

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Ataluren (Translarna™)

PTC Therapeutics, Limited

Modul 4 A (Ergänzung)

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
Tabellenverzeichnis	4
Abbildungsverzeichnis	7
Abkürzungsverzeichnis	8
4 Modul 4 – allgemeine Informationen	9
4.1 Zusammenfassung der Inhalte von Modul 4.....	10
4.2 Methodik	11
4.2.1 Fragestellung	11
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung	11
4.2.3 Informationsbeschaffung	12
4.2.3.1 Studien des pharmazeutischen Unternehmers	12
4.2.3.2 Bibliografische Literaturrecherche	12
4.2.3.3 Suche in Studienregistern	13
4.2.3.4 Selektion relevanter Studien	14
4.2.4 Bewertung der Aussagekraft der Nachweise.....	14
4.2.5 Informationssynthese und -analyse	15
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien	15
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	16
4.2.5.3 Meta-Analysen.....	17
4.2.5.4 Sensitivitätsanalysen	18
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren	18
4.2.5.6 Indirekte Vergleiche	19
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen	21
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	21
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel	21
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	21
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	22
4.3.1.1.3 Studien aus der Suche in Studienregistern	24
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	25
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	26
4.3.1.2.1 Studiendesign und Studienpopulationen	26
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	38
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	39
4.3.1.3.1 Endpunkte – RCT.....	39
4.3.1.3.1.1 Gehstrecke im 6MWT – RCT	39
4.3.1.3.1.2 Zeit-Funktionstests - RCT	51
4.3.1.3.1.3 Unerwünschte Ereignisse – RCT.....	55
4.3.1.3.2 Subgruppenanalysen – RCT.....	64

4.3.1.3.3	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	66
4.3.2	Weitere Unterlagen.....	68
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	68
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	68
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	69
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	69
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	69
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	72
4.3.2.2	Nicht randomisierte vergleichende Studien	72
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	72
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	73
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	73
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien.....	73
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	75
4.3.2.3	Weitere Untersuchungen.....	75
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	75
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	76
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	76
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	76
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	77
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen.....	77
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	77
4.4.1	Beurteilung der Aussagekraft der Nachweise	77
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	77
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	78
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	79
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	79
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	79
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	79
4.5.4	Verwendung von Surrogatendpunkten	79
4.6	Liste der eingeschlossenen Studien.....	81
4.7	Referenzliste.....	81
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		83
Anhang 4-B : Suchstrategien – Suche in Studienregistern.....		85
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		86
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....		87
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		88
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		91

Tabellenverzeichnis

	Seite
Tabelle 4-1: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	22
Tabelle 4-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	22
Tabelle 4-3: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel	25
Tabelle 4-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	26
Tabelle 4-5: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	27
Tabelle 4-6: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	29
Tabelle 4-7: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	30
Tabelle 4-8: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	38
Tabelle 4-9: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	39
Tabelle 4-10: Operationalisierung von Endpunkt 6MWT (inklusive Progressor-Analyse)	41
Tabelle 4-11: Bewertung des Verzerrungspotenzials für 6MWD (inklusive Progressor-Analyse) in RCT mit dem zu bewertenden Arzneimittel.....	42
Tabelle 4-12: Ergebnisse für Endpunkt 6MWT aus RCT mit dem zu bewertenden Arzneimittel (prä-spezifizierte MMRM-Analyse unter Anwendung einer nicht strukturierten Kovarianzstruktur - ITT)	43
Tabelle 4-13: Prä-spezifizierte Sensitivitätsanalysen der nicht transformierten 6MWD-Ergebnisse	44
Tabelle 4-14: Änderung der mittleren 6MWD über den Studienverlauf (augmentiertes gemischtes Modell mit unstrukturierter Kovarianzstruktur, kITT).....	47
Tabelle 4-15: MMRM-Analyse der Änderung der untransformierten 6MWD und Permutationstest (kITT)	48
Tabelle 4-16: Anteil der Patienten, deren 6MWD sich dauerhaft um mindestens 10% bis Woche 48 verschlechterte	49
Tabelle 4-17: Zeit bis zu einer mindestens 10%igen Verschlechterung der 6MWD in dem Cox-proportionalen Hazard-Modell.....	50
Tabelle 4-18: Operationalisierung von Endpunkt Zeit-Funktionstests („timed function tests“).....	52
Tabelle 4-19: Bewertung des Verzerrungspotenzials für den Endpunkt „TFTs“ in RCT mit dem zu bewertenden Arzneimittel	53
Tabelle 4-20: Änderung der Dauer der TFTs von BL bis Woche 48 (kITT).....	54

Tabelle 4-21: Operationalisierung des Endpunkts „Unerwünschte Ereignisse“ (UEs)	56
Tabelle 4-22: Bewertung des Verzerrungspotenzials für den Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel	57
Tabelle 4-23: Überblick der behandlungsspezifischen UEs aus RCT (Sicherheitspopulation)	58
Tabelle 4-24: Behandlungsspezifische UEs stratifiziert nach Organsystem aus RCT, die bei mehr als 5% der Patienten der Gesamtpopulation über 48 Wochen auftraten (Sicherheitspopulation)	59
Tabelle 4-25: Inzidenz der behandlungsspezifischen mit der Behandlung in Zusammenhang stehenden UEs aus RCT, die bei mehr als einem Patienten auftraten (Sicherheitspopulation)	60
Tabelle 4-26: Inzidenz der behandlungsspezifischen schweren UEs aus RCT (Sicherheitspopulation)	62
Tabelle 4-27: Inzidenz der behandlungsspezifischen schwerwiegenden UEs aus RCT (Sicherheitspopulation)	63
Tabelle 4-28: Effektgrößen der Subgruppen BL-6MWD (<300 m, ≥300 bis <400 m, ≥400 m)	66
Tabelle 4-28: Übersicht der Ergebnisse der patientenrelevanten Endpunkte aus RCT: Ataluren 10/10/20 vs. Placebo	67
Tabelle 4-29: Ergebnisse unter Einschluss der kITT Population und des prädefinierten MMRM Modells	68
Tabelle 4-30: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	69
Tabelle 4-31: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden	70
Tabelle 4-32: Operationalisierung von <Endpunkt xxx>	70
Tabelle 4-33: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	71
Tabelle 4-34: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche	71
Tabelle 4-35: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	73
Tabelle 4-36: Operationalisierung von <Endpunkt xxx>	74
Tabelle 4-37: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	74
Tabelle 4-38: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen	76
Tabelle 4-39: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens	78
Tabelle 4-40 (Anhang): Studiendesign und -methodik für Studie <Studienbezeichnung>	89
Tabelle 4-41 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PTC124-GD-007-DMD	92

Abbildungsverzeichnis

	Seite
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel	24
Abbildung 2: Ergebnisse der 6MWD bei zwei Patienten mit Verletzungen der unteren Extremitäten, die inkorrekt in die ITT-Population eingeschlossen worden waren	45
Abbildung 3: Zeit bis zu einer dauerhaften 10%igen Verschlechterung der 6MWD (kITT) ..	49
Abbildung 4: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>	51

Abkürzungsverzeichnis

Abkürzung	Bedeutung
CONSORT	Consolidated Standards of Reporting Trials
DIMDI	Deutsches Institut für Medizinische Dokumentation
EG	Europäische Gemeinschaft
ITT	Intention to treat
MTC	Mixed Treatment Comparison
RCT	Randomized Controlled Trial
SGB	Sozialgesetzbuch
SOK/SOC	Systemorganklasse / system organ class
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
WHO	World Health Organization

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

<< Angaben des pharmazeutischen Unternehmers >>

Datenquellen

<< Angaben des pharmazeutischen Unternehmers >>

Ein-/Ausschlusskriterien für Studien

<< Angaben des pharmazeutischen Unternehmers >>

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

<< Angaben des pharmazeutischen Unternehmers >>

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

<< Angaben des pharmazeutischen Unternehmers >>

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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), EU Clinical Trials Register (EU-CTR, 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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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)¹. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-² bzw. STROBE-Statements³ 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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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

² 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.

³ 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.

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 Effekten⁴ erfolgen. In begründeten Ausnahmefällen sollen zusätzlich Modelle mit festen Effekten eingesetzt werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁵ 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 Heterogenität⁶ 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.

<< Angaben des pharmazeutischen Unternehmers >>

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

⁵ 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.

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

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.

<< Angaben des pharmazeutischen Unternehmers >>

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-Regressions- 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.

<< Angaben des pharmazeutischen Unternehmers >>

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 Analyseverfahren dar, der Einsatz einfacher adjustierter indirekter Vergleiche ist möglich⁷. 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“⁸, „Multiple-Treatment-Meta-Analysen“⁹ oder auch „Netzwerk-Meta-Analysen“¹⁰, 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 ist¹¹.

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

⁷ 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.

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

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

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

¹¹ B. Schöttker, D. Lüthmann, D. Boulikhemair, and H. Raspe. Indirekte Vergleiche von Therapieverfahren. *Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.*

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 untersuchen¹².

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.*
- *Genaue 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.*

<< Angaben des pharmazeutischen Unternehmers >>

¹² 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.

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 4-1: 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
<Studie 1>	ja	ja	abgeschlossen	12 Monate	Medikament A, Medikament B, Placebo

Geben Sie an, welchen Stand die Information in Tabelle 4-1 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.

<< Angaben des pharmazeutischen Unternehmers >>

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-1 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-2: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie

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.

