuscular Dystrophy: A Review. *J Neuromuscul Dis.* 2. September 2015;2(3):313–24.
29. Food and Drug Administration. Duchenne Muscular Dystrophy and Related Dystrophinopathies: Developing Drugs for Treatment Guidance for Industry. DRAFT GUIDANCE [Internet]. 2015 [zitiert 10. Februar 2016]. Verfügbar unter: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM450229.pdf>
30. Goemans N, Klingels K, van den Hauwe M, Boons S, Verstraete L, Peeters C, u. a. Six-minute walk test: reference values and prediction equation in healthy boys aged 5 to 12 years. *PloS One.* 2013;8(12):e84120.
31. Bushby K, Finkel R, Birnkrant DJ, Case LE, Clemens PR, Cripe L, u. a. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. *Lancet Neurol.* Januar 2010;9(1):77–93.
32. Perkins KJ, Davies KE. Recent advances in Duchenne muscular dystrophy. *Degener Neurol Neuromuscul Dis.* 2012;2:141–64.
33. Humbertclaude V, Hamroun D, Bezzou K, Bérard C, Boespflug-Tanguy O, Bommelaer C, u. a. Motor and respiratory heterogeneity in Duchenne patients: implication for clinical trials. *Eur J Paediatr Neurol EJPN Off J Eur Paediatr Neurol Soc.* März 2012;16(2):149–60.
34. Elfring G, Valencia X, Ong T. A Phase 3 Efficacy and Safety Study of Ataluren (PTC124) in Patients with Nonsense Mutation Dystrophinopathy. Sponsor: PTC Therapeutics, South Plainfield, USA, Protokoll-Nr. PTC124-GD-020-DMD, Statistischer Analyseplan Version 1.0, 19. August 2015.
35. Wong B, Barth J. 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).
36. Bundesministeriums der Justiz und für Verbraucherschutz. Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - AM- NutzenV) [Internet]. 2014 [zitiert 14. Februar 2016]. Verfügbar unter: <http://www.gesetze-im-internet.de/bundesrecht/am-nutzenv/gesamt.pdf>
37. Lange S. [Challenges for clinical trials in oncology within the scope of early benefit assessment of drugs]. *Z Für Evidenz Fortbild Qual Im Gesundheitswesen.* 2015;109(6):417–30.

38. PTC Therapeutics International Limited. Dossier zur Nutzenbewertung gemäß § 35a SGB V. Ataluren (Translarna™). Modul 4 A [Internet]. 2014 [zitiert 15. Februar 2016]. Verfügbar unter: [https://www.g-ba.de/downloads/92-975-725/2014-11-24\\_Modul4\\_Ataluren.pdf](https://www.g-ba.de/downloads/92-975-725/2014-11-24_Modul4_Ataluren.pdf)
39. Sun X, Briel M, Walter SD, Guyatt GH. Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ*. 2010;340:c117.
40. Committee for Medicinal Products for Human Use (CHMP). Guideline on the investigation of subgroups in confirmatory clinical trials. Draft. EMA/CHMP/539146/2013 [Internet]. 2014 [zitiert 25. Februar 2016]. Verfügbar unter: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2014/02/WC500160523.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/02/WC500160523.pdf)
41. Skipka G, Bender R. Intervention effects in the case of heterogeneity between three subgroups. Assessment within the framework of systematic reviews. *Methods Inf Med*. 2010;49(6):613–7.
42. PTC Therapeutics International Limited. Phase 3 Study of Ataluren in Patients With Nonsense Mutation Duchenne Muscular Dystrophy (ACT DMD). Studienregistereintrag NCT01826487. [Internet]. 2014 [zitiert 9. März 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT01826487?term=ataluren&rank=3>
43. PTC Therapeutics International Limited. A Phase 3 Efficacy and Safety Study of Ataluren (PTC124) in Patients with Nonsense Mutation Dystrophinopathy. Studienregistereintrag EudraCT Nr. 2012-004527-20. [Internet]. 2015 [zitiert 9. März 2016]. Verfügbar unter: <https://www.clinicaltrialsregister.eu/ctr-search/search?query=+A+Phase+3+Efficacy+and+Safety+Study+of+Ataluren+%28PTC124%29+in+Patients+with+Nonsense+Mutation+Dystrophinopathy>
44. McDonald C, Reha A, Elfring G, Spiegel R. A Phase 3 Efficacy and Safety Study of Ataluren (PTC124) in Patients with Nonsense Mutation Dystrophinopathy. Sponsor: PTC Therapeutics, South Plainfield, USA, Protokoll-Nr. PTC124-GD-020-DMD, Studienprotokoll Version Nr. 2.0, 4. März 2014.
45. Pane M, Mazzone ES, Sivo S, Sormani MP, Messina S, D’Amico A, u. a. Long term natural history data in ambulant boys with Duchenne muscular dystrophy: 36-month changes. *PLoS One*. 2014;9(10):e108205.
46. Haas M, Vlcek V, Balabanov P, Salmonson T, Bakchine S, Markey G, u. a. European Medicines Agency review of ataluren for the treatment of ambulant patients aged 5 years and older with Duchenne muscular dystrophy resulting from a nonsense mutation in the dystrophin gene. *Neuromuscul Disord NMD*. Januar 2015;25(1):5–13.
47. 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*. 1. Juli 2002;166(1):111–7.