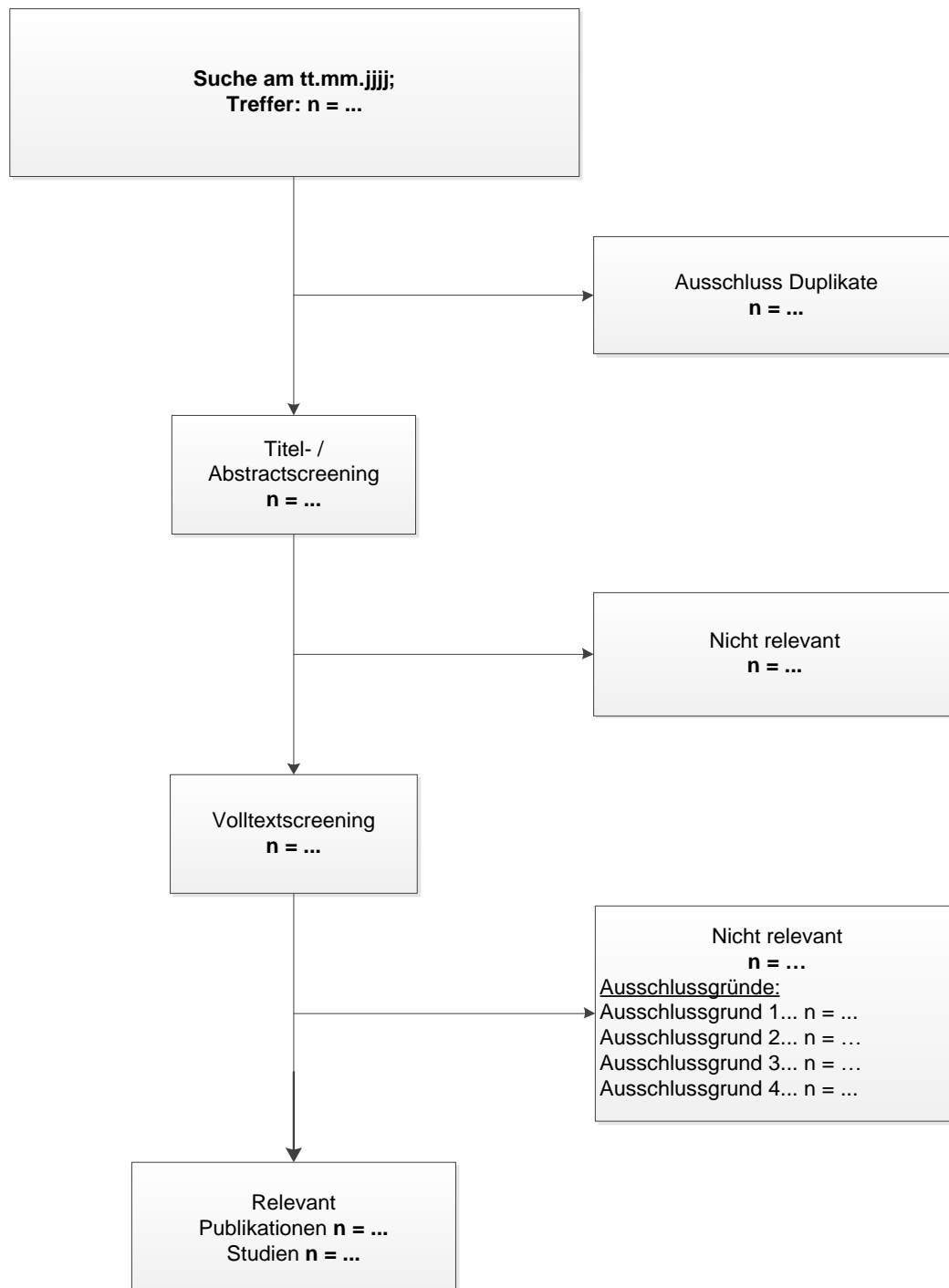


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

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 4-1) 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 4-3: 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 ^a)	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 4-3 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.

<< Angaben des pharmazeutischen Unternehmers >>

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 Zwischenzeilen-

überschriften 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 4-4: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag ^c (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
ggf. Zwischenüberschrift zur Strukturierung des Studienpools						
placebokontrolliert						
<Studie 1>	ja	ja	nein	ja [5]	ja [6, 7]	ja [8]
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
<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 Studie PTC124-GD-007-DMD und die Charakteristika der Interventionen sind in Tabelle 4-5 bzw. Tabelle 4-6 beschrieben.

Tabelle 4-5: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	Population <relevante Charakteristika, z. B. Schweregrad>	Interventionen (Zahl der randomisierten Patienten)	Studiendauer <ggf. Run-in, Behandlung, Nachbeobachtung>	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
PTC124-GD-007-DMD (9)	Phase 2b, internationale, multi-zentrische, randomisierte, doppelblinde, Placebo-kontrollierte Dosisfindungsstudie	Gehfähige männliche Patienten mit nmDMD im Alter von ≥ 5 Jahren; Fähigkeit, während des Screening-6MWTs ≥ 75 m ohne Hilfen zu gehen	Placebo (n=57) Ataluren 10/10/20 mg/d (n=57) Ataluren 20/20/40 mg/d (n=60)	Screening-Phase mit 2 Visiten Behandlung: 48 Wochen Nachbeobachtung: Follow-up-Visite 6 Wochen (± 7 d) nach der letzten Dosis oder Teilnahme an einer offenen Verlängerungsstudie; bei mit der Behandlung in Verbindung stehenden/schwerwiegenden UEs Nachbeobachtung bis UEs abgeklungen oder stabil Patienten, die ≥ 1 Dosis Ataluren erhalten hatten: Follow-up über 5 Jahre geplant (Daten noch nicht verfügbar)	International (Australien, Belgien, Deutschland, Frankreich, Israel, Italien, Kanada, Spanien, Schweden, UK, USA) 28. Februar 2008 - 17. Dezember 2009	<u>Primärer Endpunkt:</u> 6MWD; <u>Patienten-relevante sekundäre Endpunkte:</u> Körperliche Funktion: TFTs (proximale Muskelfunktion), Muskelkraft (Myometrie), Aktivität im Alltag (SAM); Durch die Patienten/Betreuer berichtete Endpunkte: Gesundheitsbezogene Lebensqualität (PedsQL), Zufriedenheit mit der Behandlung (TSQM), akzidentielle Stürze pro Tag (Tagebuch) Kognitive Funktion: „Digit span task“ Kardiale Funktion: Herzrate

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

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

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

Studie	Placebo	Ataluren Niedrig- Dosis	Ataluren Hoch-Dosis	Komedikationen
PTC124-GD-007-DMD (9)	Dreimal täglich, morgens, mittags und abends p.o. etwa 30 min nach den Mahlzeiten	10, 10, 20 mg/kg KG/d, analog Placebo	20, 20, 40 mg/kg KG/d, analog Placebo	Bei mit Kortikosteroiden (etwa 70% in allen Gruppen) oder kardialen Medikationen (15.8% Placebo, etwa 10% Ataluren) vorbehandelten Patienten sollte die Dosis während der Studie wenn möglich stabil bleiben. Nur minimale Änderungen dieser Medikationen wurden während der Studie dokumentiert.

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

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

Quelle (Studienreport PTC124-GD-007-DMD): S. 81, 82 (9)

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

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

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

Studie PTC124-GD-007-DMD (9)

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

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

Studienpopulation: Männliche Patienten im Alter von ≥ 5 Jahren mit phänotypischer Evidenz einer DMD und Dokumentation einer Nonsense-Mutation im Dystrophin-Gen via Gensequenzierung durch ein zertifiziertes Labor wurden in die Studie eingeschlossen. Nur Patienten, die in der Lage waren während der Screening-Visite ≥ 75 m ohne Hilfen zu gehen, konnten an der Studie teilnehmen.

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

Vergleichstherapie: Zum Zeitpunkt der Studienplanung stand keine medikamentöse Standardtherapie zur Behandlung der DMD zur Verfügung, sodass die Wahl von Placebo plus SOC als Vergleichstherapie angemessen war. Das CHMP stimmte der Wahl von Placebo als Referenztherapie zu, da Ataluren das erste Medikament zur Behandlung der DMD ist („first-in-class“) und keine zugelassene Standardtherapie existiert (1).

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

Änderung der 6MWD von BL bis Woche 48.

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

Sekundäre Endpunkte:

Körperliche Funktion:

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

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

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

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

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

Pharmakodynamik:

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

Plasma Ataluren-Konzentrationen

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

Fallzahlschätzung: Basierend auf Studien bei Patienten mit anderen chronischen

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

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

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

Studienpopulationen

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

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

Statistische Analyse inklusive unterstützende und Sensitivitätsanalysen:

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

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

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

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

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

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

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

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

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

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

Post-hoc-Analysen und Änderungen des SAP

Die MMRM-Methode diente als primäre Analyse der 6MWD und die ANCOVA mit LOCF-Imputation als unterstützende Analyse. Der Vergleich von Ataluren 10/10/20 und Placebo zeigte einen substantiellen Unterschied zwischen der prä-spezifizierten MMRM-Analyse der nicht transformierten 6MWD ($p=0.0905$) verglichen mit dem p-Wert der prä-spezifizierten ANCOVA-Analyse ($p=0.0445$). Da nur wenige Werte in Woche 48 fehlten, deutete diese Diskrepanz an, dass das protokollgemäße MMRM-Modell möglicherweise ungeeignet sein könnte. Nach Konsultation des Biostatistikers Prof. Gary Koch (Universität Nord-Carolina) wurde vermutet, dass ein BL-Visite Interaktionsterm in dem MMRM-Modell fehlt, da die Beziehung zwischen BL 6MWD und post-BL 6MWD mit der Zeit variieren kann. Die BL-Visite Interaktion war hoch signifikant ($p<0.001$), was andeutet, dass der Effekt der BL 6MWD über die Zeit nicht konstant ist. Die Diskrepanz zwischen der MMRM- und ANCOVA-Analyse verschwand, wenn der zusätzliche Interaktionsterm eingeschlossen wurde.

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

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

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

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

Da die kITT-Analyse wissenschaftlich und klinisch begründet ist und die tatsächlichen Fähigkeiten der Patienten widerspiegelt, bezieht sich die vorliegende Nutzenbewertung schwerpunktmäßig auf diese Analysen. Die Ergebnisse in der ITT-Population (mit den beiden invaliden 6MWD-Werten) wurden aus informativen Gründen dargestellt, wenn dies als sinnvoll erachtet wurde. Das CHMP (2014) kommentierte die kITT-Analyse dahingehend, dass die Analyse methodisch gesehen als akzeptabel eingestuft werden könne, es sich aber *de facto* um post-hoc-Ergebnisse handle. Die Ergebnisse der kITT-Analyse flossen in die Bewertung von Ataluren durch das CHMP (2014) ein (1).

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 4-8: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential wird auf Studienebene unter Berücksichtigung der in Tabelle 4-8 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Bei der Studie PTC124-GD-007-DMD handelt es sich um eine multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte internationale Parallelgruppenstudien der Phase 2b. Eine adäquate Erzeugung der Randomisierungssequenz liegt vor. Eine adaptive Randomisierung unter Einsatz akzeptierter Minimierungstechniken und Permutationsalgorithmen wurde angewandt, um mögliche Ungleichheiten zwischen den Gruppen hinsichtlich der prä-definierten Stratifikationskriterien und der Zahl der Patienten pro Arm zu vermeiden bzw. deren Risiko zu minimieren. Die Gruppenzuteilung erfolgte verdeckt. Sowohl die Patienten als auch die Prüfarzte und das medizinische Personal waren bis zur Endauswertung der Studie verblindet. Lediglich ein Patient (Ataluren 20/20/40) brach die Studie etwa in Woche 6 vorzeitig ab. Die BL- und Woche 6 Daten dieses Patienten wurden in allen MMRM-Analysen berücksichtigt. Das ITT-Prinzip wurde adäquat umgesetzt. In die Analyse der kITT-Population wurden wie im Protokoll vorgesehen nur Patienten mit validen 6MWD-Werten eingeschlossen, so dass die verletzungsbedingt unrealistisch niedrigen BL-Werte von zwei Patienten durch deren Screening-Werte ersetzt wurden. Diese Analyse wird trotz ihres post-hoc-Charakters als klinisch und wissenschaftlich begründet eingestuft (siehe 4.3.1.2.1), was sich mit der Einschätzung des CHMPs in dem Beurteilungsbericht von Ataluren deckt (1). Wie bei einer sehr seltenen Erkrankung zu erwarten, ist die Fallzahl relativ niedrig. Unter Berücksichtigung der Seltenheit der Erkrankung ist die Zahl der eingeschlossenen Patienten allerdings als akzeptabel einzuschätzen. Es liegen keine Hinweise auf eine ergebnisgesteuerte

Berichterstattung oder sonstige Aspekte, die das Verzerrungspotential beeinflussen könnten, vor.