48. McDonald CM, Henricson EK, Han JJ, Abresch RT, Nicorici A, Elfring GL, u. a. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. *Muscle Nerve*. April 2010;41(4):500–10.
49. Merlini L, Sabatelli P. Improving clinical trial design for Duchenne muscular dystrophy. *BMC Neurol*. 2015;15:153.
50. Mayhew A, Cano S, Scott E, Eagle M, Bushby K, Muntoni F, u. a. Moving towards meaningful measurement: Rasch analysis of the North Star Ambulatory Assessment in Duchenne muscular dystrophy. *Dev Med Child Neurol*. Juni 2011;53(6):535–42.
51. Ricotti V, Ridout DA, Pane M, Main M, Mayhew A, Mercuri E, u. a. The NorthStar Ambulatory Assessment in Duchenne muscular dystrophy: considerations for the design of clinical trials. *J Neurol Neurosurg Psychiatry*. Februar 2016;87(2):149–55.
52. CHMP. Translarna, INN-ataluren - Assessment report for initial marketing authorisation application [Internet]. 2014 [zitiert 21. Mai 2016]. Verfügbar unter: [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)
53. National Institute for Health and Care Excellence (NICE). Highly Specialised Technologies. Evaluation Ataluren for treating Duchenne muscular dystrophy caused by a nonsense mutation in the dystrophin gene [ID 428]. [Internet]. 2015 [zitiert 9. März 2016]. Verfügbar unter: <http://www.actionduchenne.org/wp-content/uploads/2015/10/ID-428-Ataluren-Evaluation-Report-Committee-Papers.pdf>
54. Goemans N, Kirschner J, Mercuri E. New Perspectives in the Management of Duchenne Muscular Dystrophy. Report based on Presentations at a Satellite Symposium held at the European Paediatric Neurology Society (EPNS) on 25 September 2013 in Brussels, Belgium. *Touch Med Media*. 2014;9(1):78–82.
55. Pane M, Mazzone ES, Sivo S, Fanelli L, De Sanctis R, D’Amico A, u. a. The 6 minute walk test and performance of upper limb in ambulant duchenne muscular dystrophy boys. *PLoS Curr*. 2014;6.
56. Henricson EK, Abresch RT, Cnaan A, Hu F, Duong T, Arrieta A, u. a. The cooperative international neuromuscular research group Duchenne natural history study: glucocorticoid treatment preserves clinically meaningful functional milestones and reduces rate of disease progression as measured by manual muscle testing and other commonly used clinical trial outcome measures. *Muscle Nerve*. Juli 2013;48(1):55–67.
57. Henricson E, Abresch R, Han JJ, Nicorici A, Goude Keller E, Elfring G, u. a. Percent-predicted 6-minute walk distance in duchenne muscular dystrophy to account for maturational influences. *PLoS Curr*. 2012;4:RRN1297.
58. Escolar DM, Hache LP, Clemens PR, Cnaan A, McDonald CM, Viswanathan V, u. a. Randomized, blinded trial of weekend vs daily prednisone in Duchenne muscular dystrophy. *Neurology*. 2. August 2011;77(5):444–52.

59. Mazzone E, Martinelli D, Berardinelli A, Messina S, D'Amico A, Vasco G, u. a. North Star Ambulatory Assessment, 6-minute walk test and timed items in ambulant boys with Duchenne muscular dystrophy. *Neuromuscul Disord NMD*. November 2010;20(11):712–6.
60. Mazzone ES, Pane M, Sormani MP, Scalise R, Berardinelli A, Messina S, u. a. 24 month longitudinal data in ambulant boys with Duchenne muscular dystrophy. *PloS One*. 2013;8(1):e52512.
61. Goemans N, van den Hauwe M, Wilson R, van Impe A, Klingels K, Buyse G. Ambulatory capacity and disease progression as measured by the 6-minute-walk-distance in Duchenne muscular dystrophy subjects on daily corticosteroids. *Neuromuscul Disord NMD*. August 2013;23(8):618–23.
62. Sweeney L, Vandeborne K. Developing Skeletal Muscle MRI/MRS as a biomarker for DMD therapeutic development. *Imaging DMD Project* [Internet]. 2014 [zitiert 13. Februar 2016]. Verfügbar unter: [http://www.parentprojectmd.org/site/DocServer/Session\\_8\\_-\\_Sweeney\\_b.pdf?docID=15384](http://www.parentprojectmd.org/site/DocServer/Session_8_-_Sweeney_b.pdf?docID=15384)
63. Food and Drug Administration. Guidance for Industry. Duchenne Muscular Dystrophy. Developing Drugs for Treatment over the Spectrum of Disease. [Internet]. 2014 [zitiert 8. März 2016]. Verfügbar unter: [http://www.parentprojectmd.org/site/DocServer/Guidance\\_Document\\_Submission\\_-\\_Duchenne\\_Muscular\\_Dystrop.pdf?docID=15283](http://www.parentprojectmd.org/site/DocServer/Guidance_Document_Submission_-_Duchenne_Muscular_Dystrop.pdf?docID=15283)
64. Santhera Pharmaceuticals. Efficacy and Tolerability of Idebenone in Boys With Cardiac Dysfunction Associated With Duchenne Muscular Dystrophy (DELPHI). NCT00654784, Studienregistereintrag [Internet]. 2011 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT00654784?term=walk&intr=Duchenne&rank=33>
65. Assistance Publique - Hôpitaux de Paris. Effect of Oral Glutamine on Muscle Mass and Function in Duchenne Muscular Dystrophy (MDB-GLN). NCT00296621, Studienregistereintrag [Internet]. 2007 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT00296621?term=walk&intr=Duchenne&rank=12>
66. GlaxoSmithKline. A Clinical Study to Assess the Efficacy and Safety of GSK2402968 in Subjects With Duchenne Muscular Dystrophy (DMD114044). NCT01254019, Studienregistereintrag [Internet]. 2014 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT01254019?term=6MWD&type=Intr&intr=Duchenne&rank=3>
67. GlaxoSmithKline. A Clinical Study to Assess Two Doses of GSK2402968 in Subjects With Duchenne Muscular Dystrophy (DMD) (DMD114876). NCT01462292, Studienregistereintrag [Internet]. 2015 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT01462292?term=walk&intr=Duchenne&rank=5>
68. Sarepta Therapeutics. Efficacy Study of AVI-4658 to Induce Dystrophin Expression in Selected Duchenne Muscular Dystrophy Patients. NCT01396239, Studienregister-Eintrag



- [Internet]. 2015 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT01396239?term=6MWT&type=Intr&intr=Duchenne&rank=4>
69. University Hospital, Basel, Switzerland. L-citrulline and Metformin in Duchenne's Muscular Dystrophy. NCT01995032, Studienregistereintrag [Internet]. 2015 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT01995032?term=6MWD&type=Intr&intr=Duchenne&rank=4>
70. Sarepta Therapeutics. Confirmatory Study of Eteplirsen in DMD Patients (PROMOVI). NCT02255552, Studienregister-Eintrag [Internet]. 2016 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT02255552?term=6MWT&type=Intr&intr=Duchenne&rank=3>
71. Pfizer. A Phase 2 Study to Evaluate the Safety, Efficacy, Pharmacokinetics and Pharmacodynamics of PF-06252616 in Duchenne Muscular Dystrophy. NCT02310763, Studienregistereintrag [Internet]. 2016 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT02310763?term=6MWD&type=Intr&intr=Duchenne&rank=5>
72. Sarepta Therapeutics. Study of SRP-4045 and SRP-4053 in DMD Patients (ESSENCE). NCT02500381, Studienregistereintrag [Internet]. 2016 [zitiert 18. Februar 2016]. Verfügbar unter: <https://clinicaltrials.gov/ct2/show/NCT02500381?term=6MWT&type=Intr&intr=Duchenne&rank=8>
73. Bushby K, Finkel R, Wong B, Barohn R, Campbell C, Comi GP, u. a. Ataluren treatment of patients with nonsense mutation dystrophinopathy. *Muscle Nerve*. Oktober 2014;50(4):477–87.