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

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

Tabelle 4-9: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

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

4.3.1.3.1 Endpunkte – RCT

4.3.1.3.1.1 Gehstrecke im 6MWT – RCT

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

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

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler

– Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde

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

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

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

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

Tabelle 4-10: Operationalisierung von Endpunkt 6MWT (inklusive Progressor-Analyse)

Studie	Operationalisierung
PTC124-GD-007-DMD	<p>Das Gehvermögen der Patienten wurde im 6MWT nach einer modifizierten Version der ATS-Richtlinie (47) mittels eines an der Universität Kalifornien-Davis entwickelten, standardisierten Verfahrens erfasst (48). Den Prüfbüros wurde ein Manual zur Verfügung gestellt, in dem der Test detailliert beschrieben war. Modifikationen gegenüber der ATS-Version schlossen ein kurzes Video zur Orientierung vor dem Test, eine kontinuierliche Motivation durch das Klinikpersonal und eine Person zur Sicherung, die dem Patienten während des Tests folgte, ein. Die Tests sollten in einem ausreichend breiten Flur mit harter Oberfläche, der mit einem 25 m langen Maßband markiert war, stattfinden. Der Start- bzw. Wendepunkt sollte jeweils mit Kegeln markiert werden. Die einzelnen Testschritte inklusive der Überwachung des Pulses vor, während und nach dem Test waren in dem Manual detailliert beschrieben und die Phrasen zur Motivation der Patienten waren vorgegeben, um möglichst standardisierte Testbedingungen zu schaffen. Die benötigte Zeit oder gegebenenfalls Zeit bis zu einem Sturz wurde mittels digitaler Stoppuhr erfasst. War der Patient nach einem Sturz während des Tests unverletzt und war der Patient dazu in der Lage, konnte der Test fortgeführt werden. War dies nicht der Fall wurde die Gehstrecke zum Zeitpunkt des Sturzes als 6MWD erfasst. Benötigte der Patient eine Ruhepause konnte er sich an der Wand anlehnen und wurde angehalten, den Test so bald wie möglich fortzusetzen.</p> <p>Vor Studienbeginn nahmen die Zentren an einem Endpunkttraining und einer Standardisierungssitzung teil, um die Testprotokolle und Logistik zu harmonisieren. Nach etwa einem Jahr fand ein zentrales erneutes Training statt.</p> <p>Wie im SAP festgelegt wurde der 6MWT als valide eingestuft, wenn der Patient den Test wie geplant abschloss oder den Test ausschließlich aufgrund von Erschöpfung vorzeitig abbrach. Unter folgenden Bedingungen galt der 6MWT als nicht valide: Der Patient</p> <ul style="list-style-type: none"> - folgte nicht den Anweisungen, - behielt den Kurs während der Dauer des Tests nicht ein, - ging zu irgendeinem Zeitpunkt während des Tests in eine entgegengesetzte Richtung und/oder - brach den Test aufgrund fehlender Compliance oder anderer Gründe vorzeitig ab (<6 min), die nicht mit Erschöpfung in Zusammenhang standen (z.B. Verletzung durch Sturz, Hinsetzen). <p>Die Zentren wurden angehalten im Falle eines nicht validen initialen Tests oder für den Fall, dass der Patient zu erschöpft war, um den ersten Test zu beenden zu versuchen, den Test zu wiederholen (nach einer mindestens 1-stündigen Ruhepause an Tag 1 oder an Tag 2). Bei gehfähigen Patienten ohne validen Test, sollte die 6MWD als fehlend eingegeben werden.</p> <p>Die 6MWD wurde zur Screening-Untersuchung, Tag 1 (Visite 3 vor Beginn der Behandlung), Woche 6 (Visite 4), Woche 12 (Visite 5), Woche 18 (Visite 6), Woche 24 (Visite 7), Woche 30 (Visite 8), Woche 36 (Visite 9), Woche 42 (Visite 10) und Woche 48 (Visite 11 oder Ende der Behandlung) mittels 6MWT erfasst. An Tag 1 sowie in Woche 24 und 48 sollten die 6MWTs mittels Videokamera dokumentiert werden sofern eine Einverständniserklärung der Patienten/Betreuer zur Aufzeichnung des Tests vorlag, die keinen Einfluss auf die weitere Teilnahme an der Studie hatte.</p> <p>Wie in Kapitel 4.3.1.2.1 beschrieben war die geplante primäre Analyse die Änderung der 6MWD von BL bis Woche 48 für den Vergleich der Ataluren-Gruppen vs. Placebo in der ITT-Population (mindestens eine der Ataluren-Gruppen ist Placebo signifikant überlegen). Post-hoc wurde eine so genannte kITT-Population definiert und ausgewertet, die in Übereinstimmung mit dem CHMP (2014) als methodisch begründet eingeschätzt wird (1).</p> <p>Zusätzlich waren Responder- und Progressor-Analysen und Analysen der Zeit bis zum Therapieeffekt bzw. einer Krankheitsprogression (10%ige Verbesserung oder Verschlechterung der 6MWD) im SAP spezifiziert (siehe Kapitel 4.3.1.2.1).</p> <p>Der 6MWT ist ein mittlerweile etablierter und validierter primärer Endpunkt in Studien bei DMD-Patienten, der als Maß der Krankheitsprogression angesehen (1).</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 4-11: Bewertung des Verzerrungspotenzials für 6MWD (inklusive Progressor-Analyse) in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential für den Endpunkt 6MWD wird unter Berücksichtigung der in Tabelle 4-11 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Das Verzerrungspotential auf Studienebene wurde ebenfalls als gering eingestuft. Da es sich um eine randomisierte, kontrollierte Doppelblindstudie handelte, lag eine Verblindung der Endpunkterheber vor. Die Prüfarzte, das medizinische Personal und die Patienten waren verblindet. Es fand eine adäquate Operationalisierung des Endpunkts statt. Der 6MWT wurde unter standardisierten Bedingungen nach einem Trainingsprogramm in den Studienzentren und unter Einsatz eines Manuals, in dem die Testbedingungen inklusive Angaben zur Validität eines Tests detailliert beschrieben waren, durchgeführt. Der 6MWT ist ein bei DMD-Patienten validierter und mittlerweile akzeptierter primärer Endpunkt in DMD-Studien. Auch die Wahl des Stratifizierungskriteriums (<350 m vs. \geq 350 m) und des Kriteriums zur Definition einer Krankheitsprogression (mindestens 10%ige anhaltende Verschlechterung der 6MWD) waren klinisch begründet. Eine adäquate Umsetzung des ITT Prinzips fand statt. Alle Patienten, die mindestens eine Dosis der Prüfpräparate erhalten hatten, wurden in die ITT-Analyse eingeschlossen. Die Zahl der randomisierten Patienten entsprach der ITT-Population. Wie oben angesprochen wurden in der kITT-Analyse die invaliden BL-Daten von zwei Patienten, die kurz zuvor akute Verletzungen der unteren Extremitäten erlitten hatten und inkorrekt Weise in die ITT-Population eingeschlossen worden waren, durch deren Screening-Werte ersetzt. Die unrealistisch niedrigen BL-Werte spiegelten nicht die tatsächlichen Fähigkeiten der beiden Patienten wider, so dass eine Verzerrung der protokollgemäßen Primäranalyse der

6MWD nicht angenommen werden muss. Dementsprechend wird die kITT-Analyse in der vorliegenden Bewertung als die klinisch und wissenschaftlich begründete (primäre) Analyse eingestuft und die geplante ITT-Primäranalyse als unterstützende Auswertung. Diese Vorgehensweise wird durch das CHMP (2014) bestätigt (1). Die Robustheit der Primäranalyse wurde durch verschiedene Sensitivitätsanalysen zusätzlich bestätigt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder das Vorhandensein von sonstigen Faktoren, die das Verzerrungspotential beeinflussen könnten, liegen nicht vor.

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

Tabelle 4-12: Ergebnisse für Endpunkt 6MWT aus RCT mit dem zu bewertenden Arzneimittel (prä-spezifizierte MMRM-Analyse unter Anwendung einer nicht strukturierten Kovarianzstruktur - ITT)

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

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

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

¹ Nicht transformierte Werte

² Rang-transformierte Werte

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

In der MMRM-Analyse (ITT) betrug die mittlere Änderung der nicht transformierten 6MWD von BL bis Woche 48 26.4 m (nominal p=0.0905) in der Ataluren 10/10/20-Gruppe und -0.1 m (nominal p=0.9956) in der Ataluren 20/20/40-Gruppe verglichen mit Placebo. Basierend auf den Rang-transformierten 6MWD-Daten, die aufgrund des Abweichens von der Normalverteilung als primär anzusehen gewesen wären, betragen die nominalen p-Werte der mittleren Änderung bis Woche 48 p=0.1490 bzw. p=0.4756 für Ataluren 10/10/20 bzw. 20/20/40 vs. Placebo. Auch der Dunnetts-Test zeigte keine signifikanten Unterschiede zwischen einer der Ataluren-Gruppen und Placebo (Tabelle 4-12). Dementsprechend wurde

das primäre Ziel der Studie (mindestens einer der Atalurenarme ist dem Placeboarm hinsichtlich des primären Endpunkts signifikant überlegen) formell nicht erreicht.

Da nur wenige 6MWD-Werte fehlten, ergab die ITT-Analyse mit LOCF-Imputation vergleichbare Ergebnisse. Der Unterschied der Veränderungen (Δ LSM \pm SD) in der 6MWD von BL bis Woche 48 betrug 28.36 m (14.0) zwischen der in der Placebo-Gruppe und der Ataluren 10/10/20- Gruppe (ITT-LOCF). Wie in Tabelle 4-13 dargestellt entsprachen die Sensitivitäts- und supportiven Analysen der nicht transformierten 6MWT-Ergebnisse in ihrer Ausrichtung der Primäranalyse und lieferten keine Hinweise auf eine systematische Verzerrung. Hinsichtlich der Gehstrecke im 6MWT bis Woche 48 zeigte sich übereinstimmend ein Vorteil von Ataluren 10/10/20 nicht aber von Ataluren 20/20/40 gegenüber Placebo. Ein signifikanter Unterschied zwischen Ataluren 10/10/20 und Placebo wurde lediglich in der ANCOVA-LOCF-Analyse beobachtet (nominal $p=0.0445$). Wie bereits oben angesprochen (Kapitel 4.3.1.2.1) war der Unterschied der p-Werte zwischen der ANCOVA-LOCF- und der MMRM-Analyse angesichts nur weniger fehlender Werte überraschend und führte zu einer post-hoc-Modifikation des MMRM-Modells (zusätzlicher Interaktionsterm BL 6MWD-Visite). Der zusätzliche Interaktionsterm erlaubte eine flexiblere BL-Adjustierung zu jeder Visite, da der Zusammenhang zwischen der BL 6MWD und der post-BL 6MWD über die Zeit variieren kann.

Tabelle 4-13: Prä-spezifizierte Sensitivitätsanalysen der nicht transformierten 6MWD-Ergebnisse

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

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

¹ Dunnetts-Test

² Faktoren: Behandlung, Alter und Kortikosteroid-Anwendung, BL 6MWD als Kovariate

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

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

kITT Analyse und Post-hoc-Analysen des 6MWT

Wie in Kapitel 4.3.1.2.1 dargelegt, wurden die BL 6MWD-Werte bei zwei Patienten, die kurz zuvor Verletzungen der unteren Extremitäten erlitten hatten, inkorrekt Weise durch den klinischen Beurteiler in einem Studienzentrum als valide eingestuft und dementsprechend in die ITT-Analyse eingeschlossen. Die Verletzungen der unteren Extremitäten waren bei den beiden Patienten als UE dokumentiert worden. Lediglich Patienten mit validen BL 6MWD-Werten und mindestens einem validen post-BL-Wert durften per Protokoll in die ITT-Population eingeschlossen werden. Der Vergleich der Gehstrecke zur Screening-Visite und der post-BL-Werte zeigte deutlich, dass der BL-Wert bei beiden Patienten nicht das tatsächliche Gehvermögen widerspiegelte und die niedrigen BL-Werte auf die Verletzungen der unteren Extremitäten zurückzuführen waren (Abbildung 2). Die Analysen wurden nach Ersatz der invaliden BL- durch die Screening-Werte dieser beiden Patienten in der kITT-Population wiederholt. Wie oben angesprochen wurde in der primären kITT-Analyse der 6MWD ein zusätzlicher Interaktionsterm (BL-Visite) in dem MMRM-Modell berücksichtigt (post-hoc). Die MMRM-Analyse basierte primär auf den nicht transformierten Daten und nicht den Rang-transformierten Daten wie in der Primäranalyse.

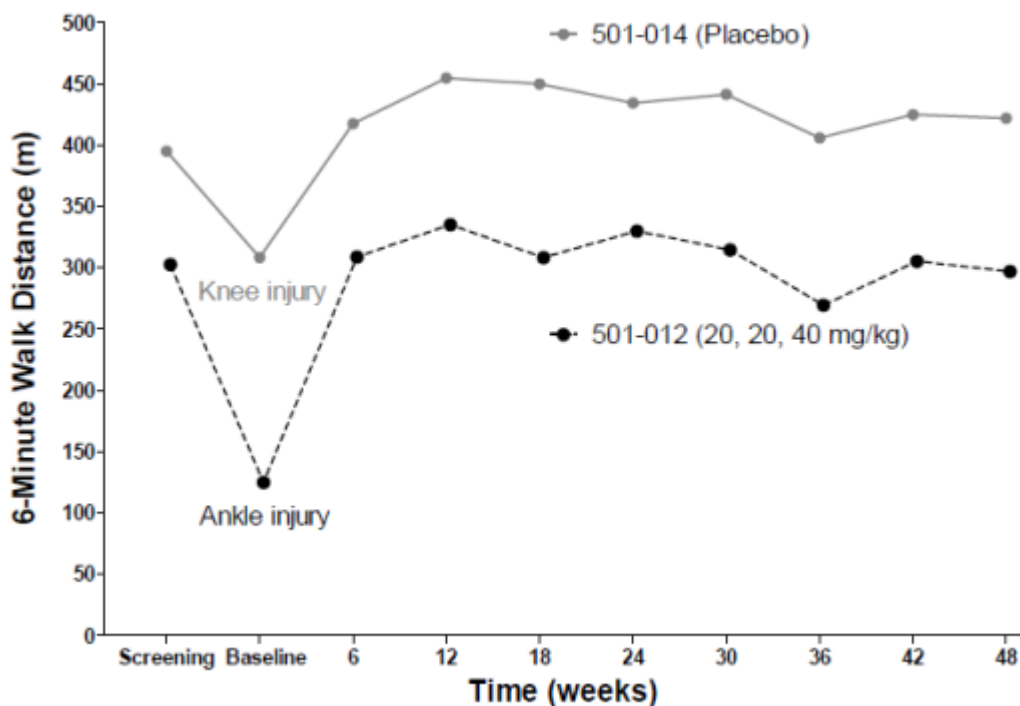


Abbildung 2: Ergebnisse der 6MWD bei zwei Patienten mit Verletzungen der unteren Extremitäten, die inkorrekt Weise in die ITT-Population eingeschlossen worden waren

Die Ergebnisse der 6MWD-Analysen in der kITT-Population und verschiedener Sensitivitätsanalysen sind in Tabelle 4-14 dargestellt.

Tabelle 4-14: Änderung der mittleren 6MWD über den Studienverlauf (augmentiertes gemischtes Modell mit unstrukturierter Kovarianzstruktur, kITT)

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

Quelle (Studienbericht): S. 555-557

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

¹ p-Wert Ataluren vs. Placebo² Dunnetts-Test

Der Unterschied zwischen Ataluren 10/10/20 und Placebo betrug unter Einschluss der kITT Population und unter Verwendung des augmentierten MMRM Modells 31.7 m (95% KI: 5.13, 58.28).

Auch der Permutationstest zur Überprüfung der Robustheit der MMRM-kITT-Analyse zeigte einen signifikanten Unterschied zugunsten von Ataluren 10/10/20 vs. Placebo ($p=0.0281$ nominal; Tabelle 4-15). Der Unterschied zwischen Ataluren 10/10/20 und Placebo zeigte sich bereits früh im Studienverlauf (etwa Woche 24-36) und blieb bis Woche 48 erhalten.

Tabelle 4-15: MMRM-Analyse der Änderung der untransformierten 6MWD und Permutationstest (kITT)

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

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

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

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

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

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

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

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

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

Innerhalb von 48 Wochen verschlechterte sich die Gehfähigkeit bei 44% der Patienten der Placebo-Gruppe dauerhaft um mindestens 10% verglichen mit 26% in der Ataluren 10/10/20-Gruppe (kITT, nominal $p=0.0386$, adjustiert 0.0652; Abbildung 3). Der Anteil der Patienten, bei denen die Erkrankung bis Woche 48 fortschritt, unterschied sich nicht zwischen der ITT- und kITT-Analyse. Die OR (95% KI; kITT) betrug 0.3 (0.1, 0.9) für Ataluren 10/10/20 und 1.1 (0.5, 2.5) für Ataluren 20/20/40 jeweils verglichen mit Placebo und der CMH-Test zeigte einen signifikanten Unterschied ($p=0.0423$) (Tabelle 4-16).

Tabelle 4-16: Anteil der Patienten, deren 6MWD sich dauerhaft um mindestens 10% bis Woche 48 verschlechterte

Gruppe	N	n (%)	OR (Ataluren vs. Placebo)	95% KI	p-Wert
<i>ITT</i>					
Placebo	57	25 (43.9%)	-	-	-
Ataluren 10/10/20	57	15 (26.3%)	0.4	0.1, 0.9	0.0423 ¹
Ataluren 20/20/40	60	29 (48.3%)	1.1	0.5, 2.6	
<i>kITT</i>					
Placebo	57	25 (43.9%)	-	-	-
Ataluren 10/10/20	57	15 (26.3%)	0.3	0.1, 0.9	0.0423¹
Ataluren 20/20/40	60	29 (48.3%)	1.1	0.5, 2.5	

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

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

¹ CMH-Test

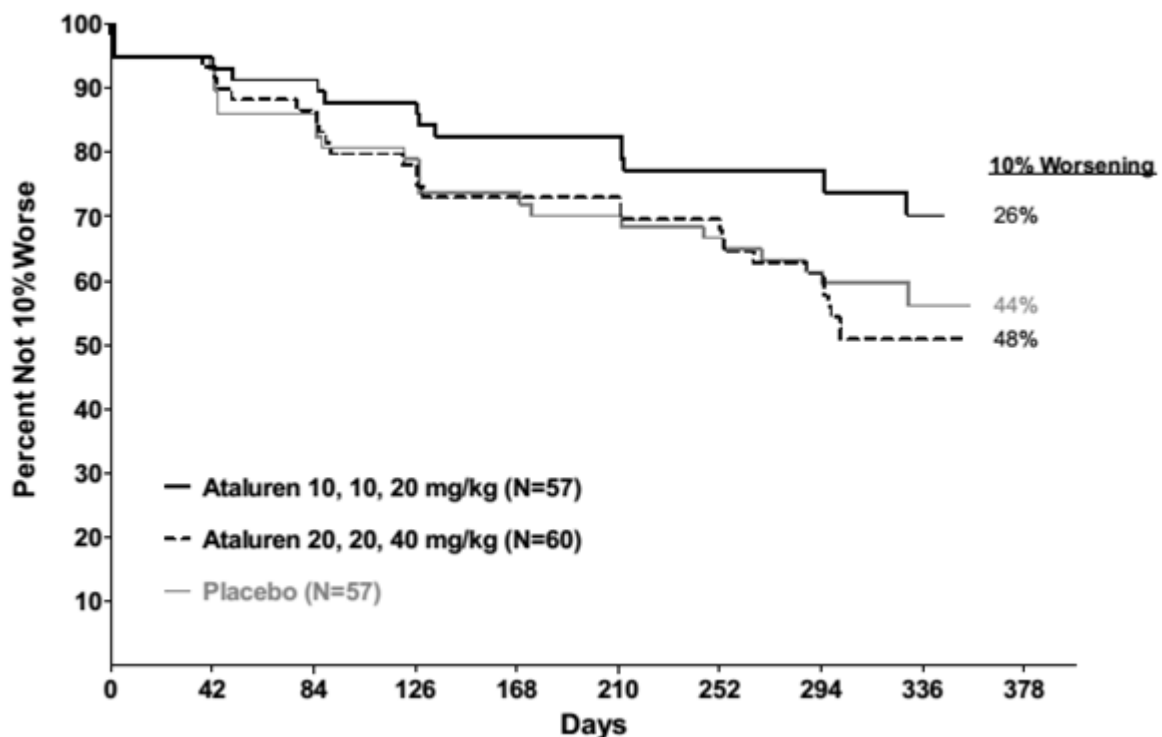


Abbildung 3: Zeit bis zu einer dauerhaften 10%igen Verschlechterung der 6MWD (kITT)