**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 4-85 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 4-85 inklusive eines Flow-Charts für den Patientenfluss.*

Tabelle 4-85 (Anhang): Studiendesign und -methodik für Studie PTC124-GD-020-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 Krankheitsprogression gemessen anhand des Verlusts der Gehfähigkeit (Abfall der 6MWD) bei Patienten mit nmDMD
<b>Methoden</b>		
<b>3</b>	Studiendesign	
<b>3a</b>	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Phase 3, multizentrisch, doppelblind, Placebo-kontrolliert, GCP Parallelgruppen-Studie, Randomisierung im Verhältnis 1:1 zu Placebo oder Ataluren (10, 20 bzw. 20 mg/kg KG)
<b>3b</b>	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<i>Amendment 1.1</i> (15.01.2013 vor Einschluss des ersten Patienten am 10.04.2013): - <i>Amendment 1.2</i> (13.06.2013): Aufnahme eines zusätzlichen Endpunkts (Zeit, um aus dem Liegen aufzustehen), weitere Änderungen als für die vorliegende Bewertung nicht relevant eingestuft <i>Amendment 2.0</i> (04.03.2014): Klarstellung des Einschlusskriteriums hinsichtlich der Gensequenzierung Klarstellung des Einschlusskriteriums 6MWD (keine mehr als 20%ige Änderung vom Screening- zum BL-Wert) Aktualisierung der Kortikosteroid-Therapie, insbesondere hinsichtlich Dosismodifikationen während der Studie Weitere Änderungen als für die vorliegende Bewertung nicht relevant eingestuft
<b>4</b>	Probanden / Patienten	
<b>4a</b>	Ein-/Ausschlusskriterien der Probanden / Patienten	<u>Einschlusskriterien</u> 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 2) Männliches Geschlecht 3) Alter $\geq 7$ Jahre und $\leq 16$ Jahre 4) Phänotypischer Nachweis einer Dystrophinopathie auf Basis des Auftretens charakteristischer klinischer Symptome oder Zeichen (d.h. proximale Muskelschwäche, schwankender Gang und Gowers Manöver) im Alter von 6 Jahren, erhöhte Serum-Kreatinkinase und anhaltende Gehschwierigkeiten



Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>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 äquivalente Organisation zertifiziert wurde</p> <p>6) Dokumentation, dass eine Blutprobe zum Nachweis der Nonsense-Mutation entnommen wurde</p> <p>7) Einnahme von systemischen Kortikosteroiden (Prednison, Prednisolon, oder Deflazacort) über mindestens sechs Monate unmittelbar vor Beginn der Studienbehandlung ohne signifikante Änderung der Dosis oder des Dosisregimes (nicht bezogen auf eine Änderung des Körpergewichts) über mindestens der Monate unmittelbar vor Beginn der Studienbehandlung und eine begründete Erwartung, dass die Dosis und das Dosisregime sich während der Dauer der Studie nicht signifikant ändern wird</p> <p>8) Valide Screening-6MWD von <math>\geq 150</math> m. Die valide Screening-6MWD muss <math>\leq 80\%</math> des nach Alter und Größe vorausgesagten Wertes liegen.</p> <p>9) Die Ergebnisse von zwei Baseline-6MWDs müssen als valide bestimmte werden und das Ergebnis des Baseline-6MWD an Tag 2 muss innerhalb von 20% des Baseline-6MWD an Tag 1 liegen.</p> <p>10) Der Baseline 6MWD (Mittelwert aus validen Werten an Tag 1 und 2) darf nicht um mehr als 20% von dem validen Screening 6MWD abweichen.</p> <p>11) Überprüfung, ob bestimmte Screening-Laborwerte (hepatisch, renal und Serumelektrolyte) innerhalb der spezifizierten Bereiche des Zentrallabors lagen</p> <p>12) 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> <p>1) Behandlung mit systemischen Aminoglykosid-Antibiotika innerhalb von drei Monaten vor Beginn der</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Studienbehandlung</p> <p>2) Einleitung einer systemischen Kortikosteroid-Therapie innerhalb von sechs Monaten vor Beginn der Studienbehandlung</p> <p>3) Änderung der systemischen Kortikosteroid-Therapie (z.B. Änderung des Wirkstofftyps, Änderung der Dosis, die nicht im Zusammenhang mit einer Änderung des Körpergewichts steht, Änderung des Regimes, Unterbrechung oder erneute Einleitung) innerhalb von drei Monaten vor Beginn der Studienbehandlung</p> <p>4) Jede Änderung (Einleitung, Änderung des Wirkstoffs, der Dosis, des Regimes, Unterbrechung, Absetzen, oder Wiedereinleitung) der Prophylaxe/Behandlung der kongestiven Herzinsuffizienz innerhalb von drei Monaten vor Beginn der Studienbehandlung</p> <p>5) Andauernde Behandlung mit Anti-Koagulantien vom Coumarin-Typ (z.B. Warfarin), Phenytoin, Tolbutamid oder Paclitaxel</p> <p>6) Frühere Therapie mit Ataluren</p> <p>7) Bekannte Überempfindlichkeit gegenüber einem der Inhaltsstoffe oder Hilfsstoffe der Studienmedikation</p> <p>8) Exposition mit einem anderen investigativen Arzneimittel innerhalb von drei Monaten vor Behandlungsbeginn</p> <p>9) Vorgeschichte eines großen chirurgischen Eingriffs innerhalb von sechs Wochen vor Beginn der Studienbehandlung</p> <p>10) Andauernde immunsuppressive Therapie (ausgenommen Kortikosteroide)</p> <p>11) Fortlaufende Teilnahme in einer anderen klinischen Studie (ausgenommen Studien, die durch PTC Therapeutics spezifisch zugelassen wurden)</p> <p>12) Erwartung eines großen therapeutischen Eingriffs (z.B. Skoliose-Operation) während der 12-monatigen Behandlungsphase der Studie</p> <p>13) Notwendigkeit einer Beatmungshilfe während des Tages</p> <p>14) Unkontrollierte klinische Symptome oder Zeichen einer kongestiven Herzinsuffizienz (ACC/AHA Stadium C oder D)</p> <p>15) Vorherige oder fortbestehende medizinische Zustände (z.B. Begleiterkrankungen, psychiatrische Beschwerden, Verhaltensauffälligkeiten, Alkoholismus, Substanzabusus), medizinische Vorgeschichte,</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		physikalische Befunde (z.B. Unterschenkelverletzung, die die Leistung im 6MWT beeinflussen könnte), 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
4b	Studienorganisation und Ort der Studiendurchführung	
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>Placebo</p> <p>Ataluren 10,10, 20 mg/kg KG</p> <p>Die Prüfpräparate sollten jeweils dreimal täglich morgens, mittags und abends oral eingenommen werden. Eine Einnahme innerhalb von ungefähr 30 min nach einer Mahlzeit wurde empfohlen. 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 nächsten Morgen eingehalten werden. Die geplante Behandlungsdauer betrug 48 Wochen.</p> <p>Die Prüfpräparate (Pulver) konnten mit Wasser, Milch oder halbfester Nahrung (Joghurt, Pudding oder Apfelkompott) gemischt werden. Der komplette Inhalt der Beutel sollte mit 30 ml der Flüssigkeit oder drei Esslöffeln der halbfesten Nahrung gut vermischt werden; je nach Vorliebe konnte die Menge der Flüssigkeit oder halbfesten Nahrung erhöht werden. Die Klinikmitarbeiter instruierten die Patienten oder deren Eltern/Betreuer ausführlich über die einzunehmende Zahl der Beutel und wie die Präparate einzunehmen sind. Zusätzlich wurden schriftliche Instruktionen ausgegeben.