Quelle (Studienbericht): S. 102

Die Zeit bis zu einer dauerhaften mindestens 10%igen Verschlechterung der 6MWD von BL bis Woche 48 wurde in einem Cox-proportionalen Hazard-Modell analysiert. Die HR (95% KI; kITT) der Zeit bis zu einer mindestens 10%igen Verschlechterung der 6MWD in Woche 48 für den Vergleich von Placebo und Ataluren 10/10/20 betrug 0.508 (0.273; 0.946; nominal

p=0.0326, adjustiert p=0.0652; Tabelle 4-17), was einer 48%igen Reduktion des Hazards unter Behandlung mit Ataluren 10/10/20 entsprach. Zwischen der Ataluren 20/20/40- und der Placebo-Gruppe ergab sich auch in dieser Auswertung kein Unterschied (kITT, HR 1.126, 95% KI 0.656; 1.933, nominal p=0.6667), was eine inverse Dosis-Effekt-Beziehung nahelegt. Die Effekte waren in der ITT-Analyse und einer Analyse unter Ausschluss der beiden Patienten mit invaliden BL 6MWD-Werten vergleichbar. In der ITT-Analyse betrug die HR für den Vergleich von Ataluren 10/10/20 vs. Placebo 0.520 (95% KI 0.280-0.966; nominal p=0.0386, adjustiert p=0.078), was einer 48%igen Reduktion des Risikos einer dauerhaften mindestens 10%igen Verschlechterung der Gehstrecke im 6MWT entsprach. Wie aus Abbildung 3 ersichtlich war der Unterschied zwischen Ataluren 10/10/20 und Placebo bereits zu einem frühen Zeitpunkt erkennbar und blieb über den gesamten Studienverlauf erhalten.

Tabelle 4-17: Zeit bis zu einer mindestens 10%igen Verschlechterung der 6MWD in dem Cox-proportionalen Hazard-Modell

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

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

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

Hinsichtlich des Anteils der Patienten, die auf die Therapie ansprachen (dauerhafte 10%ige Verbesserung der 6MWD bis Woche 48), ergab sich im Gruppenvergleich kein signifikanter Unterschied (kITT, 21% bzw. 17% Ataluren 10/10/20 bzw. 20/20/40, 9% Placebo), da die verglichenen Populationen sehr klein waren. In diesem Zusammenhang sollte betont werden, dass eine Restauration der Dystrophin-Expression durch die Behandlung mit Ataluren darauf abzielt, das noch funktionelle, gesunde Muskelgewebe vor mechanischer Belastung zu schützen. Die Wiederherstellung der Funktion eines durch die Erkrankung bereits geschädigtes Muskelgewebes ist durch eine Ataluren-Therapie auf Grund des Wirkmechanismus‘ nicht zu erwarten.

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

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

<Abbildung Meta-Analyse>

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

<< Angaben des pharmazeutischen Unternehmers >>

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 4-18: Operationalisierung von Endpunkt Zeit-Funktionstests („timed function tests“)

Studie	Operationalisierung
PTC124-GD-007-DMD	<p>Während der TFTs wurde die Zeit, die benötigt wurde, um aus der Rückenlage aufzustehen, 10 m zu gehen oder zu rennen und vier Treppenstufen von Standardgröße hinauf- oder hinabzusteigen, erfasst. Die Patienten hatten für die Tests maximal 30 sec Zeit und bei Patienten, die aufgrund von Krankheitsprogression nicht in der Lage waren, an den Tests teilzunehmen oder die Tests innerhalb von 30 sec abzuschließen, wurden ab diesem Zeitpunkt 30 sec dokumentiert. Bei Patienten, bei denen die Tests aus anderen Gründen fehlten (z.B. Knochenbrüche), sollten die Werte als fehlend eingegeben werden.</p> <p>Die Tests waren in einem Manual, das den Prüfbüros zur Verfügung gestellt wurde, detailliert beschrieben. Die Tests sollten pro Visite vorzugsweise nur einmal ausgeführt werden, ausgenommen im Falle eines Testfehlers oder der erste Versuch war nicht auswertbar. Die Tests sollten in folgender Reihenfolge stattfinden: 1) 10 m Gehen oder Rennen, 2) Aufstehen aus liegender Position, 3) 4 Treppenstufen hinauf und hinabsteigen. Ein gerader Flur mit einer Länge von mindestens 12 m, der mit einer Start- und Ziellinie (exakt 10 m) markiert war, musste zur Verfügung stehen. Der Zeitpunkt, zu dem der zweite Fuß die Ziellinie kreuzte sollte mittels Stoppuhr gemessen werden. Bei Patienten, deren Zustand sich während der Studie derart verschlechterte, dass sie lange Beinschienen oder die Unterstützung durch eine andere Person während des 10-m-Gehtests benötigten, sollte die Zeit nicht dokumentiert werden. Die Fähigkeiten in den TFTs wurden zusätzlich mittels Skalen bewertet.</p> <p>Die TFTs wurden in einem Random-Effekt-Modell für Ordinalzahlen in SAS (PROC NL-MIXED) analysiert.</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 4-19: Bewertung des Verzerrungspotenzials für den Endpunkt „TFTs“ in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential für den Endpunkt TFTs wird unter Berücksichtigung der in Tabelle 4-19 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Das Verzerrungspotential wurde auf Studienebene als niedrig eingestuft, da die Endpunkte verblindet erhoben wurden, das ITT-Prinzip adäquat umgesetzt wurde, und keine Hinweise auf eine ergebnisgesteuerte Berichterstattung vorliegen. Auch für die TFTs wird die kITT-Analyse als die primär relevante und klinisch sowie wissenschaftlich begründete Analyse eingestuft. Die Endpunkte wurden adäquat operationalisiert, die Endpunkterheber wurden trainiert und ihnen wurde ein ausführliches Manual, in dem die Tests beschrieben waren, zur Verfügung gestellt. TFTs sind ein in DMD-Studien häufig eingesetztes und etabliertes Instrument.

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

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

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

Die Zeit, um vier Treppenstufen zu ersteigen, nahm bis Woche 48 in der Ataluren 10/10/20-Gruppe deutlich weniger stark zu als in der Placebo-Gruppe (-2.57 sec. vs. Placebo), und der Unterschied zwischen der Ataluren 10/10/20- und der Placebo-Gruppe war signifikant ($p=0.0252$ nominal, $p=0.0467$ adjustiert). Insgesamt unterstützen die Ergebnisse dieses TFTs die Primäranalyse des 6MWT. Die anderen TFTs (Zeit um vier Treppenstufen hinabzugehen,

Zeit um 10 m zu gehen oder zu rennen, Zeit bis zum Aufstehen aus der Rückenlage) zeigten positive Trends zugunsten von Ataluren 10/10/20 verglichen mit Placebo (Tabelle 4-20).

Tabelle 4-20: Änderung der Dauer der TFTs von BL bis Woche 48 (kITT)

Gruppe	n BL/ Woche 48	M (±SD) BL (sec)	M (±SD) Woche 48 (sec)	Δ BL bis Woche 48, M (±SD) (sec) p-Wert	LSM (± SE) Woche 48 (sec)	Δ vs. Placebo, LSM (±SE) [95% KI] (sec)	p-Wert vs. Placebo Nominal Dunnetts
<i>Zeit bis zum Aufstehen aus Rückenlage</i>							
Placebo	57/56	11.50 (11.437)	14.60 (12.327)	3.24 (7.254)	14.85 (0.83)	-	-
ATL 10/10/ 20	57/57	10.80 (9.924)	14.03 (11.408)	3.23 (5.761) p=1.0000 ¹	14.79 (0.83)	-0.06 (1.16) [-2.35, 2.23]	0.9580 0.9980
ATL 20/20/ 40	60/59	12.25 (11.191)	15.39 (11.823)	3.00 (5.686) p= 0.9696 ¹	14.71 (0.81)	-0.15 (1.15) [-2.41, 2.12]	0.8985 0.9881
<i>Zeit, um 10 m zu gehen oder zu rennen</i>							
Placebo	57/56	6.72 (2.675)	9.85 (8.382)	3.17 (6.556)	10.95 (0.77)	-	-
ATL 10/10/ 20	57/57	7.45 (4.373)	9.13 (6.936)	1.68 (5.617) p=0.2980 ¹	9.59 (0.77)	-1.37 (1.07) [-3.48, 0.75]	0.2045 0.3393
ATL 20/20/ 40	60/59	7.44 (4.360)	10.20 (7.946)	2.74 (5.341) p=0.8901 ¹	10.28 (0.75)	-0.67 (1.07) [-2.78, 1.43]	0.5291 0.7521
<i>Zeit, um 4 Treppenstufen hinaufzusteigen</i>							
Placebo	57/56	6.01 (5.671)	10.81 (11.292)	4.82 (7.933)	12.11 (0.81)	-	-
ATL 10/10/ 20	57/57	6.94 (6.474)	9.33 (9.396)	2.39 (4.618) p=0.0905 ¹	9.54 (0.81)	-2.57 (1.14) [-4.82, -0.32]	0.0252 0.0467
ATL 20/20/ 40	60/59	7.48 (7.463)	11.21 (10.838)	3.65 (6.665) p=0.5300 ¹	10.52 (0.80)	-1.59 (1.13) [-3.83, 0.65]	0.1626 0.2751
<i>Zeit, um 4 Treppenstufen hinabzusteigen</i>							
Placebo	57/56	5.49 (5.755)	9.56 (10.708)	4.06 (7.819)	10.55 (0.89)	-	-
ATL 10/10/ 20	57/57	6.08 (5.985)	8.49 (9.334)	2.41 (6.162) p=0.3605 ¹	8.80 (0.88)	-1.74 (1.24) [-4.20, 0.71]	0.1626 0.2752
ATL 20/20/ 40	60/59	6.66 (7.202)	9.73 (10.611)	3.05 (7.220) p=0.6619 ¹	9.29 (0.87)	-1.25 (1.24) [-3.69, 1.19]	0.3120 0.4935

¹ Dunnetts-Test vs. Placebo

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

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

Fazit

Nach Einschätzung des CHMPs (2014) spiegeln TFTs typische Aktivitäten wider, die für DMD-Patienten wichtig sind und Prädiktoren für den Zeitpunkt darstellen, wann ein Patient seine Gehfähigkeit verliert (1). Der Zeit-Unterschied zwischen der Ataluren 10/10/20- und

Placebo-Gruppe (2.57 sec), um vier Treppenstufen zu ersteigen, ist größer als der von McDonald et al. (2013) mittels zweier verteilungsbasierter statistischer Verfahren ermittelten MCIDs (2.1 bzw. 2.2 sec) und der Gruppenunterschied war statistisch signifikant (18). Entsprechend der von Escolar et al. (17) für diesen Endpunkt vorgeschlagenen MCID (0.4 log sec entsprechend 1.5 sec) ist der Unterschied zwischen der Ataluren 10/10/20- und der Placebo-Gruppe ebenfalls klinisch relevant.

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

4.3.1.3.1.3 Unerwünschte Ereignisse – RCT

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

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

Tabelle 4-21: Operationalisierung des Endpunkts „Unerwünschte Ereignisse“ (UEs)

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

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

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

Tabelle 4-22: Bewertung des Verzerrungspotenzials für den Endpunkt „Unerwünschte Ereignisse“ in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotential für den Endpunkt UEs wird unter Berücksichtigung der in Tabelle 4-22 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.

In die Sicherheitspopulation (“as-treated”) wurden alle randomisierten Patienten, die irgendeines der Prüfpräparate entsprechend ihres Randomisierungscode erhalten hatten, eingeschlossen. Die Mehrzahl der Patienten aller Gruppen entwickelte mindestens ein behandlungsspezifisches UE (Placebo 98.2%, Ataluren 10/10/20 96.5%, Ataluren 20/20/40 95.0%). Die behandlungsspezifischen UEs waren mehrheitlich milden oder mäßigen Schweregrads. Ein Fortschreiten der Erkrankung war das häufigste schwere (Grad 3) UE (10.5% Placebo, 7.0% Ataluren 10/10/20, 8.3% Ataluren 20/20/40). Die Prüfarzte waren angehalten, bei Patienten, die ihre Gehfähigkeit während der Studie verloren, eine schwere (Grad 3) Progression der Erkrankung zu dokumentieren. Die Einschätzung des Zusammenhangs der UEs mit den Prüfpräparaten durch die Prüfarzte war in den Gruppen insgesamt vergleichbar. Nur 7 Patienten (n=3 Placebo, n=2 jeweils Ataluren 10/10/20 und 20/20/40)

entwickelten ein schwerwiegendes behandlungsspezifisches UE. Keiner der Patienten brach die Studie aufgrund eines behandlungsspezifischen UEs vorzeitig ab und lebensbedrohliche oder fatale behandlungsspezifische UEs wurden nicht berichtet (Tabelle 4-23). Insgesamt 44 Patienten (n=9 bzw. n=22 Ataluren 20/20/40 bzw. 10/10/20, n=13 Placebo) brachen die Behandlung aufgrund eines oder mehrerer meist gastrointestinaler UEs temporär ab.

Tabelle 4-23: Überblick der behandlungsspezifischen UEs aus RCT (Sicherheitspopulation)

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

Quelle (Studienreport): S. 134

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

Die häufigsten behandlungsspezifischen UEs, die bei mindestens 5% der Patienten auftraten, sind in Tabelle 4-24 dargestellt. Die häufigsten behandlungsspezifischen UEs, die bei 20% und mehr Patienten in einer der Gruppen auftraten, waren Erbrechen (insgesamt 46.6%), Kopfschmerzen (29.3%), Diarrhoe (24.1%), Nasopharyngitis (20.7%), Pyrexie (19.0%), Husten (19.0%) und oberer Abdominalschmerz (17.8%). Behandlungsspezifische UEs, die bei mindestens 10% der Patienten in einer Gruppe auftraten und deren Inzidenz von Placebo zu Ataluren 10/10/20 respektive 20/20/40 zunahm, waren Übelkeit (12.3%, 14.0% bzw. 16.7%), Abdominalschmerz (7.0%, 12.3% bzw. 16.7%), Schmerz in den Extremitäten (10.5%, 12.3% bzw. 13.3%), Flatulenz (7.0%, 8.8% bzw. 11.7%) und Verstopfung der Nase (7.0%, 8.8% bzw. 10.0%). Hinsichtlich der Verteilung der UEs innerhalb der verschiedenen Systemorganklassen (SOKs) ergaben sich keine auffälligen Gruppenunterschiede. Die meisten der behandlungsspezifischen UEs traten innerhalb der SOKs „Gastrointestinale Erkrankungen“ (Erbrechen, Diarrhoe und Abdominalschmerz), „Infektionen (und parasitäre Erkrankungen)“ (Nasopharyngitis und Infektionen der oberen Atemwege), „Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen“ (Stürze bedingt durch die DMD und durch die Biopsien bedingter prozeduraler Schmerz), „Muskuloskeletale Erkrankungen“ (Schmerz in den Extremitäten und Rückenschmerzen) und „Erkrankungen des Nervensystems“ (Kopfschmerzen) auf. Eine Analyse der Inzidenzen behandlungsspezifischer UEs

per 12-Wochen-Interval erbrachte keine Hinweise auf eine Zunahme der UEs während des Studienverlaufs (Daten nicht gezeigt).

Tabelle 4-24: Behandlungsspezifische UEs stratifiziert nach Organsystem aus RCT, die bei mehr als 5% der Patienten der Gesamtpopulation über 48 Wochen auftraten (Sicherheitspopulation)

SOK / UE MedDRA PT, n (%)¹	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
<i>Erkrankungen des Gastrointestinaltrakts</i>	37 (64.9)	42 (73.7)	44 (73.3)
Erbrechen	22 (38.6)	32 (56.1)	27 (45.0)
Diarrhöe	14 (24.6)	11 (19.3)	17 (28.3)
Oberer Abdominalschmerz	9 (15.8)	9 (15.8)	13 (21.7)
Übelkeit	7 (12.3)	8 (14.0)	10 (16.7)
Abdominalschmerz	4 (7.0)	7 (12.3)	10 (16.7)
Flatulenz	4 (7.0)	5 (8.8)	7 (11.7)
Magenbeschwerden	0	4 (7.0)	5 (8.3)
<i>Allgemeine Erkrankungen und Erkrankungen am Applikationsort</i>	21 (36.8)	23 (40.4)	20 (33.3)
Pyrexie	12 (21.1)	14 (24.6)	7 (11.7)
Fortschreiten der Erkrankung	6 (10.5)	4 (7.0)	5 (8.3)
Asthenie	2 (3.5)	3 (5.3)	4 (6.7)
<i>Infektionen und parasitäre Erkrankungen</i>	43 (75.4)	38 (66.7)	39 (65.0)
Nasopharyngitis	13 (22.8)	13 (22.8)	10 (16.7)
Infektionen der oberen Atemwege	10 (17.5)	9 (15.8)	11 (18.3)
Influenza	8 (14.0)	6 (10.5)	7 (11.7)
Gastroenteritis	4 (7.0)	9 (15.8)	3 (5.0)
Rhinitis	2 (3.5)	6 (10.5)	3 (5.0)
Ohrinfektion	3 (5.3)	3 (5.3)	4 (6.7)
Virale Gastroenteritis	3 (5.3)	4 (7.0)	3 (5.0)
<i>Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen</i>	26 (45.6)	28 (49.1)	31 (51.7)
Sturz	7 (12.3)	11 (19.3)	6 (10.0)
Prozeduraler Schmerz	7 (12.3)	6 (10.5)	8 (13.3)
Kontusion	3 (5.3)	6 (10.5)	4 (6.7)
Verstauchung des Gelenks	1 (1.8)	4 (7.0)	4 (6.7)
<i>Untersuchungen</i>	4 (7.0)	10 (17.5)	6 (10.0)
Körpergewicht reduziert	1 (1.8)	5 (8.8)	3 (5.0)
<i>Stoffwechsel- und Ernährungsstörungen</i>	3 (5.3)	7 (12.3)	6 (10.0)
Verminderter Appetit	2 (3.5)	5 (8.8)	5 (8.3)
<i>Sklelettmuskulatur- und Bindegeweberkrankungen</i>	19 (33.3)	25 (43.9)	28 (46.7)
Schmerz in den Extremitäten	6 (10.5)	7 (12.3)	8 (13.3)
Rückenschmerz	5 (8.8)	9 (15.8)	6 (10.0)
Arthralgie	2 (3.5)	2 (3.5)	6 (10.0)
Muskelspasmen	5 (8.8)	3 (5.3)	1 (1.7)
Muskelschwäche	1 (1.8)	3 (5.3)	5 (8.3)
<i>Erkrankungen des Nervensystems</i>	17 (29.8)	25 (43.9)	18 (30.0)

SOK / UE MedDRA PT, n (%) ¹	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
Kopfschmerzen	14 (24.6)	22 (38.6)	15 (25.0)
Schwindel („dizziness“)	4 (7.0)	3 (5.3)	3 (5.0)
<i>Erkrankungen der Atemwege, des Brustraums und Mediastinums</i>	18 (31.6)	20 (35.1)	22 (36.7)
Husten	11 (19.3)	9 (15.8)	13 (21.7)
Nasenverstopfung	4 (7.0)	5 (8.8)	6 (10.0)
Schmerzen im Oropharynx	4 (7.0)	6 (10.5)	4 (6.7)
Rhinorrhoe	6 (10.5)	4 (7.0)	0
<i>Erkrankungen der Haut und des Unterhautzellgewebes</i>	18 (31.6)	19 (33.3)	14 (23.3)
Ausschlag („rash“)	5 (8.8)	4 (7.0)	8 (13.3)
Narbe	3 (5.3)	4 (7.0)	5 (8.3)

Quelle (Studienbericht): S. 138-139

¹ Patienten, bei denen das gleiche UE mehrmals auftrat, wurden nur einmal genannt.

Die häufigsten aus Sicht der Prüfarzte möglicherweise oder wahrscheinlich mit der Behandlung in Zusammenhang stehenden UEs ($\geq 5\%$ über alle Behandlungsgruppen) waren Erbrechen (17.8%), Diarrhoe (12.6%), Schmerzen im oberen Abdomen (10.3%), Flatulenz (8.6%), Übelkeit (8.0%), Kopfschmerzen (7.5%) und verminderter Appetit (6.3%). Insgesamt war die Inzidenz der mit der Behandlung in Zusammenhang stehenden UEs zwischen den Gruppen vergleichbar, allerdings nahmen die UEs Erbrechen, Flatulenz, Magenbeschwerden und Ermüdung von Placebo über Ataluren 10/10/20 respektive 20/20/40 zu (Tabelle 4-25). Bei nur einem Patienten der Ataluren 20/20/40-Gruppe wurde aufgrund eines möglicherweise Behandlungs-bedingten UEs (Grad 2 „rash“) die Behandlung zeitweise unterbrochen und die Dosis zeitweise reduziert.