</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p><u>Primäres Zielkriterium:</u> Änderung der 6MWD von BL bis Woche 48</p> <p><u>Sekundäre Zielparame-ter:</u> <i>Sekundäre Wirksamkeitsendpunkte</i></p> <ul style="list-style-type: none"> <li>- Zeit bis zu andauernder 10%iger Verschlechterung der 6MWD</li> <li>- „Timed function tests“ (TFTs): Zeit, die benötigt wird, um <ul style="list-style-type: none"> <li>• 10 m zu gehen oder zu rennen</li> </ul> </li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>• 4 Stufen zu ersteigen</li> <li>• 4 Stufen herabzusteigen</li> </ul> <p><i>Explorative Wirksamkeitsendpunkte</i></p> <ul style="list-style-type: none"> <li>- Körperliche Funktion: “North Star Ambulatory Assessment” (NSAA) Summenscore</li> <li>- Aktivitäten des täglichen Lebens und Erkrankungs-Status mittels standardisierter Erhebung (PRO)</li> <li>- “Pediatric Outcomes Data Collection Instrument” (PODCI) für die Domänen Transfer/Mobilität und Sport/körperliche Funktion (PRO)</li> </ul> <p><i>Sicherheit und Verträglichkeit</i></p> <ul style="list-style-type: none"> <li>- Größe, Gewicht, Vitalzeichen und körperliche Untersuchungen</li> <li>- UEs klassifiziert nach MedDRA sowie Schweregrad klassifiziert nach CTCAE Version 3.0 wenn möglich; UEs (inklusive Laborauffälligkeiten, Vitalzeichen, Befunde bei den körperlichen Untersuchungen und EKGs) nach Typ, Häufigkeit, Schweregrad, Zeitpunkt und Beziehung zu den Studienpräparaten</li> <li>- Hämatologische Parameter</li> <li>- Serum-Biochemie</li> <li>- Urinanalyse</li> <li>- EKG und Ultraschall der Nieren</li> </ul> <p><i>Ataluren Plasmakonzentrationen</i></p> <p>Probennahme vor den morgendlichen Dosen</p>
<b>6b</b>	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Definition eines zusätzlichen Endpunkts (Zeit, die benötigt wurde, um aus dem Liegen aufzustehen; TFT)
<b>7</b>	Fallzahl	
<b>7a</b>	Wie wurden die Fallzahlen bestimmt?	<p>Der Studie lag die Hypothese zugrunde, dass die mittlere Änderung der 6MWD von BL bis Woche 48 in der Ataluren-Gruppe 30 m größer ist als in der Placebo-Gruppe (jeweils Mittelwert aus zwei validen 6MWTs an Tag 1 und 2 zu BL und in Woche 48). Durch den Einsatz von wiederholten Tests (zu BL und in Woche 48) wurde der Bias, der mit der Verwendung des besten Testergebnisses verbunden ist, umgangen und die natürliche Variabilität des Tests wurde berücksichtigt. In der Studie 007 betrug die SD der Änderung der 6MWD von BL bis Woche 48 72.0 m in der Ataluren-Gruppe. Bei einer 1:1 Randomisierung waren 210 Patienten (105 pro Gruppe) erforderlich, um einen Unterschied</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		von 30 m der 6MWD mit einer Power von 85% zu detektieren ( $\alpha = 0.05$ ). Ausgehend von einer angenommenen Abbruchrate von circa 5% konnten bis zu 220 Patienten (110 pro Gruppe) eingeschlossen werden.
<b>7b</b>	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	Abhängig von der Rekrutierung konnten in Abständen von etwa sechs Monaten verblindete Interim-Analysen zur Sicherheit vorgenommen werden, sofern das DMC nicht anderweitig Analysen wünschte. Falls die verblindeten Interim-Analysen Hinweise auf mögliche Sicherheitsbedenken ergaben, konnte das DMC die Sicherheitsdaten bestimmter oder aller Patienten offenlegen.
<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 Jahre vs. $\geq 9$ Jahre), Kortikosteroid-Begleitmedikation zu BL ( $\geq 12$ Monate vs. <12-6 Monate bezogen auf das Datum der BL-Visite), Gehfähigkeit bestimmt als Mittelwert aus zwei validen BL-6MWTs an Tag 1 und 2 ( $\geq 350$ m vs. <350 m)
<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 gewährleistet. Eine permutierte Blockrandomisierung wurde eingesetzt. Bei Zwillingen erfolgte die Randomisierung für die ganze Familie, wenn der erste Zwilling in die Studie aufgenommen wurde. Nachfolgende Zwillinge der Familie wurden demselben Behandlungsarm zugeteilt wie der erste Zwilling.
<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, erfahrene 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?	Eine Verblindung war durch den Einsatz eines Placebos, welches äußerlich und hinsichtlich des Geschmacks, Geruchs, der Verpackung, der Beschriftung und Anwendungsart nicht unterscheidbar war, gegeben. Nur die CRO, die das IVR/IWR-System erstellt hatte, hatte Zugang zu Randomisierungsinformationen. Eine Entblindung war nur im Falle einer Notfallsituation, auf Verlangen des DMCs

Item <sup>a</sup>	Charakteristikum	Studieninformation
		zum Zeitpunkt der verblindeten Interim-Analysen und nach Beendigung der Studie möglich. Falls das DMC sich entschließen sollte, die Daten während des Sicherheitsreviews zu entblenden, sollten nur die Sicherheitsdaten analysiert werden und nur das DMC hatte Zugang zu diesen Daten. Mit Ausnahme einer Notfallsituation, waren die Patienten, Eltern/Vormund und das Studienpersonal über die randomisierte Zuordnung nicht informiert. Die nicht verblindete Analyse der Daten fand erst statt, wenn der Datensatz komplett war und im Falle einer Entblindung wurde der Patient aus der Studie ausgeschlossen.
<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>Primärer Endpunkt:</u> ANCOVA mit multipler Imputation für fehlende Werte in der ITT-Population unter Berücksichtigung der Stratifizierungsfaktoren Alter (&lt;9 vs. ≥9 Jahre), Dauer der Kortikosteroidaufnahme (circa ≥6-12 vs. ≥12 Monate) und BL-6MWD-Kategorie (≥350 m vs. &lt;350 m) und BL-6MWD als Kovarianten unter Einsatz des MIANALYZE-Verfahrens</p> <p><u>Sekundäre Endpunkte:</u> Wie primärer Endpunkt in der ITT-Population unter Berücksichtigung des jeweiligen BL-Wertes als Kovariante (wenn nicht anders angemerkt)</p>
<b>12b</b>	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p><u>Supportive Analysen des primären Endpunkts:</u></p> <ul style="list-style-type: none"> <li>- MMRM-Analyse unter Berücksichtigung der Stratifizierungsfaktoren (s.o.) und der BL-6MWD, Behandlung, Visite, Interaktion zwischen Behandlung und Visite, zwischen Visite und BL-6MWD und zwischen jedem der Stratifizierungsfaktoren und Visite</li> <li>- Änderung der 6MWD von BL bis zu den einzelnen Visiten (sukzessive jeweils die der finalen zeitlich nächste Visite) vorausgesetzt der Unterschied in Woche 48 ist verglichen mit Placebo signifikant</li> <li>- Randomisierungs-basierte ANCOVA-Analyse (optional)</li> </ul> <p><u>Subgruppen-Analysen:</u> ANCOVA zusätzlich unter Berücksichtigung der Interaktion zwischen Behandlung und Subgruppe:</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation	
		<ul style="list-style-type: none"> <li>- BL-6MWD-Stratifizierungsfaktoren (≥350 m vs. &lt;350 m)</li> <li>- BL-6MWD (&lt;300 m, ≥300 bis &lt;400 m, ≥400 m)</li> <li>- Dauer der vorausgegangenen Kortikosteroid-Therapie zu BL (circa ≥6-12 vs. ≥12 Monate)</li> <li>- Alters-Gruppe zu BL (&lt;9 vs. ≥9 Jahre)</li> </ul> <p><u>Sensitivitätsanalysen des primären Endpunkts:</u></p> <ul style="list-style-type: none"> <li>- Effekt der Zuteilung von Zwillingen in dieselbe Behandlungsgruppe, ANCOVA-Analyse, in die nur die Daten des ersten Zwillings/Drillings eingeschlossen wurden</li> </ul>	
<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) 115, b) 111, c) 114 Ataluren: a) 115, b) 110, c) 114 b) Patienten, die die Studie beendet haben	
<b>13b</b>	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Placebo (n=4): 1 UE 1 verloren („lost to follow-up“) 1 Patientenwunsch 1 Nicht-Compliance	Ataluren (n=5): 1 UE 3 Patientenwunsch 1 Nicht-Compliance
<b>14</b>	Aufnahme / Rekrutierung		
<b>14a</b>	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	26. März 2013 (erste Patientenvisite) bis 20. August 2015 (letzte Patientenvisite)	
<b>14b</b>	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde protokollgemäß beendet.	
<b>a: nach CONSORT 2010.</b>			

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

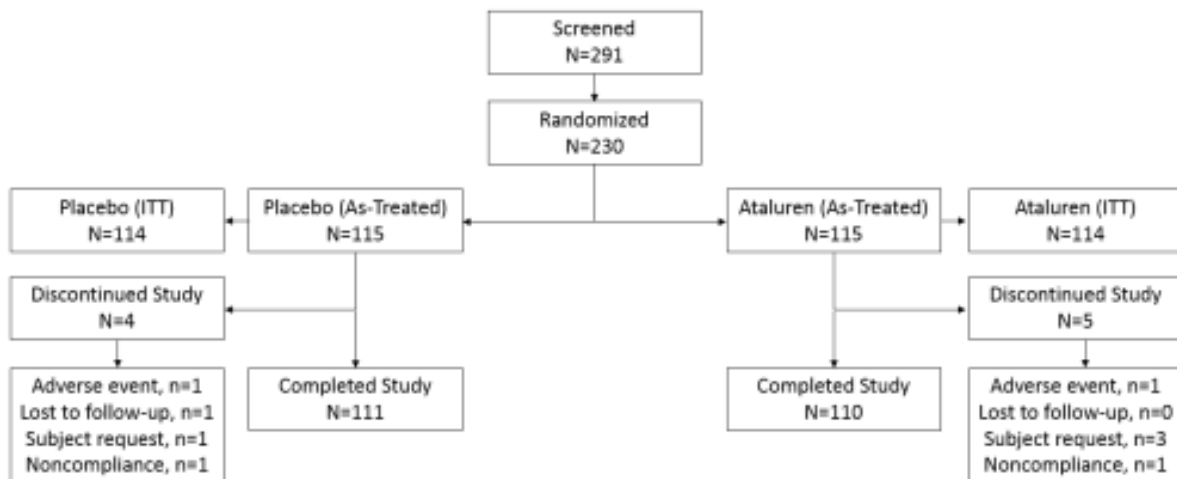


Abbildung 69: Patientenfluss der Studie PTC124-GD-020-DMD (8)



**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 4-86 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PTC124-GD-020-DMD

**Studie: PTC124-GD-020-DMD**

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

Genauere Benennung der Quelle	Kürzel
PTC124-GD-020-DMD Klinischer Studienbericht (8)	A
PTC124-GD-020-DMD Studienprotokoll (37)	B
PTC124-GD-020-DMD Statistischer Analyseplan (32)	C
EMA Richtlinie (2)	D
FDA Richtlinienentwurf (25)	E
IQWiG Allgemeine Methoden – Version 4.2 (7)	F

**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 /IVR unter Einsatz einer permutierten Blockrandomisierung und die Randomisierung ist detailliert im Protokoll und SAP beschrieben (B,C).

**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,C).

---

**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:

---

Es handelte sich um eine doppelblinde Studie (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 Endpunkte wurden wie im Protokoll und SAP festgelegt ausgewertet. Subgruppen-Analysen nach 6MWD spielen für die Bewertung des Ausmaßes des Zusatznutzens von Ataluren eine entscheidende Bedeutung und wurden ausführlich im Studienbericht dargestellt und diskutiert. Die relevanten Subgruppen-Analysen waren im SAP *a priori* definiert (A,B,C). Dieses Vorgehen steht im Einklang mit den Empfehlungen des IQWiGs: „Gibt es a priori Kenntnis von einem möglichen Effektmodifikator (z. B. Alter, Pathologie), so ist es sogar erforderlich, eine mögliche Heterogenität bezüglich des Effekts in den verschiedenen Patientengruppen vorab zu untersuchen“ (F). Auch die FDA und EMA empfehlen Subgruppen-Analysen in klinischen Studien bei DMD (D,E). Daher ergeben sich auch für die Subgruppen-Analysen keine Hinweise auf eine ergebnisgesteuerte Berichterstattung.

---

**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).

---

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

---

Die Erhebung und Auswertung des Endpunkts erfolgte verblindet (A,B,C).

---

**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: 0.8%). Nur ein Patient der randomisierten Patienten pro Gruppe wurde protokollgemäß aus der ITT-Analyse ausgeschlossen. Bei beiden Patienten konnte eine Mutation des Dystrophin-Gens genetisch nicht nachgewiesen werden. Es lagen für beide Patienten keine Post-BL-6MWD-Werte vor, sodass sie nicht der protokollgemäßen Definition der ITT-Population entsprachen. Die Strategie zum Ersatz fehlender Werte war sinnvoll angelegt (A,B,C).

---

**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.

---

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

---

Die Erhebung und Auswertung des Endpunkts erfolgte verblindet (A,B,C).

---

**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: 0.8%) (A,B,C).

---

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

---

Die Erhebung und Auswertung des Endpunkts erfolgte verblindet (A,B,C).

---

**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: 0.8%) (A,B,C).

---

**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.

---

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 zu 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:

---

---