Tabelle 4-25: Inzidenz der behandlungsspezifischen mit der Behandlung in Zusammenhang stehenden UEs aus RCT, die bei mehr als einem Patienten auftraten (Sicherheitspopulation)

SOK / UE MedDRA PT, n (%) ¹	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
<i>Patienten mit ≥ 1 mit der Behandlung in Zusammenhang stehendem UE</i>	26 (45.6)	30 (52.6)	33 (55.0)
<i>Erkrankungen des Gastrointestinaltrakts</i>	18 (31.6)	21 (36.8)	28 (46.7)
Abdominale Beschwerden	1 (1.8)	0	2 (3.3)
Abdominalschmerz	3 (5.3)	1 (1.8)	4 (6.7)
Oberer Abdominalschmerz	5 (8.8)	5 (8.8)	8 (13.3)
Abdominale Druckempfindlichkeit	1 (1.8)	0	2 (3.3)
Aerophagie	2 (3.5)	0	1 (1.7)
Konstipation	1 (1.8)	1 (1.8)	2 (3.3)
Diarrhoe	7 (12.3)	3 (5.3)	12 (20.0)
Flatulenz	3 (5.3)	5 (8.8)	7 (11.7)
Häufiger Stuhlgang / -drang	1 (1.8)	0	1 (1.7)
Übelkeit	0	7 (12.3)	7 (11.7)
Regurgitation	0	1 (1.8)	1 (1.7)
Magenbeschwerden	0	2 (3.5)	4 (6.7)

SOK / UE MedDRA PT, n (%)¹	Placebo (n=57)	Ataluren 10/10/20 (n=57)	Ataluren 20/20/40 (n=60)
Erbrechen	7 (12.3)	11 (19.3)	13 (21.7)
<i>Allgemeine Erkrankungen und Erkrankungen am Applikationsort</i>	2 (3.5)	2 (3.5)	5 (8.3)
Ermüdung („fatigue“)	0	1 (1.8)	2 (3.3)
Pyrexie	0	1 (1.8)	1 (1.7)
<i>Untersuchungen</i>	3 (5.3)	6 (10.5)	1 (1.7)
Farbe des Urins nicht normal	1 (1.8)	1 (1.8)	0
Körpergewicht vermindert	1 (1.8)	4 (7.0)	1 (1.7)
<i>Stoffwechsel- und Ernährungsstörungen</i>	2 (3.5)	5 (8.8)	4 (6.7)
Verminderter Appetit	2 (3.5)	5 (8.8)	4 (6.7)
<i>Sklelettmuskulatur- und Bindegeweberkrankungen</i>	0	3 (5.3)	2 (3.3)
Schmerz in den Extremitäten	0	2 (3.5)	1 (1.7)
<i>Erkrankungen des Nervensystems</i>	4 (7.0)	8 (14.0)	5 (8.3)
Schwindel („dizziness“)	1 (1.8)	1 (1.8)	1 (1.7)
Kopfschmerzen	2 (3.5)	7 (12.3)	4 (6.7)
<i>Erkrankungen der Niere und der Harnwege</i>	2 (3.5)	6 (10.5)	1 (1.7)
Enuresis	0	2 (3.5)	0
Pollakisurie	1 (1.8)	1 (1.8)	1 (1.7)
Nierenzyste	1 (1.8)	1 (1.8)	0
<i>Erkrankungen der Atemwege, des Brustraums und Mediastinums</i>	0	4 (7.0)	3 (5.0)
Husten	0	2 (3.5)	2 (3.3)
Epistaxis	0	1 (1.8)	1 (1.7)
<i>Erkrankungen der Haut und des Unterhautzellgewebes</i>	2 (3.5)	1 (1.8)	2 (3.3)
Erythem	0	1 (1.8)	1 (1.7)
Ausschlag („rash“)	1 (1.8)	0	2 (3.3)
<i>Gefäßkrankungen</i>	2 (3.5)	3 (5.3)	0
Hypertension	1 (1.8)	3 (5.3)	0

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

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

Die Inzidenz schwerer UEs (Grad 3) war in den Gruppen vergleichbar (15.8% Placebo, 14.0% bzw. 16.7% Ataluren 10/10/20 bzw. 20/20/40; Tabelle 4-26). Bei den meisten der schweren UEs handelte es sich um Einzelfälle und eine Krankheitsprogression wurde in allen Gruppen als häufigste Grad 3-UE berichtet.

Tabelle 4-26: Inzidenz der behandlungsspezifischen schweren UEs aus RCT (Sicherheitspopulation)

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

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

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

In der Placebo-Gruppe wurden zwei möglicherweise mit der Behandlung in Zusammenhang stehende schwerwiegende UEs berichtet (Grand-Mal-Anfall [re-klassifiziert: unwahrscheinlicher Zusammenhang] und Schmerzen im Abdomen). Jeweils zwei Patienten beider Ataluren-Gruppen entwickelten ein schwerwiegendes UE (Tabelle 4-27), die von den

Prüfärzten als nicht mit der Behandlung in Zusammenhang stehend klassifiziert wurden oder ein Zusammenhang unwahrscheinlich war. Der Fall einer paroxysmalen supraventrikulären Tachykardie unter Ataluren 20/20/40 wurde als wahrscheinlich nicht mit der Behandlung in Zusammenhang stehend bewertet, da der Patient bereits früher ähnliche Episoden hatte.

Tabelle 4-27: Inzidenz der behandlungsspezifischen schwerwiegenden UEs aus RCT (Sicherheitspopulation)

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

Quelle (Studienbericht): S. 2556

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

Da Patienten mit DMD häufig mit Kortikosteroiden behandelt werden und auch in der RCT die Mehrzahl der Patienten Kortikosteroide als Begleitmedikation erhielten (71.3%), wurde die Inzidenz von UEs in Abhängigkeit von der Behandlung mit Kortikosteroiden analysiert. Hinsichtlich dieser Analyse ist zu beachten, dass die Gruppengrößen nicht ausgeglichen waren (Kortikosteroide: ja / nein) und die Patienten unterschiedliche Kortikosteroide und Regimes erhielten. Der deutlichste Unterschied zeigte sich beim Auftreten von Kopfschmerzen, welche bei 35.5% vs. 14.0% der Patienten, die Kortikosteroide erhielten bzw. nicht erhielten, auftrat. Die mit Kortikosteroiden behandelten Patienten berichteten öfter über Rückenschmerzen (15.3% vs. 2.0%), Schmerzen in den Extremitäten (13.7% vs. 8.0%) und Muskelkrämpfe (6.5% vs. 2.0%). Hypertension trat als UE nur bei den Kortikosteroid-behandelten Patienten auf (4.0% vs. 0%). Folgende UEs traten bei den Patienten, die keine Kortikosteroide erhielten, häufiger auf als unter Begleitmedikation mit Kortikosteroiden:

Pyrexie (26.0% vs. 16.1%), Fortschreiten der Erkrankung (12.0% vs. 7.3%), Rhinitis (10.0% vs. 4.8%), virale Gastroenteritis (10.0% vs. 4.0%), Rhinorrhoe (8.0% vs. 4.8%), Lymphadenopathie (6.0% vs. 0.8%), Sinusitis (6.0% vs. 3.2%), virale Infektionen (6.0% vs. 3.2%), Haut-Läsionen (6.0% vs. 0%) und saisonale Allergien (10.0% vs. 0%).

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

Fazit

Das UE-Profil von Ataluren war hinsichtlich Inzidenz, Art, Schwere und möglichem Kausalzusammenhang mit den Prüfpräparaten insgesamt demjenigen unter Placebo vergleichbar. Übelkeit, Abdominalschmerzen, Schmerzen in den Extremitäten, Flatulenz und Verstopfung der Nase traten unter Ataluren verglichen mit Placebo etwas häufiger auf. Bei keinem der wenigen schwerwiegenden UEs wurde ein Zusammenhang mit der Therapie mit Ataluren vermutet und keiner der Patienten brach die Therapie aufgrund von UEs ab. Lebensbedrohliche oder fatale UEs wurden nicht berichtet. In dem Beurteilungsbericht zu Ataluren schloss das CHMP (2014), dass die häufigsten mit der Behandlung in Zusammenhang stehenden UEs (Erbrechen, Diarrhoe, Schmerz im oberen Abdomen, Flatulenz, Nausea, Kopfschmerzen und verminderter Appetit) bei einer äußerst belastenden und lebensbedrohlichen Erkrankung wie der DMD keine bedeutenden Sicherheitsbedenken hervorrufen. Nach Ansicht des CHMPs (2014) sind die identifizierten Risiken von Ataluren akzeptabel (1).

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

4.3.1.3.2 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.

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.

Bei der Planung der Phase 2b Studie wurde davon ausgegangen, dass die Krankheitsprogression einen Effektmodifikator für die Effektgrößen der einzelnen Endpunkte darstellt. Die folgenden Stratifizierungsfaktoren (Trennpunkte für Subgruppenanalysen) wurden für die Studie PTC124-GD-007-DMD a priori festgelegt:

- Alter (<9 Jahre vs ≥ 9 Jahre)
- 6-Minuten Gehstrecke zur Baseline (≥ 350 Meter vs <350 Meter)
- Kortikosteroidbehandlung (ja vs. nein)

Während die Kortikoidbehandlung einen unabhängigen Effektmodifikator darstellt, sind das Alter und die 6MWD zur Baseline abhängige Effektmodifikatoren, die von der Krankheitsprogression abhängig sind (39).

Obwohl für die oben angeführten Subgruppen der einzelnen Stratifizierungsfaktoren in der Studie PTC124-GD-007-DMD unterschiedliche Effekte nachweisbar waren, waren die Interaktionen zwischen den Subgruppen eines Stratifizierungsmerkmals statistisch nicht signifikant. Für die Planung der Trennpunkte von Effektmodifikatoren wurden bei der Planung der Phase 3 Studie PTC124-GD-007-DMD aus diesem Grund neue Erkenntnisse der Arbeitsgruppe von Sweeney et al. (62) herangezogen. Sweeney et al. studierten unter Anwendung von Magnetresonanztomographie in einer natürlichen Verlaufsstudie den Zusammenhang zwischen der Gehfähigkeit im 6MWT und der Degeneration des gesunden Muskelgewebes (Krankheitsprogression) in DMD Patienten. Die Ergebnisse deuten darauf hin, dass folgender Effektmodifikator (Trennpunkte) den Zusammenhang zwischen der Krankheitsprogression und den jeweiligen Effektgrößen am besten beschreibt:

- BL-6MWD (<300 m, ≥ 300 m bis <400 m, ≥ 400 m)

Für die Subgruppenanalysen der Phase 3 Studie, und die für dieses Dossier durchgeführten Metaanalysen der beiden eingeschlossenen Studien PTC124-GD-007-DMD und PTC124-GD-002-DMD wurden daher die oben angeführten Trennpunkte (prospektiv) definiert.

Zudem wurden für die Auswertung der Effektgrößen aus der Phase 2b Studie PTC124-GD-007-DMD die prädefinierte MMRM Analyseverfahren und die kITT Population eingesetzt (10-13). Die Ergebnisse der einzelnen Subgruppen für die bewertungsrelevanten Endpunkte sind in Tabelle 4-29 zusammengefasst.

Tabelle 4-28: Effektgrößen der Subgruppen BL-6MWD (<300 m, ≥300 bis <400 m, ≥400 m)

Endpunkt	LS Mean Differenz (SE) [p-Wert]		
	< 300m	300-400m	≥ 400m
	6MWD	20.8996 (27.6798) [0.4513]	49.9183 (21.8003) [0.0233]
4 Stufen hinauf	-4.5125 (2.1805) [0.0401]	-3.3951 (1.7407) [0.0528]	-0.5525 (1.7744) [0.7559]
4 Stufen hinab	-0.7267 (2.3468) [0.5538]	-4.0840 (1.8730) [0.0306]	-0.3633 (1.9116) [0.8495]
10 Meter gehen/laufen	-1.1537 (1.9285) [0.5505]	-3.0868 (1.5398) [0.0466]	-0.1490 (1.5698) [0.9245]
kITT Population, prädefiniertes MMRM Modell in PTC124-007 p-Wert: t-Test ($Pr > t $)			

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.

Bei der DMD handelt es sich um eine sehr schwerwiegende, progredient verlaufende Erkrankung mit hohem Leidensdruck für die Patienten sowie deren familiäres Umfeld. Die Behandlung mit Ataluren führt verglichen mit Placebo (SOC) zu einer Abschwächung schwerwiegender (bzw. schwerer) Symptome (bzw. Folgekomplikationen), in diesem Falle

eines für die Patienten klinisch bedeutsamen Effekts auf die Gehfähigkeit, der Einfluss auf die Krankheitsprogression hat. Der Therapieeffekt ist klinisch relevant, was sich auch im EPAR der CHMP bestätigt findet (1).

In Tabelle 4-29 sind die Ergebnisse für die einzelnen patientenrelevanten Endpunkte der Studie PTC124-GD-007-DMD zusammenfassend dargestellt. Der Einschluss der korrigierten ITT Population (kITT) für die Berechnung der Effektgrößen wurde in den tragenden Gründen zum ersten Beschluss über die Nutzenbewertung für Ataluren (21. Mai 2015) als grundsätzlich plausibel eingestuft. Andererseits wurden die im statistischen Auswertungsmodell *post-hoc* durchgeführten Änderungen (augmentiertes MMRM Modell) als potentiell verzerrt eingestuft.

Tabelle 4-29: Übersicht der Ergebnisse der patientenrelevanten Endpunkte aus RCT: Ataluren 10/10/20 vs. Placebo

Patientenrelevanter Endpunkt	Ataluren 10/10/20 vs. Placebo	p-Wert (nominal)
Änderung der 6MWD (augmentiertes MMRM Modell, kITT)	Mittelwert 31.7 m (95% KI: 5.1; 58.3)	0.0197
Zeit bis zu einer 10%igen Verschlechterung der 6MWD (Cox-proportionales Hazard-Modell, ITT)	HR 0.508 (95% KI: 0.273; 0.946)	p=0.0326
Anteil der Patienten mit $\geq 10\%$ iger Verschlechterung der 6MWD (CMH-Test; ITT)	OR 0.3 (95% KI: 0.1; 0.9)	p=0.0423
Zeit, um 4 Treppenstufen hinaufzusteigen (sec) (augmentiertes MMRM Modell, kITT)	LSM Differenz -2.57 s (95% KI: -4.8; -0.32)	p=0.0252
Zeit, um 4 Treppenstufen hinabzusteigen (sec) (augmentiertes MMRM Modell, kITT)	LSM Differenz -1.74 s (95% KI: -4.2; 0.71)	p=0.1626
Zeit, um 10 m zu gehen oder zu laufen (sec) (augmentiertes MMRM Modell, kITT)	LSM Differenz -1.37 s (95% KI: -3.48; 0.75)	p=0.2045

Für dieses Dossier wurden daher im Nachgang zur ersten Nutzenbewertung von Ataluren die Ergebnisse der bewertungsrelevanten Endpunkte der Studie PTC124-GD-007-DMD unter Einschluss der kITT Population und des im SAP *prädefinierten* MMRM Modell nachberechnet. Die Ergebnisse hierzu sind in Tabelle 4-30 dargestellt (10-13). Für die in diesem Dossier durchgeführten Meta-Analysen der beiden eingeschlossenen Studien PTC124-GD-007-DMD und PTC124-GD-020-DMD wurden diese Auswertungen eingesetzt.

Tabelle 4-30: Ergebnisse unter Einschluss der kITT Population und des prädefinierten MMRM Modells

Patientenrelevanter Endpunkt	LSM Differenz (SE) 95% KI Ataluren 10/10/20 vs. Placebo	p-Wert
Änderung der 6MWD (m) (prädefiniertes MMRM Modell, kITT)	28.6259 (14.7030) 95% KI: -0.4019; 57.6537	p=0.0532
Zeit bis zu einer 10%ige Verschlechterung der 6MWD (Cox-proportionales Hazard- Modell, ITT)	HR 0.508 (95% KI: 0.273; 0.946)	p=0.0326
Anteil der Patienten mit ≥10%iger Verschlechterung der 6MWD (CMH-Test; ITT)	OR 0.3 (95% KI: 0.1; 0.9)	p=0.0423
Zeit, um 4 Treppenstufen hinaufzusteigen (sec) (prädefiniertes MMRM Modell, kITT)	-2.4626 (1.1395) 95% KI: -4.7124; -0.2128	p=0.0321
Zeit, um 4 Treppenstufen hinabzusteigen (sec) (prädefiniertes MMRM Modell, kITT)	-1.6956 (1.2390) 95% KI: -4.1419; 0.7506	p=0.1730
Zeit, um 10 m zu gehen oder zu laufen (sec) (prädefiniertes MMRM Modell, kITT)	-1.3696 (1.0510) 95% KI: -3.4446; 0.7053	p=0.1943

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

<< Angaben des pharmazeutischen Unternehmers >>

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.**

<< Angaben des pharmazeutischen Unternehmers >>

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 4-31: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

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

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

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

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

Tabelle 4-32: 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.

<< Angaben des pharmazeutischen Unternehmers >>

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 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 4-34: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> 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.

<< Angaben des pharmazeutischen Unternehmers >>

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 4-35: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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*

<< Angaben des pharmazeutischen Unternehmers >>

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 4-36: 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.

<< Angaben des pharmazeutischen Unternehmers >>

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 4-37: 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 4-38: 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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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

4.3.2.3.2 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.

<< Angaben des pharmazeutischen Unternehmers >>

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

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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 4-39: 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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.**

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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 4-40: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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.

<< Angaben des pharmazeutischen Unternehmers >>

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 2005¹³,

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

Molenberghs 2010¹⁴). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrunde liegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁵) 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 2006¹⁶) 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.

<< Angaben des pharmazeutischen Unternehmers >>

¹⁴ 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.

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

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

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).

<< Angaben des pharmazeutischen Unternehmers >>

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
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.
17. Escolar DM, Hache LP, Clemens PR, Cnaan A, McDonald CM, Viswanathan V, et al. Randomized, blinded trial of weekend vs daily prednisone in Duchenne muscular dystrophy. *Neurology*. 2011 Aug 2;77(5):444–52.
18. 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.
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.
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.
62. Sweeney L, Vandenborne 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

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):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	08.12.2010	
Zeitsegment	1980 to 2010 week 50	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ¹⁷] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
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

¹⁷ 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

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-A4: Suche nach weiteren Untersuchungen

<< Angaben des pharmazeutischen Unternehmers >>

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:

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

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

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-B4: Suche nach weiteren Untersuchungen

<< Angaben des pharmazeutischen Unternehmers >>

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

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-C4: Suche nach weiteren Untersuchungen

<< Angaben des pharmazeutischen Unternehmers >>

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

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

<< Angaben des pharmazeutischen Unternehmers >>

Anhang 4-D4: Suche nach weiteren Untersuchungen

<< Angaben des pharmazeutischen Unternehmers >>

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 4-41 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 4-41 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-41 (Anhang): Studiendesign und -methodik für Studie <Studienbezeichnung>

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2 b	Genauere Ziele, Fragestellung und Hypothesen	
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	
4b	Studienorganisation und Ort der Studiendurchführung	
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	
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	
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	
9	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	
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm	

Item ^a	Charakteristikum	Studieninformation
	die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	
11	Verblindung	
11a	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?	
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	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	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	
14b	Informationen, warum die Studie endete oder beendet wurde	
a: nach CONSORT 2010.		

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

<< Angaben des pharmazeutischen Unternehmers >>

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 4-42 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie PTC124-GD-007-DMD

Studie: PTC124-GD-007-DMD

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genauere Benennung der Quelle	Kürzel
PTC124-GD-007-DMD Klinischer Studienbericht (9)	A

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)

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

ja **unklar** **nein**

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

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

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/IWR und die Prüfpräparate waren äußerlich nicht unterscheidbar (A)

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)

behandelnde bzw. weiterbehandelnde Personen:

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

ja unklar nein

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

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

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

ja nein

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

Es existieren keine Hinweise auf sonstige (endpunktübergreifende) Aspekte, die das Verzerrungspotenzial beeinflussen könnten (A)

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

niedrig **hoch**

Begründung für die Einstufung:

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

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: 6MWD inklusive Progressor-Analysen

1. Verblindung der Endpunkterheber

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Erhebung und Auswertung des Endpunktes erfolgte verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

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

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

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.

Endpunkt: TFTs

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Erhebung und Auswertung des Endpunktes erfolgte verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

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

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

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

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

niedrig hoch

Begründung für die Einstufung:

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

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Erhebung und Auswertung des Endpunktes erfolgte verblindet.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

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

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

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

Es existieren keine Hinweise auf eine ergebnisgesteuerte Berichterstattung dieses Endpunkts.

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

ja nein

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

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

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

niedrig hoch

Begründung für die Einstufung:

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

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

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

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

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

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

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

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

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)** **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

 unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau. **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**für nicht randomisierte Studien:****Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren** **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

 unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau. **nein:** Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:**3. Verblindung von Patienten und behandelnden Personen****Patient:** **ja:** Die Patienten waren verblindet. **unklar:** Es finden sich keine diesbezüglichen Angaben. **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

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

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein (ansonsten ggf. im Nachhinein von der Biometrie durchzuführen).
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

ja

nein

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

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

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

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

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

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

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

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

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

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

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:
