

Dokumentvorlage, Version vom 18.04.2013

# Dossier zur Nutzenbewertung gemäß § 35a SGB V

*Elosulfase alfa (Vimizim<sup>®</sup>)*

BioMarin Deutschland GmbH

## **Modul 4 A**

*Behandlung von Mukopolysaccharidose IVA (Morquio  
A Syndrom, MPS IVA)*

Medizinischer Nutzen und  
medizinischer Zusatznutzen,  
Patientengruppen mit therapeutisch  
bedeutsamem Zusatznutzen

Stand: 18 Mai 2014

# Inhaltsverzeichnis

	Seite
<b>Tabellenverzeichnis</b> .....	<b>4</b>
<b>Abbildungsverzeichnis</b> .....	<b>9</b>
<b>Abkürzungsverzeichnis</b> .....	<b>11</b>
<b>4 Modul 4 – allgemeine Informationen</b> .....	<b>16</b>
4.1 Zusammenfassung der Inhalte von Modul 4.....	17
4.2 Methodik .....	25
4.2.1 Fragestellung .....	25
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung .....	27
4.2.3 Informationsbeschaffung .....	28
4.2.3.1 Studien des pharmazeutischen Unternehmers .....	28
4.2.3.2 Bibliografische Literaturrecherche .....	28
4.2.3.3 Suche in Studienregistern .....	29
4.2.3.4 Selektion relevanter Studien .....	30
4.2.4 Bewertung der Aussagekraft der Nachweise.....	30
4.2.5 Informationssynthese und -analyse .....	32
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien .....	32
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	33
4.2.5.3 Meta-Analysen.....	40
4.2.5.4 Sensitivitätsanalysen .....	41
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren .....	43
4.2.5.6 Indirekte Vergleiche .....	44
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen .....	47
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	47
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel .....	47
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	47
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	48
4.3.1.1.3 Studien aus der Suche in Studienregistern .....	50
4.3.1.1.4 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	50
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel .....	51
4.3.1.2.1 Studiendesign und Studienpopulationen .....	51
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	57
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	58
4.3.1.3.1 Gehstrecke im 6MWT – RCT .....	58
4.3.1.3.2 Responderanalyse 6MWT – RCT .....	69
4.3.1.3.3 Ausdauer im 3MSCT – RCT.....	82
4.3.1.3.4 KS im Urin – RCT .....	92
4.3.1.3.5 Respiratorische Funktion – RCT.....	101

4.3.1.3.6	MPS HAQ (Selbstversorgung, Hilfe durch die Betreuer und Mobilität) – RCT .....	110
4.3.1.3.7	Endpunkt Anthropometrie – RCT .....	117
4.3.1.3.8	Inzidenz nicht schwerwiegender und schwerwiegender UEs – RCT ...	124
4.3.1.3.9	Inzidenz von Überempfindlichkeits- und mit der Infusion assoziierten Reaktionen (IARs) – RCT .....	134
4.3.1.3.10	Subgruppenanalysen – RCT .....	142
4.3.1.3.11	Zusammenfassung der Ergebnisse aus randomisierten kontrollierten Studien.....	156
4.3.2	Weitere Unterlagen.....	160
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien .....	160
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche .....	161
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche.....	161
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen .....	161
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT .....	161
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT .....	164
4.3.2.2	Nicht randomisierte vergleichende Studien .....	164
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien .....	164
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien.....	164
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien .....	165
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien....	165
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien .....	166
4.3.2.3	Weitere Untersuchungen.....	166
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen .....	166
4.3.2.3.2	Charakteristika der weiteren Untersuchungen .....	168
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen .....	169
4.3.2.3.3.1	Zwischenanalyse Langzeitwirksamkeit – 6MWT– weitere Untersuchungen .....	169
4.3.2.3.3.2	Zwischenanalyse Langzeitwirksamkeit – FVC– weitere Untersuchungen .....	173
4.3.2.3.3.3	Subgruppenanalysen – weitere Untersuchungen .....	176
4.3.2.4	Zusammenfassung der Ergebnisse aus weiteren Unterlagen .....	176
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	177
4.4.1	Beurteilung der Aussagekraft der Nachweise .....	177
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	177
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht .....	182
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte .....	182
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	182
4.5.2	Begründung für die Vorlage nichtrandomisierter vergleichender Studien und weiterer Untersuchungen.....	182
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen .....	183
4.5.4	Verwendung von Surrogatendpunkten .....	183

4.6	Liste der eingeschlossenen Studien.....	185
4.7	Referenzliste.....	186
<b>Anhang 4-A</b>	<b>: Suchstrategien – bibliografische Literaturrecherche .....</b>	<b>192</b>
<b>Anhang 4-B</b>	<b>: Suchstrategien – Suche in Studienregistern.....</b>	<b>194</b>
<b>Anhang 4-C</b>	<b>: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....</b>	<b>196</b>
<b>Anhang 4-D</b>	<b>: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern).....</b>	<b>197</b>
<b>Anhang 4-E</b>	<b>: Methodik der eingeschlossenen Studien – RCT .....</b>	<b>204</b>
<b>Anhang 4-F</b>	<b>: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten .....</b>	<b>215</b>

**Tabellenverzeichnis**

	<b>Seite</b>
Tabelle 4-1: Ein- und Ausschlusskriterien zur Auswahl relevanter klinischer Studien zur Bewertung des Ausmaßes des Zusatznutzens von Vimizim.....	18
Tabelle 4-2: Darstellung der patientenrelevanten Endpunkte .....	20
Tabelle 4-3: Ein- und Ausschlusskriterien zur Auswahl relevanter klinischer Studien zur Bewertung des Ausmaßes des Zusatznutzens von BMN 110.....	27
Tabelle 4-4: Darstellung der patientenrelevanten Endpunkte .....	35
Tabelle 4-5: Änderung des 6MWT stratifiziert nach PODCI-Ausgangswert, die bei Kindern mit Duchenne-Erkrankung mit einer bedeutsamen Änderung des PODCI um 4.5 Punkte (MCID, Transfer und basale Mobilität) einhergehen (32) .....	37
Tabelle 4-6: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel .....	48
Tabelle 4-7: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel .....	50
Tabelle 4-8: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	51
Tabelle 4-9: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel .....	52
Tabelle 4-10: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	53
Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel.....	54
Tabelle 4-12: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	57
Tabelle 4-13: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel .....	58
Tabelle 4-14: Operationalisierung der Gehstrecke im 6MWT.....	59
Tabelle 4-15: Bewertung des Verzerrungspotenzials für 6MWT in RCT mit dem zu bewertenden Arzneimittel .....	61
Tabelle 4-16: Ergebnisse für 6MWT (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik).....	62
Tabelle 4-17: Ergebnisse für 6MWT (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (Primär- und Sensitivitätsanalysen) .....	63
Tabelle 4-18: Ergebnisse für 6MWT (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel: Einfluss zusätzlicher Kovariaten (ITT-Population) .....	66
Tabelle 4-19: Operationalisierung der Responder-Analyse des 6MWT .....	70
Tabelle 4-20: Bewertung des Verzerrungspotenzials für Responder-Analyse 6MWT in RCT mit dem zu bewertenden Arzneimittel .....	71

Tabelle 4-21: 6MWT Responder-Analyse basierend auf den Delphi-Kriterien (ITT, „observed cases“, Post-hoc-Analyse).....	72
Tabelle 4-22: Wilcoxon-Mann-Whitney-U-Test (Gesamtpopulation) .....	77
Tabelle 4-23: Ableitung des relativen Risikos aus der OR (Gesamtpopulation).....	78
Tabelle 4-24: Wilcoxon-Mann-Whitney-U-Test (weiße Population (50)) .....	81
Tabelle 4-25: Effektgrößenunterschiede der Gehstreckenveränderungen von Baseline zu Woche 24 (BMN 2.0 mg/kg/Woche versus Placebo) .....	81
Tabelle 4-26: Ableitung des relativen Risikos aus der OR (weiße Population (50)).....	82
Tabelle 4-27: Operationalisierung des Endpunkts „3MSCT“ .....	84
Tabelle 4-28: Bewertung des Verzerrungspotenzials für den Endpunkt „3MSCT“ in RCT mit dem zu bewertenden Arzneimittel .....	85
Tabelle 4-29: Ergebnisse für den Endpunkt „3MSCT“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik) .....	85
Tabelle 4-30: Ergebnisse für den Endpunkt „3MSCT“ (mittlere Änderung der pro Minute erklommenen Stufen von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (Primär- und Sensitivitätsanalysen) .....	86
Tabelle 4-31: Ergebnisse für den Endpunkt „3MSCT“ (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel: Einfluss zusätzlicher Kovariaten (ITT-Population) .....	90
Tabelle 4-32: Operationalisierung des Endpunkts „uKS“ .....	94
Tabelle 4-33: Bewertung des Verzerrungspotenzials für uKS in RCT mit dem zu bewertenden Arzneimittel .....	94
Tabelle 4-34: Ergebnisse für den Endpunkt „uKS“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik) .....	95
Tabelle 4-35: Ergebnisse für den Endpunkt „uKS“ (normalisiert, prozentuale Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (Primär- und Sensitivitätsanalysen) .....	96
Tabelle 4-36: Ergebnisse für den Endpunkt „uKS“ (normalisiert, prozentuale Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel: Einfluss zusätzlicher Kovariaten (ITT-Population) .....	98
Tabelle 4-37: Operationalisierung des Endpunktes „respiratorische Funktion“ .....	102
Tabelle 4-38: Bewertung des Verzerrungspotenzials für den Endpunkt „respiratorische Funktion“ in RCT mit dem zu bewertenden Arzneimittel .....	103
Tabelle 4-39: Ergebnisse für den Endpunkt „respiratorische Funktion“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik) .....	103
Tabelle 4-40: Ergebnisse für den Endpunkt „MVV“ (mittleres Volumen der ausgeatmeteten Luft in l/min während wiederholter maximaler respiratorischer Anstrengungen von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (Primär- und Sensitivitätsanalysen) .....	105

Tabelle 4-41: Ergebnisse für andere respiratorische Endpunkte aus RCT mit dem zu bewertenden Arzneimittel (ITT) .....	107
Tabelle 4-42: Operationalisierung des Endpunkts „MPS HAQ“ .....	111
Tabelle 4-43: Bewertung des Verzerrungspotenzials des Endpunkts „MPS HAQ“ (Selbstversorgung, Hilfe durch die Betreuer und Mobilität) in RCT mit dem zu bewertenden Arzneimittel .....	112
Tabelle 4-44: Ergebnisse für den Endpunkt „MPS HAQ“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik).....	113
Tabelle 4-45: Ergebnisse des MPS HAQ Scores von Baseline bis Woche 24 aus RCT mit dem zu bewertenden Arzneimittel.....	114
Tabelle 4-46: Ergebnisse des MPS HAQ (Benutzung eines Rollstuhls oder von Gehilfen zu Beginn der Studie und in Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT) .....	116
Tabelle 4-47: Operationalisierung des Endpunkts „Anthropometrie“ .....	118
Tabelle 4-48: Bewertung des Verzerrungspotenzials für den Endpunkt „Anthropometrie“ in RCT mit dem zu bewertenden Arzneimittel .....	119
Tabelle 4-49: Ergebnisse für den Endpunkt „Anthropometrie“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik)	119
Tabelle 4-50: Ergebnisse für den Endpunkt „Anthropometrie“ von Baseline bis Woche 24 aus RCT mit dem zu bewertenden Arzneimittel (ITT, jugendliche Männer ≤18 Jahre, jugendliche Frauen ≤15 Jahre) .....	122
Tabelle 4-51: Operationalisierung des Endpunkts „Inzidenz nicht schwerwiegender und schwerwiegender UEs“ .....	125
Tabelle 4-52: Bewertung des Verzerrungspotenzials für den Endpunkt „Inzidenz nicht schwerwiegender und schwerwiegender UEs“ in RCT mit dem zu bewertenden Arzneimittel.....	126
Tabelle 4-53: Ergebnisse für den Endpunkt „Inzidenz der gesamten nicht schwerwiegenden und der gesamten schwerwiegenden UEs“ (MedDRA PT) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation, MOR-004) (9).....	127
Tabelle 4-54: Endpunkt „Inzidenz der gesamten nicht schwerwiegenden und der gesamten schwerwiegenden UEs“: Placebo vs. BMN 110 2.9 mg/kg/Woche (9).....	128
Tabelle 4-55: Ergebnisse für den Endpunkt „Inzidenz gesamter behandlungsspezifischer UEs“ (MedDRA PT) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation) (9).....	128
Tabelle 4-56: Ergebnisse für den Endpunkt „Inzidenz der Therapie-assoziierten UEs“ (MedDRA PT) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation) (9) .....	130
Tabelle 4-57: Endpunkt „Inzidenz der mit der Behandlung in Zusammenhang stehenden UEs“: Placebo vs. BMN 110 2.0 mg/kg/Woche (9) .....	133
Tabelle 4-58: Operationalisierung des Endpunkts „Überempfindlichkeitsreaktionen und IARs“ .....	135

Tabelle 4-59: Bewertung des Verzerrungspotenzials für den Endpunkt „Überempfindlichkeitsreaktionen und IARs“ in RCT mit dem zu bewertenden Arzneimittel.....	136
Tabelle 4-60: Ergebnisse für Endpunkt „Inzidenz von Überempfindlichkeitsreaktionen“ aus RCT mit dem zu bewertenden Arzneimittel (9) .....	136
Tabelle 4-61: Endpunkt „Inzidenz von Überempfindlichkeitsreaktionen“: Placebo vs. BMN 110 (9) .....	138
Tabelle 4-62: Ergebnisse für Endpunkt „IARs“ aus RCT mit dem zu bewertenden Arzneimittel.....	139
Tabelle 4-63: Endpunkt „Inzidenz von IAR“: Placebo vs. BMN 110 (9).....	141
Tabelle 4-64: Subgruppenanalysen für den Endpunkt „6MWT“ (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel in der ITT-Population (9) .....	144
Tabelle 4-65: Subgruppenanalysen für den Endpunkt „3MSCT“ (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel in der ITT-Population (9) .....	147
Tabelle 4-66: Subgruppenanalysen für den Endpunkt „uKS“ aus RCT mit dem zu bewertenden Arzneimittel in der ITT-Population (9) .....	150
Tabelle 4-67: Subgruppenanalysen für den Endpunkt „MVV“ aus RCT mit dem zu bewertenden Arzneimittel in der ITT-Population (9) .....	153
Tabelle 4-68: Approximierte relative Risiken für den Vergleich der individuellen Gehstreckenunterschiede in den Behandlungsgruppen BMN 110 (einmal wöchentlich) versus Placebo .....	157
Tabelle 4-69: Zusammenfassung der Ergebnisse der Wirksamkeitsanalysen aus RCT mit dem zu bewertenden Arzneimittel (9) .....	158
Tabelle 4-70: Zusammenfassung der Ergebnisse der Sicherheitsanalysen aus RCT mit dem zu bewertenden Arzneimittel (9) .....	160
Tabelle 4-71: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche .....	161
Tabelle 4-72: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	162
Tabelle 4-73: Operationalisierung von <Endpunkt xxx>.....	162
Tabelle 4-74: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche .....	162
Tabelle 4-75: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	163
Tabelle 4-76: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien .....	165
Tabelle 4-77: Operationalisierung von <Endpunkt xxx>.....	165
Tabelle 4-78: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien .....	166

Tabelle 4-79: Liste der Studien des pharmazeutischen Unternehmers – nicht RCT mit dem zu bewertenden Arzneimittel.....	167
Tabelle 4-80: Studienpool – nicht RCT mit dem zu bewertenden Arzneimittel.....	167
Tabelle 4-81: Operationalisierung von Langzeitwirksamkeit – 6MWT - weitere Untersuchungen.....	169
Tabelle 82: Gehstrecken im 6MWT von Baseline bis Woche 72 (ITT-Population) .....	170
Tabelle 4-83: Operationalisierung von Langzeitwirksamkeit – FVC - weitere Untersuchungen.....	173
Tabelle 4-84: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	182
Tabelle 4-85 (Anhang): Studiendesign und -methodik für Studie MOR-004.....	205
Tabelle 4-86 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MOR-004 .....	216

**Abbildungsverzeichnis**

	<b>Seite</b>
Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel .....	49
Abbildung 2: Änderung der Gehstrecke im 6MWT über die Zeit bis Woche 24 (ITT-Population) .....	68
Abbildung 3: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	68
Abbildung 4: Responder-Analyse des 6MWTs kategorisiert nach der Änderung der Gehstrecke von Baseline bis Woche 24 („observed cases“) .....	73
Abbildung 5: Responder-Analyse des 6MWTs: KVF der Veränderung von Baseline bis Woche 24 (ITT-Population) .....	75
Abbildung 6: Percentage/Percentage Plot für die Gesamtpopulation (BMN 110 2.0 mg/kg/Woche versus Placebo) .....	76
Abbildung 7: Responder-Analyse des 6MWTs: KVF der Veränderung von Baseline bis Woche 24 (Weiße-Population (50)) .....	79
Abbildung 8: Percentage/Percentage Plot für die weiße Population (BMN 110 2.0 mg/kg/Woche versus Placebo) .....	80
Abbildung 9: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	82
Abbildung 10: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	92
Abbildung 11: Modell-basierte ANCOVA mit Messwiederholung der prozentualen Veränderung des normalisierten uKS (ITT) .....	100
Abbildung 12: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	100
Abbildung 13: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	110
Abbildung 14: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	116
Abbildung 15: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	123
Abbildung 16: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	133
Abbildung 17: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie> .....	142
Abbildung 18: Zusammenfassung der Behandlungseffekte von BMN 110 (einmal wöchentlich) gegenüber Placebo für die Wirksamkeitsendpunkte. Analysepopulation: ITT	159
Abbildung 19: Studiendesign der Studien MOR-004 und MOR-005 (52) .....	169

Abbildung 20: Änderung der Gehstrecke im 6MWT über 72 Wochen (Repeated Measures ANCOVA Mittelwert $\pm$ SE). Analyse Population: ITT (A); Per Protocol (B) .....	172
Abbildung 21: ANCOVA mit Messwiederholung – Prozentuale Veränderung des Mittelwertes (least squares mean $\pm$ SE) der forcierten Vitalkapaazität. Analysepopulation ITT .....	174
Abbildung 22: ANCOVA mit Messwiederholung – Prozentuale Veränderung des Mittelwertes (least squares mean $\pm$ SE) der forcierten Vitalkapaazität. Analysepopulation PP .....	175
Abbildung 23: Prozentuale Veränderung des Mittelwertes ( $\pm$ SE) der forcierten Vitalkapaazität. Analysepopulation ITT .....	176
Abbildung 24: Patientenfluss in Studie MOR-004 .....	214

**Abkürzungsverzeichnis**

<b>Abkürzung</b>	<b>Bedeutung</b>
%	Prozent
§	Paragraph
µg	Mikrogramm
3/6/12-MWT	3/6/12-Minuten-Gehtest
3MSCT	3-Minuten Treppensteigen Test
Abs.	Absatz
al.	alii
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANCOVA	Analysis of Covariance
ARRB	Allergic Reaction Review Board
ATS	American Thoracic Society
BL	Baseline
BMN	BioMarin Pharmaceutical Inc.
BMN 110	recombinant human <i>N</i> -acetylgalactosamine-6-sulfatase (Elosulfase alfa)
bpm	beats per minute
bzw.	beziehungsweise
CDC	Centers for Disease Control and Prevention
cm	Zentimeter
CONSORT	Consolidated Standards of Reporting Trials
COPD	Chronic Obstructive Pulmonary Disease
CRF	Case Report Form
CTX1	type I collagen C-terminal crosslinked C-telopeptide
d	day(s)
d.h.	das heißt
DIMDI	Deutsches Institut für Medizinische Dokumentation
DMC	Data Monitoring Committee
EG	Europäische Gemeinschaft
EKG	Elektrokardiogramm

<b>Abkürzung</b>	<b>Bedeutung</b>
ELISA	Enzyme Linked Immunosorbent Assay
EMA	European Medicines Agency
ERT	Enzyme Replacement Therapy
etc.	et cetera
EUCTR	European Union Clinical Trials Register
FDA	Food and Drug Administration
FET	Forced Expiratory Time
FEV1	Forced Expiratory Volume in 1 Second
FIVC	Forced Inspiratory Vital Capacity
FVC	Forced Vital Capacity
g	Gramm
GALNS	<i>N</i> -Acetylgalactosamin-6-Sulfatase
G-BA	Gemeinsamer Bundesausschuss
ggf.	gegebenenfalls
h	hour(s)
H2	Histamin 2
HAQ	Health Assessment Questionnaire
i.v.	intravenös
IAR(s)	Infusion-Associated Reaction(s)
ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
ICTRP	International Clinical Trials Registry Platform
IgE/G/M	Immunoglobulin E/G/M
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
ITT	Intention to treat
IVRS	Interactive Voice Response System
IWRS	Interactive Web Response System
J	Jahre
KDF	Kumulative Distributionsfunktion
kg	Kilogramm

<b>Abkürzung</b>	<b>Bedeutung</b>
KG	Körpergewicht
KI	Konfidenzintervall
KS	Keratansulfat
KVF	kumulative Verteilungsfunktion
l	Liter
LOCF	Last Observation Carried Forward
LOWESS	Locally Weighted Scatterplot Smoothing
LSM	Least Squares Mean
m	Meter
MAR	Missing At Random
MCID	Minimal clinical important difference (minimal klinisch relevanter Unterschied)
MedDRA	Medical Dictionary for Regulatory Activities
mg	Milligramm
min	Minute(n)
ml	Milliliter
mmHg	Millimeter Quecksilber(säule)
MorCAP	Morquio Clinical Assessment Program
MPS	Mukopolysaccharidose
MTC	Mixed Treatment Comparison
MVV	Maximum Voluntary Ventilation
MW	Mann-Whitney Effektgröße
n / N	Nummer/Anzahl
NAk	neutralisierende Antikörper
NCT	National Clinical Trial
ng	Nanogramm
NIH	National Institute of Health
Nr.	Nummer
OC	Observed Cases
OR	Odds Ratio

<b>Abkürzung</b>	<b>Bedeutung</b>
p	probability
PIIANP	type IIA collagen N-propeptide
PK	Pharmakokinetik
PODCI	Pediatric Outcomes Data Collection Instrument
POR	Proportional odds ratio
PP	Per Protocol
PT(s)	Preferred Term(s)
qow	every other week
RCT	Randomized Controlled Trial
RFT	Respiratory Function Test
rh	recombinant human
RR	Relative Risk
SAP	Statistical Analysis Plan
SD	Standard Deviation
SE	Standard Error
sec.	Sekunde(n)
SGB	Sozialgesetzbuch
SMP	Safety Management Plan
SMQ	Standardised MedDRA Query
SOC	System Organ Class
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes unerwünschtes Ereignis
Tab.	Tabelle
TNF $\alpha$	Tumornekrosefaktor alpha
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
u.a.	unter anderem
u.U.	unter Umständen
UE	unerwünschtes Ereignis

<b>Abkürzung</b>	<b>Bedeutung</b>
UK	United Kingdom
uKS	urinäres Keratansulfat
USA	United States of America
VerfO	Verfahrensordnung
vs.	versus
WHO	World Health Organization
z.B.	zum Beispiel

#### 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 0)
- 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**

Vimizim<sup>®</sup> ist zur Behandlung (Enzymersatztherapie, ERT) der Mucopolysaccharidose IVA (MPS IVA, Morquio A Syndrom) zugelassen. Vimizim enthält rekombinante humane *N*-Acetylgalactosamin-6-Sulfatase (rh GALNS) als Wirkstoff. Die MPS IVA ist eine autosomal rezessiv vererbte Erkrankung, die durch einen Mangel des lysosomalen Enzyms GALNS ausgelöst wird. Die GALNS katalysiert einen Schritt innerhalb des Abbaus von Glykosaminoglykanen. Der Mangel an GALNS führt zur Akkumulation von Keratansulfat (KS) und Chondroitin-6-Sulfat in unterschiedlichen Geweben und Organen, welche zu multisystemischen Defekten führt (1,2). Die Krankheitssymptomatik ist heterogen und vielfältig, und betrifft unter anderem das Wachstum, die Ausdauer und Belastbarkeit, die respiratorischen Funktionen, das Herz- Kreislaufsystem, und die chirurgische Krankheitsgeschichte. Die MPS IVA ist somit eine sehr schwerwiegende multifaktorielle Erkrankung, die mit Kleinwuchs und zunehmender Behinderung bis zur Immobilität und cardiorespiratorischem Versagen einhergeht (3,4).

Bei der MPS IVA handelt es sich um ein seltenes Leiden. Dementsprechend wurde Vimizim am 24. Juli 2009 als Arzneimittel für seltene Leiden gemäß der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 zugelassen (5). Entsprechend gilt der Zusatznutzen von Vimizim durch die Zulassung als belegt. Ziel der vorliegenden Nutzenbewertung ist die Quantifizierung des Zusatznutzens von Vimizim.

Zur Behandlung der MPS IVA ist bislang in Deutschland keine Therapie zugelassen. Die Bewertung für die Intervention Vimizim im Vergleich zu Placebo erfolgte anhand der patientenrelevanten Zielgröße Morbidität (Nutzendimensionen Verbesserung des Gesundheitszustands und unerwünschte Ereignisse).

Die Erstellung des vorliegenden Nutzendossiers erfolgte auf Grundlage der Zulassungsstudien zu Vimizim, die den in Tabelle 4-1 festgelegten Kriterien entsprachen. In die Bewertung einbezogen wurde eine randomisierte, Placebo-kontrollierte Doppelblindstudie bei Patienten mit gesicherter MPS IVA im Alter von fünf Jahren und älter mit einer Gehstrecke zwischen 30 und 325 m. Die Studiendauer war auf 24 Wochen limitiert, da größere chirurgische Eingriffe während der Studie nicht möglich waren, und daher aus ethischen Gründen eine längere Studiendauer nicht vertretbar war.

In diesem Dossier soll das Ausmaß des Zusatznutzens einer Therapie mit Elosulfase alfa für Patienten mit Mucopolysaccharidose IVA (MPS IVA, Morquio A Syndrom) abgeschätzt werden. Für Arzneimittel zur Behandlung eines seltenen Leidens (Orphan Drugs), die nach der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 zugelassen sind, gilt gemäß § 35a SGB V Abs. 1 Satz 10 Halbs 1 der medizinische Zusatznutzen durch die Zulassung als belegt. Das Ausmaß des Zusatznutzens wurde aus den für die Zulassung relevanten Studien nach Maßgabe der in Kapitel 5, § 5 Absatz 7 Satz 1 Nr. 1 bis 4 der Verfo festgelegten Kriterien in Hinblick auf ihre

therapeutische Relevanz bewertet. Auf die vom IQWiG vorgeschlagene Methodik gemäß Anhang A der Dossierbewertung zu Ticagrelor wurde bei der Nutzenbewertung von Elosulfase alfa nicht abgestellt.

### Datenquellen

Die Quantifizierung des Zusatznutzens basiert auf dem Studienbericht der pivotalen Studie MOR-004. Für Arzneimittel zur Behandlung von seltenen Leiden sind nur die Abschnitte 4.4.2 und 4.4.3 auszufüllen, daher war eine systematische Literaturrecherche nicht erforderlich. Eine Studiensuche in den öffentlich zugänglichen Studienregistern wurde jedoch durchgeführt, um die Registereinträge der berücksichtigten Studie zu identifizieren.

Eine vollständige Auflistung der Datenquellen ist in Kapitel 4.7 Referenzliste zu finden.

### Ein-/Ausschlusskriterien für Studien

Da randomisierte, kontrollierte Studien den höchsten Grad der evidenzbasierten Medizin repräsentieren und damit ein direkter Kausalitätsbeweis potentieller Unterschiede zwischen der Intervention und der Vergleichstherapie ermöglicht wird, wurden in der folgenden Nutzenbewertung nur randomisierte kontrollierte Studien berücksichtigt.

Es wurden folgende Einschlusskriterien für die Aufnahme von Studien in das vorliegende Nutzendossier definiert:

Tabelle 4-1: Ein- und Ausschlusskriterien zur Auswahl relevanter klinischer Studien zur Bewertung des Ausmaßes des Zusatznutzens von Vimizim

Ein- oder Ausschlusskriterium	Begründung
E1 Gesicherte Diagnose einer MPS IVA (für Patienten im Alter von $\geq 5$ Jahren)	Das Einschlusskriterium entspricht den Einschlusskriterien der pivotalen Studie. Die Bewertung des Ausmaßes des Zusatznutzens bezieht auf die Gesamtpopulation der in die Zulassungsstudien aufgenommenen Patienten.
E2 Intervention mit BMN 110	Elosulfase alfa ist die derzeit einzige medikamentöse Intervention zur Behandlung des Morquio A Syndroms
E3 Placebo-kontrollierte Studie	Zum Zeitpunkt der Dossiererstellung war zur kausalen Therapie der MPS IVA kein Arzneimittel in Deutschland zugelassen, so dass eine Vergleichstherapie nicht zur Verfügung stand. Die Placebo-Kontrolle gewährleistet die Assaysensitivität der Studie.
E4 Randomisierte, kontrollierte Studien	Nur eine randomisierte kontrollierte Studie ermöglicht eine Strukturgleichheit der zu untersuchenden Gruppen und erlaubt einen Kausalitätsnachweis hoher Ergebnissicherheit.
E5 Mindestens über einen der in Kapitel 4.2.1 vordefinierten Endpunkte wird berichtet.	Für die Quantifizierung des Zusatznutzens werden nur patientenrelevante Endpunkte adressiert.

Ein- oder Ausschlusskriterium	Begründung
E6 Studiendauer von mindestens 24 Wochen	Die Studiendauer entsprach derjenigen einer vorangegangenen Studie (MOR-002) mit BMN 110, in der nach 24 Wochen eine Steigerung der Gehstrecke und der Ausdauer beobachtet wurde. Die Zulassungsstudien anderer bei anderen MPS-Erkrankungen eingesetzter Arzneimittel waren von vergleichbarer Dauer. Bei einer schwerwiegenden Erkrankung wie MPS IVA ist eine längere Verabreichung eines Placebos ethisch kaum vertretbar.
E7 Abgeschlossene und ausgewertete Studien	Nur abgeschlossene und ausgewertete Studien können in die Nutzenbewertung eingeschlossen werden.

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

Die als Ergebnis der Einschlusskriterien (Tabelle 4-1) selektionierte randomisierte, kontrollierte klinische Studie wurde gemäß den Anforderungen des CONSORT 2010 Statements (Items 2b bis 14) und der CONSORT-Flow-Charts beschrieben (Anhang 4-E). Eine zusammenfassende Beurteilung der Methodik und des Designs der eingeschlossenen Studie fand statt. Die Bewertung der Aussagekraft der Nachweise erfolgte für jede eingeschlossene Studie entsprechend der Vorgaben der Verfahrensordnung des G-BA (6) zur Beurteilung des Verzerrungspotentials endpunktübergreifend und endpunktspezifisch (Abschnitt 4.3, Anhang Anhang 4-F). Es wird angemerkt, dass diese Angaben für ein Arzneimittel für seltene Leiden nicht erforderlich gewesen wären. Diese Daten werden auf rein freiwilliger Basis zur strukturierten Darstellung sowie zur Bewertung der Ergebnissicherheit präsentiert.

Die prospektiv definierten primären, sekundären, und tertiären Endpunkte wurden den patientenrelevanten Endpunkten wie Verbesserung des Gesundheitszustandes, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen oder Verbesserung der Lebensqualität zugeordnet (6). Die Endpunkte der Studie MOR-004 wurden nach dem Kriterium ausgewählt, in einer 24-wöchigen Studie signifikante und therapierelevante Effekte demonstrieren zu können. Für die patientenrelevante Zielgröße „Lebensqualität“ liegen daher noch keine Daten vor. Die patientenrelevante Zielgröße Mortalität kann derzeit aufgrund der 24-wöchigen Studiendauer und der Seltenheit der Erkrankung MPS IVA in der durchgeführten Placebo-kontrollierten klinischen Studie MOR-004 nicht evaluiert werden.

In Tabelle 4-2 sind die in der Nutzenbewertung berücksichtigten patientenrelevanten Endpunkte zusammenfassend dargestellt.

Tabelle 4-2: Darstellung der patientenrelevanten Endpunkte

Patientenrelevanter Endpunkt	Endpunktparameter
Morbidität (Nutzendimension: Abschwächung schwerwiegender Krankheitssymptome bzw. Folgekomplikationen)	<ul style="list-style-type: none"> <li>• Gehstrecke im 6-Minuten Gehtest (6MWT inklusive Responderanalysen)</li> <li>• Ausdauer im 3MSCT („3-minute step climbing test“)</li> <li>• Respiratorische Funktion</li> <li>• Verrichtung von Alltagsaktivitäten („MPS Health Assessment Questionnaire“)</li> <li>• Anthropometrie</li> </ul>
	<ul style="list-style-type: none"> <li>• Normalisiertes KS im Urin (uKS); dieser Endpunktparameter dient supportiv der Demonstration der Wirksamkeit des eingesetzten Arzneimittels</li> </ul>
Nebenwirkungen (Nutzendimension: Vermeidung von nicht schwerwiegenden bzw. schwerwiegenden Nebenwirkungen)	<ul style="list-style-type: none"> <li>• Nicht schwerwiegende bzw. schwerwiegende unerwünschte Arzneimittelwirkungen</li> </ul>

Zur Bewertung des Ausmaßes des Zusatznutzens von Vimizim im Vergleich zu Placebo wurde die pivotale Studie MOR-004 herangezogen. Es handelte sich um eine randomisierte, Placebo-kontrollierte Doppelblindstudie der Phase 3 und der Evidenzstufe 1b. In allen Studien wurde das ITT-Prinzip („intention to treat“) angewendet. Nur ein Patient hat die Studie vorzeitig abgebrochen, so dass sich zwischen der ITT-Population und den Patienten mit vollständigen Datensätzen keine nennenswerten Unterschiede ergaben. Es wurden verschiedene Methoden eingesetzt, um fehlende Werte nach vorzeitigem Abbruch der Therapie auszugleichen.

Die Ergebnisse der patientenrelevanten Endpunkte wurden jeweils separat dargestellt. In der Studie wurde besonderer Wert auf eine adäquate Operationalisierung der Endpunkte gelegt. Beispielsweise wurden spezielle Trainings- und Zertifizierungsprogramme entwickelt, um die Variabilität des primären Endpunkts „6MWT“ und des sekundären Endpunkts „3MSCT“ möglichst gering zu halten. Überempfindlichkeitsreaktionen wurden mittels zweier zuvor festgelegter standardisierter Suchalgorithmen identifiziert. Infusionsreaktionen („infusion-associated reaction“, IAR), die als Nebenwirkungen von Enzymersatztherapien bekannt sind, galten ebenfalls als unerwünschte Ereignisse (UEs) besonderen Interesses. Diese umfassten jedes UE, das nach Beginn der Infusion bis einen Tag nach Ende der Infusion auftrat, unabhängig von der Kausalitätsbewertung (auch beispielsweise ein Sturz konnte als IAR gewertet werden wenn er innerhalb eines Tages auftrat).

Die Bewertung des Ausmaßes des Zusatznutzens von Vimizim erfolgte nach den Kriterien des §5 Absatz 7 (2) des 5. Kapitels der VerfO. Der Vorschlag des IQWiG zur quantitativen, Zielgrößen-spezifischen Operationalisierung fand dabei keine Anwendung (6,7).

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

Die vorliegende Nutzenbewertung basiert auf einer randomisierten, Placebo-kontrollierten Doppelblindstudie bei Patienten mit gesicherter MPS IVA im Alter von fünf Jahren und älter mit einer Gehstrecke zwischen 30 und 325 m (N=176, ITT- und Sicherheitspopulation). Die Patienten wurden in folgende Behandlungsarme im Verhältnis 1:1:1 randomisiert: Behandlung mit Placebo, Behandlung mit Vimizim 2 mg/kg KG einmal wöchentlich oder jede zweite Woche über 24 Wochen. Die Dosis von 2 mg/kg KG einmal pro Woche entspricht der in Deutschland zugelassenen Dosierung von Vimizim. Die Studiendauer war auf 24 Wochen limitiert, da chirurgische Eingriffe während der Studie nicht möglich waren, und daher aus ethischen Gründen eine längere Studiendauer nicht vertretbar war.

Die Studie entspricht nach der in der Verfahrensordnung des G-BA dargestellten Evidenzhierarchie der Evidenzstufe 1b (6). Das Verzerrungspotential wurde für die eingeschlossene Studie endpunktübergreifend als gering eingeschätzt. Als Endpunkte wurden nur patientenrelevante Endpunkte für die Nutzenbewertung herangezogen. Einzelne Surrogatparameter (uKS und respiratorische Parameter) wurden berücksichtigt, das Ausmaß des Nutzens konnte für diese Endpunkte allerdings nicht quantifiziert werden. Das Verzerrungspotential der für die Nutzenbewertung herangezogenen Endpunkte wurde als niedrig eingestuft. Die für die Nutzenbewertung relevanten Parameter sind patientenrelevante Parameter. Aufgrund der hohen Qualität der Studie wird die Ergebnissicherheit der vorgelegten Daten als hoch eingestuft.

Das primäre Ziel der Studie war es, den Effekt von Vimizim 2.0 mg/kg KG einmal wöchentlich auf die Ausdauer im 6MWT über 24 Wochen im Vergleich zu Placebo zu ermitteln. Dieser etablierte Endpunkt wurde auch in anderen Studien zu lysosomalen Speicherkrankheiten eingesetzt, da er der sensitivste Endpunkt ist, um in einer zeitlich begrenzten Studie einen signifikanten und therapielevanten Effekt zu zeigen. Das primäre Ziel der Studie wurde erreicht und Vimizim 2.0 mg/kg KG/Woche verlängerte die mittlere Gehstrecke im 6MWT um 22.5 m (LSM, 95% KI 4.0, 40.9; p=0.0174) verglichen mit Placebo, während der Effekt von Vimizim 2.0 mg/kg KG jede zweite Woche nicht signifikant war. Gegenüber Placebo betrug die prozentuale Änderung der Gehstrecke im 6MWT bis Woche 24 14.9% (95% KI 2.7, 27.2; p=0.017) bei den Patienten, die Vimizim einmal wöchentlich erhalten hatten.

Die Sensitivitätsanalysen für die Gehstreckenveränderungen bestätigten die Robustheit der Primäranalyse. Die *a priori* festgelegten Subgruppenanalysen lieferten einen Hinweis auf eine größere Wirksamkeit von Vimizim einmal wöchentlich bei denjenigen Patienten, deren Gehfähigkeit deutlicher eingeschränkt war ( $\leq 200$ ). Da Subgruppenanalysen mit erheblichen Unsicherheiten behaftet sind und aufgrund der kleinen Fallzahl bei hoher Variabilität der

Grunderkrankung, ergeben sich keine Konsequenzen für die Effektgrößenschätzung in dieser Nutzenbewertung.

Eine Responder-Analyse unter Einsatz einer kumulativen Verteilungsfunktion zeigte eine klare und statistisch signifikante Trennung der individuellen Gehstreckenveränderungen zwischen den Behandlungsgruppen Vimizim 2.0 mg/kg KG/Woche und Placebo über das gesamte Responder-Kontinuum. Die Überlegenheit von Vimizim 2.0 mg/kg KG/Woche gegenüber Placebo war statistisch signifikant und mit Sicherheit klinisch relevant.

Die Bestimmung des Ausmaßes des Zusatznutzens für Elosulfase alfa erfolgte anhand der Kriterien des §5 Absatz 7 (2) des 5. Kapitels der Verfo auf Basis einer therapierelevanten „Abschwächung schwerwiegender (krankheitsbedingter) Symptome“, und zeigte einen betächtlichen Zusatznutzen für den patientenrelevanten Endpunkt der Gehstreckenveränderung.

Die Ergebnisse der Ausdauer im 3MSCT zeigten eine numerische Verbesserung unter Behandlung mit Vimizim 2.0 mg/kg KG/Woche gegenüber Placebo.

Die Behandlung mit Vimizim 2.0 mg/kg KG einmal wöchentlich führte bereits nach zwei Wochen zu einer hoch signifikanten Reduktion der normalisierten uKS, die über 24 Wochen stabil blieb (-30.2%, 95% KI -38.5, -22.0;  $p < 0.0001$  bzw. -40.7%, 95% KI -49.0, -32.4;  $p < 0.0001$ ).

Hinsichtlich der respiratorischen Parameters MVV („maximum voluntary ventilation“), FVC, und FEV<sub>1</sub> ergab sich ein positiver Trend zugunsten von Vimizim verglichen mit Placebo, der das Signifikanzniveau allerdings auf Grund der 24-wöchigen Studiendauer nicht erreichte.

Die Auswertung des MPS HAQ zur Funktion der Patienten im Alltag ergab keine signifikanten Gruppenunterschiede, allerdings zeigten sich positive Trends in den Dimensionen ‚Mobilität‘ und ‚Unterstützung durch die Betreuer‘. Auch nahm die Zahl der Patienten, die in Woche 24 einen Rollstuhl benutzten, in der Placebogruppe um fünf Patienten zu, während die Zahl in den Vimizim-Gruppen konstant blieb.

Bei den jüngeren Patienten ergab sich ein positiver Trend in Richtung einer höheren Wachstumsrate unter Behandlung mit Vimizim verglichen mit Placebo. Diese Ergebnisse sind von hoher klinischer Relevanz, waren jedoch auf Grund der 24-wöchigen Studiendauer statistisch nicht signifikant.

Keiner der Patienten brach die Studie aufgrund eines unerwünschten Ereignisses (UE) ab und mehr als 99% der Patienten beendeten die Studie. Fast alle Patienten entwickelten mindestens ein UE und der Prozentsatz der Patienten, die ein mit den Prüfpräparaten in Zusammenhang stehendes UE entwickelten, war in der Vimizim-Gruppe (72.4%) etwas höher als in der Placebo-Gruppe (61.0%) (RR=1.19, 95% KI: 0.92, 1.54;  $p=0.1942$ ). Die häufigsten mit der Behandlung in Zusammenhang stehenden UEs waren Fieber, Erbrechen, Kopfschmerzen und Übelkeit. Nur drei Patienten der Vimizim-Gruppen entwickelten ein schwerwiegendes UE, bei dem ein Zusammenhang mit der Behandlung vermutet wurde (Therapie-assoziiert), ein

Patient in der Gruppe BMN 110 2.0 mg/kg KG jede zweite Woche, und zwei Patienten in der Gruppe BMN 110 2.0 mg/kg KG/Woche. Überempfindlichkeitsreaktionen traten in den Vimizim-Gruppen häufiger aus als in der Placebo-Gruppe (11.9% Placebo, 20.7% Vimizim einmal wöchentlich; RR=1.74, 95% KI: 0.74, 4.12; p=0.2044). Fast alle Patienten entwickelten mindestens ein IAR mehrheitlich milder oder mittelschwerer Ausprägung, was angesichts der breiten Definition eines IAR nicht unerwartet war. Alle Infusionsreaktionen waren durch eine Modifikation der Infusionsrate und/oder medizinische Interventionen kontrollierbar. Nur wenige der Vimizim-Infusionen (0.64% bzw. 1.26% bei Anwendung jede zweite Woche bzw. einmal pro Woche) wurden unterbrochen oder abgebrochen und erforderten gleichzeitig eine medizinische Intervention (meist i.v. Antihistaminika oder i.v. Steroide). Mit Ausnahme von IARs ergab sich kein kausaler Zusammenhang zwischen der Behandlung und klinisch bedeutungsvollen Änderungen der Laborparameter oder Vitalzeichen.

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

Der Zusatznutzen von Vimizim zur Behandlung des seltenen Leidens MPS IVA gilt als durch die Zulassung belegt. Bei der Bewertung des Ausmaßes des Zusatznutzens von Vimizim ist zu berücksichtigen, dass es sich um eine sehr schwerwiegende Erkrankung handelt, für die bislang kein Arzneimittel in Deutschland zugelassen ist und die bis dato hauptsächlich palliativ behandelt wurde (8). MPS IVA-Patienten sind gegenüber der Normalbevölkerung stark in ihrer Mobilität eingeschränkt und bei Fortschreiten der Erkrankung ist ein nicht unerheblicher Teil der Patienten auf einen Rollstuhl angewiesen. Dementsprechend ist die Mobilität für MPS IVA-Patienten von hoher Bedeutung (9) und bei einer derartig schwerwiegenden Erkrankung können bereits kleine Verbesserungen klinisch bedeutsam sein (10). Für die Gehstrecke im 6MWT liegt bislang kein Schwellenwert einer minimal klinisch relevanten Änderung bei MPS-Patienten vor. Dennoch ist festzustellen, dass bei Patienten mit anderen Erkrankungen, die in ihrer Gehfähigkeit weniger eingeschränkt waren als die Teilnehmer der MOR-004-Studie, bereits geringere als die in der Studie beobachteten Verbesserungen im Bereich von 5% bis 10% von den Patienten als positiv wahrgenommen wurden (11). Da eine verminderte Gehfähigkeit die Patienten relevant in ihrem Leben beeinträchtigt und deren Lebensqualität wahrscheinlich deutlich einschränkt (d.h. es handelt sich um ein schwerwiegendes Krankheitssymptom), wird die um nahezu 15% verlängerte Gehstrecke unter Behandlung mit Vimizim einmal wöchentlich als für die Patienten relevant eingestuft. Hier sei jedoch deutlich festgehalten, dass es auf Grund der heterogenen und multisystemischen Krankheitssymptomatik, und auf Grund der unterschiedlichen Progression schwierig und fragwürdig wäre, einen diskreten numerischen Schwellenwert zu bestimmen, der für alle Patienten klinisch relevant sein soll. Aus diesem Grund wurde in diesem Dossier eine quantitative Auswertung der kontinuierliche Responder-analyse durchgeführt, die eine signifikante, und therapeutisch relevante Überlegenheit von Vimizim gegenüber Placebo in der gesamten behandelten Patientenpopulation demonstrierte. Für diesen Endpunkt ergibt sich somit ein beträchtlicher Zusatznutzen. Dies ist in striktem Einklang mit den Kriterien des §5 Absatz 7 (2) des 5. Kapitels der Verfo.

Hinsichtlich anderer patientenrelevanter Endpunkte, die in der pivotalen Studie einen positiven Trend zugunsten von Vimizim zeigten, kann das Ausmaß des Zusatznutzens derzeit nicht quantifiziert werden. Eine Zwischenanalyse einer Verlängerungsstudie zeigte jedoch, dass der therapeutische Effekt der Gehstreckenverbesserung bis zu einer Behandlungsdauer von mindestens 72 Wochen erhalten blieb; auch für den respiratorischen Parameter ‚forcierte Vitalkapazität‘ ist für die Patienten, die mit BMN 100 (einmal wöchentlich) bis zu Studienende behandelt wurden, eine zunehmende und signifikante Verbesserung gegenüber Studienbeginn erkennbar.

Bei den unter Behandlung mit Vimizim beobachteten Infusionsreaktionen handelt es sich um für Enzyersatztherapien bekannte UEs, die dementsprechend auch für Vimizim erwartet wurden. Nur einzelne Patienten entwickelten schwerwiegende mit der Behandlung in Zusammenhang stehende UEs und keiner der Patienten brach die Behandlung aufgrund eines UEs ab. Die Infusionsreaktionen sind durch Modifikation der Infusionsrate und in Einzelfällen medizinische Interventionen kontrollierbar, wobei dies in der Studie MOR-004 nur einen kleinen Bruchteil der Infusionen betraf (1.26% Vimizim einmal wöchentlich).

In der Zusammenschau wird das Ausmaß des Zusatznutzens von Vimizim als beträchtlich eingestuft, da zur Therapie der MPS IVA in Deutschland bislang kein Arzneimittel zugelassen ist und eine bisher nicht erreichte deutliche Verbesserung des therapielevanten Nutzens demonstriert wurde, insbesondere eine Abschwächung schwerwiegender, krankheitsbedingter Symptome.

## 4.2 Methodik

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

### 4.2.1 Fragestellung

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

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

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

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

Ziel der vorliegenden Nutzenbewertung ist es, das Ausmaß des Zusatznutzens von Vimizim<sup>®</sup> 1 mg/ml Konzentrat Lösung zur Infusion mit dem Wirkstoff rekombinante humane *N*-Acetylgalactosamin-6-Sulfatase (rh GALNS) in der zugelassenen Indikation zur Behandlung der Mukopolysaccharidose Typ IVA (MPS IVA, Morquio A Syndrom) zu bewerten. Bei der MPS IVA handelt es sich um ein seltenes Leiden. Vimizim wurde am 24. Juli 2009 als Arzneimittel für seltene Leiden entsprechend der Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999 zugelassen (5). Die MPS IVA ist eine autosomal rezessiv vererbte Erkrankung, die durch einen Mangel des lysosomalen Enzyms GALNS ausgelöst wird. Die GALNS katalysiert einen Schritt innerhalb des Abbaus von Glykosaminoglykanen. Der Mangel an GALNS führt zur Akkumulation von Keratansulfat (KS) und Chondroitin-6-Sulfat in unterschiedlichen Geweben und Organen, die zu einer multisystemischen Erkrankung führt. Im Gegensatz zu anderen MPS-Erkrankungen ist bei MPS IVA die Intelligenz der Patienten üblicherweise nicht beeinträchtigt (1,2).

Nach den gesetzlichen Vorgaben (§ 35a Abs. 1 Satz 10 SGB V) gilt für Arzneimittel zur Behandlung seltener Erkrankungen der medizinische Zusatznutzen durch die Zulassung als belegt. Nachweise zum medizinischen Nutzen und zum medizinischen Zusatznutzen im

Verhältnis zur zweckmäßigen Vergleichstherapie müssen nicht vorgelegt werden. Lediglich das Ausmaß des Zusatznutzens ist für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, nachzuweisen (5. Kapitel § 12 Nr. 1 Satz 2 Verfahrensordnung [VerfO] des Gemeinsamen Bundesausschusses [G-BA]) (6). Das Ausmaß des Zusatznutzens wird auf der Grundlage der Zulassungsstudien durch den G-BA bewertet.

Bei Arzneimitteln für seltene Leiden ist eine Nutzenbewertung bei zuvor festgelegter Vergleichstherapie erst dann erforderlich, wenn der Umsatz des betreffenden Arzneimittels die gesetzliche Grenze von 50 Millionen Euro überschreitet und damit einer uneingeschränkten Nutzenbewertung unterliegt (Anlage II - Format und Gliederung des Dossiers).

Die Patientenpopulation des zu bewertenden Arzneimittels entspricht der zugelassenen Indikation „Behandlung der MPS IVA“. Die Quantifizierung des Zusatznutzens bezieht sich auf die Gesamtpopulation der Patienten mit MPS IVA.

Vimizim ist eine rekombinante Form der humanen GALNS, deren Aminosäuresequenz und *N*-Glykosylierungsstellen dem humanen lysosomalen Enzym entsprechen. Zur Behandlung des seltenen Leidens MPS IVA war zum Zeitpunkt der Dossiererstellung keine Vergleichstherapie zugelassen. Bei Patienten mit MPS IVA stehen zur Zeit nur supportive oder palliative Behandlungsoptionen zur Verfügung, wie korrektive orthopädische oder HNO Eingriffe oder Medikationen zur Behandlung der Schmerzen, von Infektionen oder respiratorischer Erkrankungen oder Sauerstoff bei pulmonalen Manifestationen (1,3,8).

Die Bewertung des Ausmaßes des Zusatznutzens von Vimizim erfolgte anhand patientenrelevanter Endpunkte und derjenigen Zulassungsstudien, die der Zulassung zugrunde lagen und die den zuvor festgelegten Kriterien entsprachen. Die vorliegende Bewertung wurde für die Intervention Vimizim im Vergleich zu Placebo anhand der patientenrelevanten Endpunkte 6-Minuten-Gehtest („6-min walk test“, 6MWT), Ausdauer beim Treppensteigen („3-min step climbing test“, 3MSCT), eines Fragebogens zu Alltagsaktivitäten („MPS Health Assessment Questionnaire“, MPS HAQ), respiratorische und anthropometrische Parameter und unerwünschte Ereignisse (UEs), insbesondere Infusionsreaktionen, sowie schwerwiegende unerwünschte Ereignisse (SUEs) durchgeführt. Die Endpunkte sind den Zielgrößenkategorien Morbidität und Nebenwirkungen zuzuordnen. Daten zur Lebensqualität liegen für Vimizim nicht vor. Der patientenrelevante Endpunkt Mortalität konnte innerhalb der Dauer der eingeschlossenen Studie nicht evaluiert werden. Gesicherte Aussagen zur Mortalität sind erst in Jahren, wenn nicht in Jahrzehnten zu erwarten.

Da randomisierte, kontrollierte Studien den höchsten Grad der evidenzbasierten Medizin repräsentieren und damit ein direkter Kausalitätsbeweis potentieller Unterschiede zwischen der Intervention und der Vergleichstherapie ermöglicht wird, wurden in der folgenden Untersuchung nur randomisierte kontrollierte Studien berücksichtigt. In die Bewertung wurde eine randomisierte, doppelblinde Placebo-kontrollierte Vergleichsstudie mit Vimizim einbezogen. Aufgrund der niedrigeren Ergebnissicherheit werden keine offenen, nicht kontrollierten Studien für die Bewertung des Ausmaßes des Zusatznutzens herangezogen.

Da es sich bei Vimizim um ein Arzneimittel zur Behandlung seltener Leiden handelt sind die Abschnitte 4.4.2 und 4.4.3 zu bearbeiten. Zur strukturierten Darstellung wurden dennoch weitere Abschnitte des Dossiers bearbeitet. Im Folgenden wird Vimizim mit dessen experimentellem Namen „BMN 110“ bezeichnet.

#### 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.*

Die Grundlage dieses Dossiers bilden die der Zulassung von BMN 110 zugrunde liegenden Studien, die den in Tabelle 4-3 definierten Kriterien entsprechen. Die Auswahl der Studien sowie deren Auswertung hinsichtlich patientenrelevanter Endpunkte erfolgten auf Basis der festgelegten Ein- oder Ausschlusskriterien.

Tabelle 4-3: Ein- und Ausschlusskriterien zur Auswahl relevanter klinischer Studien zur Bewertung des Ausmaßes des Zusatznutzens von BMN 110

Ein- oder Ausschlusskriterium	Begründung
E1 Gesicherte Diagnose einer MPS IVA (für Patienten im Alter von $\geq 5$ Jahren)	Das Einschlusskriterium entspricht der zugelassenen Indikation von BMN 110 in Deutschland. Die Bewertung des Ausmaßes des Zusatznutzens bezieht auf die Gesamtpopulation der in die Zulassungsstudien aufgenommenen Patienten.
E2 Intervention mit BMN 110	Elosulfase alfa ist die derzeit einzige medikamentöse Intervention zur Behandlung des Morquio A Syndroms
E3 Placebo-kontrollierte Studie	Zum Zeitpunkt der Dossiererstellung war zur kausalen Therapie der MPS IVA kein Arzneimittel in Deutschland zugelassen, so dass eine Vergleichstherapie nicht zur Verfügung stand. Die Placebo-Kontrolle gewährleistet die Assaysensitivität der Studie.
E4 Randomisierte, kontrollierte Studien	Nur eine randomisierte kontrollierte Studie ermöglicht eine Strukturgleichheit der zu untersuchenden Gruppen und erlaubt einen Kausalitätsnachweis mit hoher Ergebnissicherheit.
E5 Mindestens über einen der in Kapitel 4.2.1 vordefinierten Endpunkte wird berichtet.	Für die Quantifizierung des Zusatznutzens werden nur patientenrelevante Endpunkte adressiert.

Ein- oder Ausschlusskriterium	Begründung
E6 Studiendauer von mindestens 24 Wochen	Die Studiendauer entsprach derjenigen einer vorangegangenen Studie (MOR-002) mit BMN 110, in der nach 24 Wochen eine Steigerung der Gehstrecke und der Ausdauer beobachtet wurde. Die Zulassungsstudien anderer bei anderen MPS-Erkrankungen eingesetzter Arzneimittel waren von vergleichbarer Dauer. Bei einer schwerwiegenden Erkrankung wie MPS IVA ist eine längere Verabreichung eines Placebos ethisch kaum vertretbar.
E7 Abgeschlossene und ausgewertete Studien	Nur abgeschlossene und ausgewertete Studien können in die Nutzenbewertung eingeschlossen werden.

### 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 0) sowie weitere Untersuchungen (Abschnitt 0) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

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

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials

(Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

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

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

Für die Bewertung des Ausmaßes des Zusatznutzens eines Arzneimittels zur Behandlung seltener Leiden sind die Zulassungsstudien relevant. Für das zu bewertende Arzneimittel wurde im Zeitraum von 12 Monaten vor dem Zeitpunkt der Dossiererstellung keine Zulassung in dem für das vorliegende Dossier relevanten Anwendungsgebiet erteilt. Für das vorliegende Arzneimittel ist eine bibliografische Literaturrecherche nicht erforderlich. Da eine zweckmäßige Vergleichstherapie zum Zeitpunkt der Dossiererstellung nicht existierte, entfällt eine Literaturrecherche nach der zweckmäßigen Vergleichstherapie.

#### **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 0) sowie weitere Untersuchungen (Abschnitt 0) muss eine Suche in Studienregistern immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

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

Die Suche soll mindestens in den Studienregistern [clinicaltrials.gov](http://www.clinicaltrials.gov) ([www.clinicaltrials.gov](http://www.clinicaltrials.gov)), EU Clinical Trials Register (EU-CTR, [www.clinicaltrialsregister.eu](http://www.clinicaltrialsregister.eu)), Klinische Prüfungen PharmNet.Bund (<http://www.pharmnet-bund.de/dynamic/de/klinische-pruefungen/index.htm>) sowie über das International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: <http://apps.who.int/trialsearch/>) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern (z. B.

krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern anderer pharmazeutischer Unternehmen ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

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

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

Die Suche in den öffentlich zugänglichen Studienregistern zu BMN 100 erfolgte in folgenden Datenbanken:

- Studienregister Clinicaltrials.gov des National Institute of Health (NIH) ([www.clinicaltrials.gov](http://www.clinicaltrials.gov)); die letzte Suche durchgeführt am 10.02.2014.
- International Clinical Trials Registry Platform Search Portal (ICTRP) der WHO ([apps.who.int/trialsearch/](http://apps.who.int/trialsearch/)); die letzte Suche durchgeführt am 10.02.2014.
- EU Clinical Trials Register ([www.clinicaltrialsregister.eu](http://www.clinicaltrialsregister.eu)); die letzte Suche durchgeführt am 10.02.2014.

Eine Einschränkung der Suchstrategie fand während der Recherche nicht statt. Die Ergebnisse der Recherche sind im Anhang 4-B dargestellt.

#### **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.*

Für die Bewertung des gegenständlichen Arzneimittels wurden diejenigen Zulassungsstudien, die den in Kapitel 4.2.2 festgelegten Kriterien entsprachen, herangezogen. Eine bibliographische Literaturrecherche war nicht erforderlich.

#### **4.2.4 Bewertung der Aussagekraft der Nachweise**

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert

werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

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

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

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

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

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

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

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

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

Der medizinische Zusatznutzen von Arzneimitteln zur Behandlung seltener Leiden gilt nach § 35a Absatz 1 Satz 3 Nummer 2 und 3 SGB V durch die Zulassung als belegt. Nachweise gemäß § 5 Absatz 1 bis 6 Verfo müssen nicht vorgelegt werden (6). Eine Bewertung der Aussagekraft der Nachweise ist für Arzneimittel zur Behandlung seltener Leiden daher nicht erforderlich. Dennoch wurde im vorliegenden Dossier die Aussagekraft der vorgelegten Nachweise bewertet, um die hohe Ergebnissicherheit bei der Bewertung des Ausmaßes des Zusatznutzens zu demonstrieren.

Für Arzneimittel zur Behandlung seltener Leiden ist das Ausmaß des Zusatznutzens für die Anzahl der Patienten und Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, nachzuweisen (6). In früheren Nutzenbewertungen von Arzneimitteln zur Behandlung seltener Leiden hat der G-BA die für die Zulassung relevanten Studien nach Maßgabe der im 5. Kapitel § 5 Abs. 7 Satz 1 Nr. 1 bis 4 Verfo festgelegten Kriterien in Hinblick auf ihre therapeutische Relevanz (qualitativ) bewertet, um das Ausmaß des Zusatznutzens zu bestimmen. Die seitens des IQWiG vorgeschlagene Methodik gemäß Anhang A der Dossierbewertung zu Ticagrelor wurde bei der Nutzenbewertung durch den G-BA nicht angewendet (12). Entsprechend des Vorgehens des G-BA wurde das Ausmaß des Zusatznutzens von BMN 110 (qualitativ) nach Maßgabe der im 5. Kapitel § 5 Abs. 7 Satz 1 Nr. 1 bis 4 Verfo festgelegten Kriterien bewertet.

#### 4.2.5 Informationssynthese und -analyse

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

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)<sup>1</sup>. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-<sup>2</sup> bzw. STROBE-Statements<sup>3</sup> folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

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

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

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

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

Das Design und die Methodik der eingeschlossenen Studie wurden in Abschnitt 4.3.1.2 und in Tabelle 4-85 detailliert beschrieben. Im Anschluss fand eine zusammenfassende Beurteilung der Methodik und des Designs der eingeschlossenen Studie statt. Von den oben beschriebenen Vorgaben wurde nicht abgewichen.

#### 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.*

Hinsichtlich der Patientencharakteristika werden Angaben zum Alter, zu Altersgruppen (5-11, 12-18 und  $\geq 19$  Jahre), Geschlecht, Rasse, Herkunft, der Ergebnisse der Gehstrecke im 6MWT, der Gehstrecken-Kategorien ( $\leq 200$  m,  $> 200$  m), der Ergebnisse des 3MSCT und des normalisierten uKS gemacht.

#### **Patientenrelevante Endpunkte und deren Validität**

Die Verfahrensordnung des G-BA nach § 35a SGB V definiert patientenrelevante Effekte in Paragraph 3 Absatz 1 als Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen oder Verbesserung der Lebensqualität (6). Für die Bewertung des Ausmaßes des Zusatznutzens von BMN 110 wurden die patientenrelevanten Zielgrößen der Morbidität (Nutzendimensionen: Verbesserung des Gesundheitszustands [Mobilität, respiratorische Funktion, Verrichtung von Alltagsaktivitäten und anthropometrische Parameter], sowie der Nebenwirkungen) herangezogen.

Bei der MPS IVA handelt es sich um eine progredient verlaufende Multisystemerkrankung. Die mit MPS IVA assoziierte Morbidität umfasst schwere Atemwegsobstruktionen und Herz-Kreislaufkrankungen, bis hin zu respiratorischem und Herz-Kreislaufversagen. Schwere Knochendysplasien, schwere Kompressionen des Rückenmarks, schwere

Gelenkerkrankungen, leichte bis mittelschwere Erkrankungen der Ohren, leichte Augenerkrankungen, und schwere dentale Probleme sind weiter klassische Morbiditätsmerkmale.

Patienten mit schwerer MPS IVA - charakterisiert durch Herz-Kreislaufkrankungen und Atemwegserkrankungen - überleben häufig die 2. oder 3. Dekade nicht, weniger als 5% werden älter als 40 Jahre (1,4,8). Studien mit dem patientenrelevanten Endpunkt „Mortalität“ liegen derzeit nicht vor, da sie auf Grund der aus ethischen Überlegungen eingeschränkten Studiendauer nicht sinnvoll erhoben werden konnten. Auch würden Mortalitätsstudien in dieser Patientenpopulation Jahre, wenn nicht Jahrzehnte andauern.

Ein validiertes Instrument bzw. validierte Responsekriterien zur Erfassung der Lebensqualität existieren bei MPS IVA-Patienten nicht. Einige Autoren stufen die Ausdauer im 6MWT und 3MSCT bei MPS IV-Patienten unter der Kategorie Lebensqualität ein (2,13), da die Ausdauer Einfluss auf die Unabhängigkeit und folglich die Lebensqualität haben kann (2).

Der in der Studie MOR-004 evaluierte kombinierte Endpunkt aus 6MWT, 3MSCT und MVV wurde trotz möglicher Patientenrelevanz nicht berücksichtigt, da eine formelle Anforderung der evidenzbasierten Medizin – nämlich die Validierung des zusammengesetzten Endpunktes – nicht erfüllt ist (7,14). Zudem kann nicht abgeschätzt werden, welche Änderung für den Patienten klinisch relevant sein könnte.

Prinzipiell können auch andere in der Studie erhobene tertiäre Endpunkte wie radiographische Messungen oder das Ausmaß der Hornhauttrübung als patientenrelevant eingestuft werden. Hinsichtlich dieser Endpunkte ergaben sich keine signifikanten Gruppenunterschiede. Auf die Darstellung dieser Endpunkte wird daher verzichtet. Daten zur Entwicklung von gegen BMN 110 gerichteten Antikörpern wurden zur allgemeinen Information rein deskriptiv dargestellt. Diese Daten wurden bei der Bewertung des Ausmaßes des Zusatznutzens von BMN 110 nicht berücksichtigt; nach derzeitigem Kenntnisstand scheint allerdings kein Zusammenhang zwischen dem Auftreten BMN 110-spezifischer Antikörper und der Wirksamkeit oder Sicherheit von BMN 110 zu bestehen.

Die nachfolgende Tabelle 4-4 gibt eine Übersicht der für die vorliegende Bewertung herangezogenen patientenrelevanten Endpunkte.

Tabelle 4-4: Darstellung der patientenrelevanten Endpunkte

<b>Patientenrelevante Zielgrößenkategorie</b>	<b>Endpunktparameter</b>
Morbidity (Nutzendimension: Abschwächung schwerwiegender Krankheitssymptome bzw. Folgekomplikationen)	<ul style="list-style-type: none"> <li>• Gehstrecke im 6MWT (inklusive Responderanalysen)</li> <li>• Ausdauer im 3MSCT</li> <li>• Respiratorische Funktion</li> <li>• Verrichtung von Alltagsaktivitäten (MPS HAQ)</li> <li>• Anthropometrie</li> </ul>
	<ul style="list-style-type: none"> <li>• uKS dieser Endpunktparameter dient supportiv der Demonstration der Wirksamkeit des eingesetzten Arzneimittels</li> </ul>
Nebenwirkungen (Nutzendimension: Vermeidung von (schwerwiegenden) Nebenwirkungen)	<ul style="list-style-type: none"> <li>• Nicht schwerwiegende (bzw. schwerwiegende) unerwünschte Arzneimittelwirkungen</li> </ul>

*Ausdauerfähigkeiten (6MWT und 3MSCT)*

Bei einer chronischen Erkrankung, die durch fortschreitende Behinderung, häufig bis zur Immobilität, charakterisiert ist, werden Ausdauertests als bedeutende Indikatoren funktionaler Einschränkungen und somit als klinisch relevant eingestuft (3). MPS IVA-Patienten zeigten unabhängig von der Altersgruppe Einschränkungen der Ausdauer im 6MWT und 3MSCT. Bei MPS IVA-Patienten im Alter von 18 Jahren und jünger lag die mittlere Gehstrecke ( $\pm$  SD) im 6MWT bei 218.0 (153.2) m und bei älteren Patienten bei 193.1 (148.5) m (gesamt  $212 \pm 152$  m). Bei Gesunden im Alter zwischen vier und 16 Jahren betrug der untere Normalwert 470-664 m und bei gesunden Erwachsenen etwa 500-580 m (3). Sowohl die EMA (s. EPAR), als auch die FDA sehen die verwendeten Wirksamkeitsendpunkte als klinisch relevant an (15).

Bei dem 6MWT und 3MSCT handelt es sich um submaximale Ausdauertests (10,16), die klinisch eingesetzt werden, um den funktionellen Status des kardiovaskulären, pulmonalen und muskuloskeletalen Systems sowie bestimmter Komponenten des Nervensystems zu evaluieren (10). Der 6MWT wird als globales, integrales Instrument zur Erfassung der Gehfähigkeit eingestuft (17). Mittels Ausdauertests können multiple Systeme simultan klinisch erfasst werden und dies werden daher als effektive Methode angesehen, um die Krankheitsprogression insgesamt zu überwachen (2,10). Gehtests wurden bei einer Vielzahl von Erkrankungen eingesetzt, um die Schwere der Erkrankung und deren Verlauf zu erfassen wie z.B. COPD, pulmonale Hypertension, idiopathische pulmonale Fibrose, idiopathische oder fibrotische interstitielle Pneumonie, kongestive und chronische Herzinsuffizienz, bei

Kindern mit chronischen Erkrankungen wie zystischer Fibrose, Muskelatrophie vom Typ Duchenne sowie MPS I und VI (10,18–22). Bei MPS-Patienten werden zeitlich begrenzte Gehtests (6MWT oder 12MWT) oder die Fähigkeit, Treppen innerhalb eines vorgegebenen Zeitrahmens zu ersteigen, routinemäßig eingesetzt, um die Belastungstoleranz zu überprüfen oder das Ausmaß der Behinderung und die Progression der Erkrankung zu erfassen (10,23,24). Eine europäische Expertengruppe hat den 6MWT als essentielles Instrument zur routinemäßigen Überwachung von MPS II- und MPS VI- Patienten empfohlen (24,25). Der 6MWT (und 3MSCT) wird in einer nicht-interventionellen Verlaufsstudie bei MPS IVA-Patienten zur Überprüfung der körperlichen Funktion eingesetzt und wird bei MPS IVA-Patienten als prognostisch relevant eingestuft (3). In Phase III Zulassungsstudien bei Patienten mit MPS I, II und VI waren der 6MWT oder 12MWT der primäre Zielparame-ter und der 3MSCT war ein Hauptsekundärendpunkt bei MPS VI-Patienten (20,24,26–28). In einem aktuellen Übersichtsartikel (2013) zu MPS IVA werden jährliche Ausdauer- tests empfohlen, um die Mobilität der Patienten und deren Potential zur Selbständigkeit zu überprüfen (2). Die meisten Alltagsaktivitäten werden bei submaximaler Anstrengung ausgeführt und es wird angenommen, dass der 6MWT die bei Alltagsaktivitäten übliche Belastung widerspiegelt (16,29). McDonald et al. (2013) stufen den 6MWT bei Patienten mit Muskelatrophie vom Duchenne Typ als einen klinisch bedeutenden Parameter ein, der die Gehfähigkeit direkt erfasst (17).

Da MPS als schwerwiegende Multiorgan-Erkrankung mit strukturellen kardialen und pulmonalen Fehlbildungen sowie skeletalen Veränderungen einhergeht, können bereits kleine Änderungen der Gehstrecke klinisch signifikant sein und bedeutsame Änderungen der Fähigkeit, Alltagstätigkeiten zu verrichten, widerspiegeln (10). Für andere Erkrankungen wurden minimal klinisch relevante Änderungen (MCID) der Gehstrecke im 6MWT bestimmt, z.B. 54 m (95% KI 37-71 m) oder 35 m bei Patienten mit schwerer COPD (relative MCID 10% oder 15%), 43 m bei älteren Patienten mit Herzversagen, 25 m bei Patienten mit koronarer Herzkrankheit nach koronarem Syndrom (relative MCID 5%), 45 m oder 24-45 m bei idiopathischer pulmonaler Fibrose (relative MCID 6% bis 11%), etwa 33 m bei pulmonaler arterieller Hypertension (relative MCID 10%) und etwa 30 m bei Patienten mit Muskelatrophie vom Duchenne-Typ (11,16,17,30,31). In der Mehrzahl dieser Studien wurden Patienten eingeschlossen, deren mittlere 6MWT-Ausgangsgehstrecke mit >350 m (3,11) über der durchschnittlichen Gehstrecke bei MPS IVA-Patienten lag (218.0±153.2 m bei Patienten ≤18 Jahre, 193.1±148.5 m bei Patienten >18 Jahre, gesamt 212.6±152.2 m). Auf Grund des höheren Behinderungsgrades von MPS IV Patienten sollte der MCID bei letzteren niedriger liegen.

Eine Extrapolation dieser Daten auf Patienten mit MPS wird jedoch als nicht zulässig oder als zumindest fraglich eingeschätzt (10,23). Die Mehrzahl der Daten zu klinisch relevanten Veränderungen im 6MWT wurde bei Erwachsenen, typischerweise älteren Patienten mit entweder kardialen oder pulmonalen Erkrankungen erhoben, deren Symptomatik nicht direkt mit der multisystemischen Erkrankung der MPS vergleichbar ist (10). Bei Patienten mit Muskelatrophie vom Duchenne-Typ wird eine fortbestehende Reduktion der Gehstrecke im 6MWT um 10% als klinisch relevantes Fortschreiten der Erkrankung eingestuft (17). In einer

Studie bei 24 Jungen mit Duchenne-Erkrankung war eine bedeutungsvolle Änderung bei niedrigen PODCI-Ausgangswerten (Transfer/Mobilität; „Pediatric Outcomes Data Collection Instrument“) von 30 auf 34.5 mit einer um 5.6 m verlängerten Gehstrecke im 6MWT assoziiert (150.3 vs. 155.9 m). Bei Patienten, deren PODCI-Ausgangswert nahe dem Normwert war, war eine bedeutsame Änderung des PODCI mit einer um fast 46 m verlängerten Gehstrecke im 6MWT verbunden. Der PODCI ist ein Instrument, welches zur Überprüfung der funktionalen Fähigkeiten bei Kindern mit orthopädischen Behinderungen eingesetzt wird. Bei einer Gehstrecke von etwa 200 m im 6MWT betrug die Änderung im 6MWT, die mit einer klinisch bedeutsamen Änderung des PODCI (4.5 Punkte) verbunden war, 20.9 m. Daher scheint bei Duchenne-Patienten das Ausmaß einer für die Patienten bedeutsamen Änderung der Gehstrecke im 6MWT (5.6 bis zu 45.9 m) von der Ausgangsmobilität abzuhängen, wobei bei Patienten mit größeren Funktionseinschränkungen bereits geringe Änderungen die gesundheitsbezogene Lebensqualität verbessern können (Tabelle 4-5) (32). In einer Untersuchung bei 174 Patienten mit Duchenne-Erkrankung ( $\geq 5$  Jahre) wurde hinsichtlich der Gehstrecke im 6MWT ein MCID von 28.5 oder 31.7 m ermittelt, basierend auf zwei unterschiedlichen statistischen Testverfahren (33).

Tabelle 4-5: Änderung des 6MWT stratifiziert nach PODCI-Ausgangswert, die bei Kindern mit Duchenne-Erkrankung mit einer bedeutsamen Änderung des PODCI um 4.5 Punkte (MCID, Transfer und basale Mobilität) einhergehen (32)

<b>PODCI Bereich (Punkte)</b>	<b>6MWT Bereich (m)</b>	<b>Änderung des 6MWT (m)</b>
30-34.5	150.3-155.9	5.6
40-44.5	165.0-174.7	9.7
50-54.5	189.5-202.2	12.7
60-64.5	225.9-246.8	20.9
70-74.5	276.7-304.9	28.2
80-84.5	344.3-380.8	36.5
90-94.5	431.1-477.0	45.9

Der Antragsteller hat eine Delphi-Konferenz initiiert, um tentative Größenordnungen für MCIDs bei MPS IVA-Patienten festzulegen (9), diese spiegeln allerdings die Erfahrungen und Einschätzungen der Ärzte/Kliniker wider. Daher ist festzustellen, dass bei Patienten mit MPS IVA bislang keine eindeutigen MCIDs hinsichtlich der Änderung der Gehstrecke im 6MWT sowie anderer Patienten-relevanter Endpunkte aus Sicht der Patienten vorliegen. Für konfirmatorische Studien mit relativ kleinen Patientenpopulationen und sehr heterogenen Krankheitsausprägungen ist die Bestimmung/Verwendung eines MCID (der für alle untersuchten Patienten gleichermassen gilt) nicht sinnvoll. In diesem Dossier wird daher die Ausmaßschätzung u.a. mittels Dichotomisierung der primären Endpunktes vorgenommen (s. Kapitel 4.3.1.3.1).

Der 3MSCT wurde in Studien bei MPS-Patienten (MPS I und VI) eingesetzt und wird als robustes Instrument eingeschätzt, um die Ausdauer unter Therapie mit ERTs zu erfassen

(18,19). Bei dem 3MSCT handelt es sich allerdings nicht um einen standardisierten Test (18). In einem aktuellen Übersichtsartikel zu MPS IVA wird empfohlen, den 3MSCT in Anlehnung an eine Studie bei MPS VI-Patienten (Harmatz et al., 2006) (27) auszuführen (2). Bei Patienten mit schwerer COPD reflektierten die pathophysiologischen Veränderungen während des 6MWT nicht diejenigen während eines Treppentests. Die Gehstrecke im 6MWT erlaubt daher keine Vorhersagen über die Fähigkeit von Patienten mit schwerer COPD, Treppen zu steigen (34).

Der 6MWT und 3MSCT werden bei MPS-Patienten routinemäßig zur Therapiekontrolle oder Überwachung des Krankheitsverlaufs eingesetzt und stellen somit Standardinstrumente in der klinischen Praxis dar. Der 6MWT und 3MSCT werden als patientenrelevant eingestuft, da die Ausdauer bei einer mit fortschreitender Behinderung assoziierten Erkrankung für die Patienten von unmittelbarer Bedeutung ist. Eine unter Behandlung weniger progrediente Verschlechterung der Mobilität kann sich sowohl auf die Unabhängigkeit als auch die Lebensqualität der Patienten positiv auswirken.

#### *Urinäre Keratinsulfatausscheidung*

Die urinaire Glykosaminoglykaneausscheidung war auch in anderen Phase III-Studien z.B. bei MPS II- oder MPS VI-Patienten ein sekundärer Endpunkt (24,25), während die KS-Spiegel im Blut nicht erfasst wurden (35).

Wie bereits oben angemerkt, wird die uKS-Elimination supportiver Parameter gewertet, da das uKS das zugrunde liegende pathologische Geschehen widerspiegelt und sich Hinweise für einen Zusammenhang zwischen der uKS-Ausscheidung und der Schwere der Erkrankung sowie der Krankheitsprogression ergeben. Weitere Daten zur Validität des uKS sind in Abschnitt 4.5.4 dargestellt.

#### *Respiratorische Parameter*

Obstruktive Lungenerkrankungen werden als eine der bedeutendsten und schwierigsten Komplikationen bei MPS IVA-Patienten eingestuft (4), wobei die Ätiologie der respiratorischen Dysfunktionen bei MPS IVA-Patienten multifaktoriell ist (2). Obwohl Lungenfunktionstests nicht als diagnostisch eingeschätzt werden, werden serielle Lungenfunktionstests bei MPS IVA-Patienten zur Überprüfung des natürlichen Verlaufs oder zur Therapiekontrolle empfohlen (2). Bei MPS-Patienten stellt die Spirometrie die am häufigsten eingesetzte Methode zur Messung der Vitalkapazität dar (2,23), obwohl bei sehr jungen Patienten oder älteren Patienten aufgrund nachlassender Ausdauer passive Tests geeigneter sein könnten (2). Das MVV ist das maximale Volumen, das ein Patient in einem vorgegebenen Zeitrahmen, üblicherweise 12-15 Sekunden, unter maximaler, wiederholter respiratorischer Anstrengung atmen kann (2). Das MVV-Manöver wird als Korrelat maximaler Atmung während einer Anstrengung angesehen (36). Das MVV ist bei restriktiven Atemwegserkrankungen vermindert, während das Forcierte Expiratorische Sekundenvolumen bei obstruktiven Atemwegserkrankungen vermindert ist (2). Das mittlere MVV ( $\pm$ SD) ist bei

Patienten mit MPS IVA gegenüber Gesunden deutlich reduziert ( $32.4 \pm 22.1$  l/min bei Patienten  $\leq 18$  Jahre,  $42.1 \pm 32.8$  l/min bei Patienten  $> 18$  Jahre) (3).

Die Änderung des MVV war ein Zielparamester in einer ERT-Zulassungsstudie bei MPS VI-Patienten (37). Unter ERT besserte sich der MVV innerhalb von 24 Wochen um etwa 15%, während FEV1 und FVC bis Woche 24 nur minimale Änderungen zeigten (37). Die Ergebnisse der Studie MOR-002 und der entsprechenden Extensionsstudie MOR-100 mit BMN 110 bestätigten dieses Ergebnis und zeigten, dass das MVV innerhalb von 24 Wochen sensibler auf die Behandlung anspricht, während sich die FVC erst unter längerer Behandlung zu verbessern scheint (38).

Bei dem in der pivotalen Studie MOR-004 erhobenen MVV und den spirometrischen Parametern handelt es sich Endpunkte, für die klinisch relevante Änderungen bei MPS IVA-Patienten nicht eindeutig definiert sind. Hinsichtlich anderer respiratorischer Parameter ergaben sich über die 24-wöchige Studiendauer keine Unterschiede zwischen BMN 110 und Placebo. Es wird daher darauf verzichtet, Daten zur minimalen klinischen Relevanz dieser Endpunkte (bei anderen Krankheitsbildern) darzustellen.

#### *Funktion bei Alltagsaktivitäten*

Der in der pivotalen Studie MOR-004 evaluierte Endpunkt MPS HAQ erfasst Funktionen bei üblichen Alltagsaktivitäten oder inwieweit der Patient im Alltag selbständig ist, wird allerdings nicht als klassisches Instrument der Lebensqualität eingestuft. Die Funktion im Alltag wird bei einer mit zunehmender Behinderung assoziierten Erkrankung als unmittelbar patientenrelevant eingestuft. Der MPS HAQ wurde ursprünglich als klinisches Instrument bei MPS I-Patienten entwickelt und wurde anschließend zum Einsatz bei Patienten mit MPS VI adaptiert. Der MPS HAQ wird in einem internationalen Patientenregister zum natürlichen Verlauf der MPS IVA (MorCAP) erfasst (3). Eine formelle Bestimmung eines MCID liegt für den MPS HAQ bei MPS-Patienten jedoch nicht vor. Vorläufige Untersuchungen zeigten jedoch, dass die Domänen Mobilität und Unterstützung durch die Betreuer sensibler sind, als die Domäne Selbstversorgung.

#### *Anthropometrische Parameter*

Bei einer Erkrankung, die sehr häufig mit Kleinwuchs verbunden ist, kann davon ausgegangen werden, dass Änderungen der Körpergröße für die Patienten von sehr hoher klinischer Relevanz sind.

#### *Unerwünschte Ereignisse*

Nebenwirkungen stellen einen Aspekt der therapiebedingten Morbidität dar und sind daher unmittelbar patientenrelevant. Bei der Bewertung wurden insbesondere die unter ERT häufig auftretenden Infusionsreaktionen sowie schwerwiegende UEs berücksichtigt.

### 4.2.5.3 Meta-Analysen

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

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

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

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

Nicht zutreffend

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

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

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

#### 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.*

Sensitivitätsanalysen sind Analysen, in denen festgestellt wird, wie stabil sich eine Modellrechnung gegenüber Schwankungen der Eingangsparameter oder Veränderungen in der Berechnungsmethodik verhält. Dadurch kann die Robustheit der Ergebnisse eingeschätzt werden. Im statistischen Analyseplan der randomisierten, Placebo-kontrollierten Studie MOR-004 (9) waren prospektiv Sensitivitätsanalysen der Primäranalyse festgelegt.

In die Primäranalyse der Studie wurden alle randomisierten Patienten, die mindestens eine Infusion der Prüfpräparate erhalten hatten, einbezogen (ITT). Bei Patienten, die gestorben waren oder aus körperlichen Gründen unfähig waren an den Ausdauertests (6MWT und 3MSC) teilzunehmen, wurde für fehlende Werte ein Score von 0 eingegeben. Wenn nur ein Ergebnis der beiden geplanten Ausdauertests pro Untersuchungszeitpunkt vorlag, wurde der Einzelscore eingegeben, um den fehlenden Wert zu ersetzen. Hinsichtlich der respiratorischen Funktion wurde bei Patienten, die gestorben waren oder körperlich unfähig waren an den Tests teilzunehmen, unabhängig von der Gruppenzugehörigkeit und der Studienwoche, der schlechteste im Datensatz vorhandene Wert für fehlende Werte eingesetzt. Für aus anderen Gründen fehlende Werte (nicht Tod und nicht körperliches Unvermögen) wurde in der primären Analyse des 6MWT, 3MSCT, des normalisierten KS im Urin und der respiratorischen Parameter multiple Imputation eingesetzt, unter der Annahme, dass nur wenige Werte fehlen und die Werte zufällig verteilt fehlen (MAR, „missing at random“).

#### Sensitivitätsanalysen

Auswertung der Einzeltests: Um die Robustheit der Primäranalyse basierend auf den Mittelwerten der 6MWTs im Duplikat zu jedem Untersuchungszeitpunkt zu überprüfen, wurden die Tests separat mittels des primären ANCOVA-Modells analysiert (1. bzw. 2. Test). Zusätzlich wurde das individuell bessere Testergebnis der beiden Tests in das ANCOVA-Modell eingegeben.

Placebo-Mittelwert: Um die Robustheit der primären Analyse zu überprüfen, wurden fehlende Werte durch den mittleren Wert unter Placebo ersetzt (modifiziertes „worst-case“ Szenario).

Eine „worst-case“ Imputation wurde als nicht angemessen erachtet, da angenommen wurde, dass die Mehrzahl der fehlenden Daten zufällig verteilt ist (z.B. aufgrund von Entwicklungsdefiziten).

LOCF („last observed value carried forward“): Zusätzlich wurde die LOCF-Imputation für fehlende Werte eingesetzt, bei der fehlende Werte nach der Baseline-Visite durch den vorherigen Wert (*post*-Baseline) ersetzt wurden. Die Ausgangswerte wurden dabei nicht als Ersatz für fehlende Werte nach der Baseline-Visite eingegeben.

OC („observed cases“): Da wie erwartet nur wenige Datensätze fehlten, unterschied sich die OC- nicht nennenswert von der ITT-Population.

PP („per protocol“): Als zusätzliche Sensitivitätsanalyse wurde die Primäranalyse auf die PP-Population beschränkt. In Übereinstimmung mit der ICH-Richtlinie E9 stellte die PP-Population eine Untergruppe der ITT-Population dar, bei denen keine schweren Protokollverletzungen vorlagen. Zur Bestimmung der PP-Population wurden vordefinierte schwere Protokollverletzungen, die die Interpretierbarkeit der Daten infrage stellen könnten, sowie wiederholtes Auslassen der Infusionen herangezogen. Die Gründe für einen Ausschluss der Patienten wurden vor der Entblindung der Studie dokumentiert.

Evaluation der Annahmen zur Verteilung: Für den Fall, dass die Annahmen zur Verteilung (z.B. Normalverteilung, Ausreißer) verletzt war, wurde die Änderung des 6MWT bis Woche 24 in dem für die Primäranalyse eingesetzten ANCOVA-Modell ausgewertet.

Evaluation der Behandlung allein: Die Daten wurden zusätzlich in einem ANCOVA-Modell, in dem nur die Behandlung als Faktor eingegeben wurde, analysiert.

Evaluation unter Verwendung kontinuierlicher Baseline-Werte: Das ANCOVA-Modell wurde wiederholt, indem die kategoriale 6MWT-Stratifizierung ( $\leq 200$  m,  $> 200$  m) durch kontinuierliche Baseline-Werte ersetzt wurde.

Evaluation zusätzlicher Kovariaten: In weiteren Sensitivitätsanalysen wurde geprüft, ob sich das Ergebnis der Primäranalyse ändert, wenn eine zusätzliche Kovariate in dem ANCOVA-Modell enthalten war. Die Kovariaten wurden auf Basis von Daten zum natürlichen Verlauf der MPS IVA (MOR-001) ausgewählt. Folgende zusätzliche Kovariaten wurden geprüft: Interaktion zwischen den Altersgruppen und der 6MWT-Stratifizierung zum Ausgangspunkt, Geschlecht, Größe im Stehen, normalisierte Größe im Stehen (z-Score) entsprechend der CDC-Standards, 3MSCT, FVC, klinisch signifikant abnormales EKG oder Echokardiogramm, niedriges, abnormales, normalisiertes KS im Urin ( $< 19$  Jahre:  $< 20$   $\mu\text{g}/\text{mg}$ ,  $\geq 19$  Jahre:  $< 10$   $\mu\text{g}/\text{mg}$ ) und Domänen im MPS HAQ.

Evaluation der Heterogenität des Behandlungseffekts: Um die Heterogenität des Behandlungseffekts zu analysieren, wurde die Interaktion zwischen der Behandlung und der 6MWT Stratifizierung ( $\leq 200$  m,  $> 200$  m; Baseline) als zusätzliche Kovariate in das primäre ANCOVA-Modell eingegeben. Der Behandlungseffekt und die 95% KIs wurden für jedes der

beiden 6MWT-Strata berechnet. Es wird angemerkt, dass es sich bei dieser Analyse eher um eine Subgruppen- als eine Sensitivitätsanalyse handelt.

Evaluation von Ausreißern: Entsprechend des statistischen Analyseplans wurde mittels graphischer Methoden für jeden Endpunkt geprüft, ob möglicher Ausreißer vorliegen. Für jeden Endpunkt wurde der Baseline-Wert vs. die studentisierten Residuen dargestellt und die primäre Analyse ohne die größten Ausreißer wiederholt. Nach Öffnen der Datenbank wurde ein studentisiertes Residual von  $>4$  gewählt.

#### **Unterstützende Analysen:**

ANCOVA mit Messwiederholung: Als unterstützende Analyse wurde unter der Annahme einer zufälligen Verteilung fehlender Werte ein ANCOVA-Modell mit Messwiederholung eingesetzt. Bei der Analyse wurden die Behandlung, der Zeitpunkt (Woche der Visite), die Interaktion zwischen Behandlung und Zeitpunkt, die Altersklassen (5-11, 12-18,  $\geq 19$  Jahre) und die 6MWT-Stratifizierung zum Ausgangspunkt ( $\leq 200$  m,  $>200$  m) als Faktoren berücksichtigt. Die Analyse entsprach der primären Analyse mit einer Ausnahme: Fehlende Werte von Patienten, die während der Studie verstarben oder die physisch nicht in der Lage waren an dem Test teilzunehmen, wurden nicht ersetzt.

Die unterstützenden Analysen und Sensitivitätsanalysen des 3MSCT, des normalisierten KS im Urin und MVV entsprachen denjenigen des 6MWT, mit Ausnahme einer zusätzlichen Kovariate des jeweiligen Ausgangswerts. Darüber hinaus wurde der Parameter MVV nicht mittels Varianzanalyse mit Messwiederholung ausgewertet und für diesen Parameter keine LOCF-Imputation eingesetzt, da nach der Baseline-Visite bis Woche 24 keine weiteren Werte erhoben wurden. In das primäre ANCOVA-Modell zur Analyse der Heterogenität des Behandlungseffekts wurde ein zusätzlicher Interaktions-Term eingegeben (Behandlung - Baseline 3MSC, Behandlung - Baseline normalisiertes uKS bzw. Behandlung - Baseline MVV).

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

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

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

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

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

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

In der pivotalen MOR-004-Studie wurden Subgruppenanalysen im statistischen Analyseplan *a priori* festgelegt. In den Subgruppenanalysen wurde das für die Primäranalyse genutzte ANCOVA-Modell mit zusätzlichen Interaktions-Terms eingesetzt und die Effektschätzer mit den zugehörigen 95% KIs wurden innerhalb der Subgruppen kalkuliert. Folgende Subgruppen wurden hinsichtlich der Ergebnisse des 6MWT, 3MSCT, des normalisierten KS im Urin sowie des MVV analysiert:

- 6MWT Kategorien ( $\leq 200$  m und  $> 200$  m, Screening)
- Altersgruppe bei der Baseline-Visite (5-11, 12-18,  $\geq 19$  Jahre)
- Geschlecht (weiblich vs. männlich)
- Rasse (weiß vs. nicht weiß)
- Region (Nordamerika, Europa, Andere)

#### **4.2.5.6 Indirekte Vergleiche**

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Nicht adjustierte indirekte Vergleiche (d. h. Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator) stellen dabei keine valide Analysemethode dar, der Einsatz einfacher adjustierter

indirekter Vergleiche ist möglich<sup>7</sup>. Komplexe Verfahren für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche werden in der Literatur unterschiedlich bezeichnet, z. B. als „Mixed-Treatment-Comparison(MTC)-Meta-Analysen“<sup>8</sup>, „Multiple-Treatment-Meta-Analysen“<sup>9</sup> oder auch „Netzwerk-Meta-Analysen“<sup>10</sup>, sie gehen aber im Prinzip von denselben wesentlichen Annahmen aus.

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

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

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

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

- *Benennung des Brückenkomparators bzw. der Brückenkomparatoren und Begründung für die Auswahl.*
- *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.*

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

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

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

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

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

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

- *Art der Prüfung der Konsistenz zwischen den Ergebnissen direkter und indirekter Vergleiche.*
- *Bilden Sie den Code des Computerprogramms in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Nicht zutreffend

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

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

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

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 0)
- Weitere Untersuchungen (Abschnitt 0)

#### 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 4-6: 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
MOR-004	ja	ja	abgeschlossen	27 Wochen (24 Behandlungswochen)	BMN 110 2.0 mg/kg pro Woche 2.0 mg/kg jede zweite Woche Placebo

*Geben Sie an, welchen Stand die Information in Tabelle 4-6 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.*

Die Tabelle 4-6 spiegelt den Stand der Information der Studien mit BMN 110 vom 10.02.2014 wider.

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

Nicht zutreffend

#### **4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche**

*Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.*

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

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

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

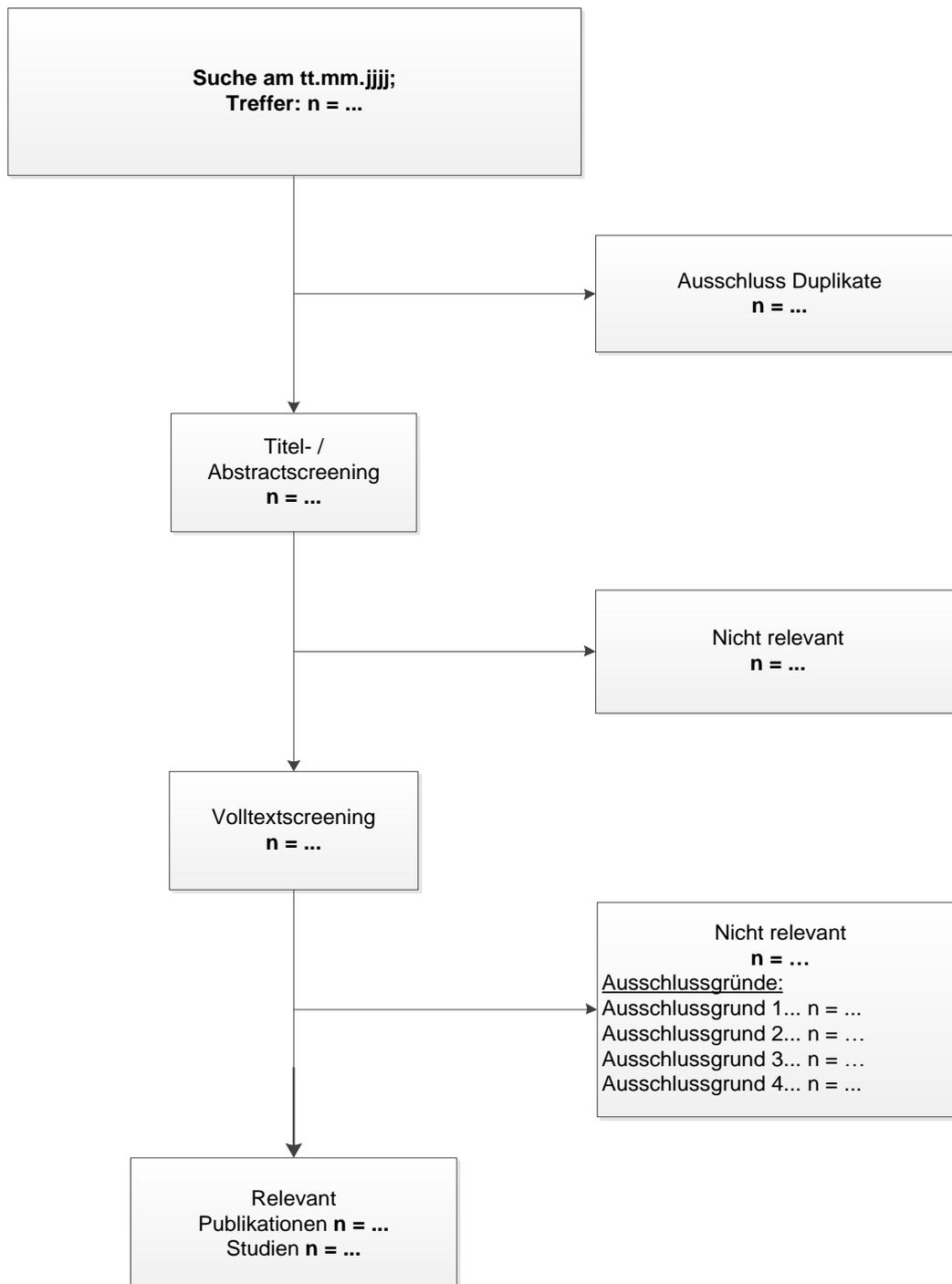


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

Für ein Arzneimittel zur Behandlung seltener Leiden sind die Zulassungsstudien relevant. Für das zu bewertende Arzneimittel wurde im Zeitraum von 12 Monaten vor dem Zeitpunkt der Dossiererstellung keine Zulassung in dem für die Nutzenbewertung relevanten Anwendungsgebiet erteilt. Eine bibliografische Literaturrecherche war nicht erforderlich.

#### 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 4-6) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

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

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

Tabelle 4-7: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters und Angabe der Zitate <sup>a</sup> )	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja / nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja / nein)	Status (abgeschlossen / abgebrochen / laufend)
MOR-004	clinicaltrials.gov: NCT01275066, Registereintrag (39) ICTRP: EUCTR2010- 020198-18-GB, Registereintrag (40)	ja	nein	abgeschlossen
a: Zitat des Studienregistereintrags sowie, falls vorhanden, der im Studienregister aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.				

Geben Sie an, welchen Stand die Information in Tabelle 4-7 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Tabelle 4-7 spiegelt den Stand der Information der Studien mit BMN 110 vom 10.02.2014 wider.

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

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2 und 4.3.1.1.3) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich der verfügbaren Quellen in Abschnitt 4.6 auf. Alle durch die vorhergehenden

Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

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

Tabelle 4-8: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
<b>ggf. Zwischenüberschrift zur Strukturierung des Studienpools</b>						
<b>placebokontrolliert</b>						
MOR-004	ja	ja	nein	ja (9)	ja (39,40)	nein
<b>aktivkontrolliert, zweckmäßige Vergleichstherapie(n)</b>						
Entfällt						
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.7 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.6 (Liste der eingeschlossenen Studien) aufgeführt werden.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p>						

#### 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.

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

Tabelle 4-9: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

<b>Studie</b>	<b>Studiendesign</b> <RCT, doppelblind/einfach verblindet/offen, parallel/cross-over etc.>	<b>Population</b> <relevante Charakteristika, z. B. Schweregrad>	<b>Interventionen</b> (Zahl der randomisierten Patienten)	<b>Studiendauer</b> <ggf. Run-in, Behandlung, Nachbeobachtung>	<b>Ort und Zeitraum der Durchführung</b>	<b>Primärer Endpunkt; patientenrelevante sekundäre Endpunkte</b>
MOR-004	RCT, doppelblind, parallel, multi- zentrisch, multi- national	Patienten mit der klinischen Diagnose MPS IV A auf Basis der klinischen Zeichen und Symptome und dokumentierte reduzierte Fibroblasten- oder Leukozyten-GALNS, ≥5 Jahre, 6MWT ≥30 und <325 m	BMN 110 2.0 mg/kg KG/Woche i.v. (58) BMN 110 2.0 mg/kg KG i.v. alternierend mit Placebo jede zweite Woche (59) Placebo (59)	Bis zu 3 Wochen Screening, 24 Wochen Doppel- blindphase	33 Studienzentren in 17 Ländern (Argentinien, Bra- silien, Kanada, Kolumbien, Däne- mark, Frankreich, Deutschland, Italien, Japan, Portugal, Katar, Saudi Arabien, Südkorea, Taiwan, Niederlande, UK, USA) 01.02.2011 - 16.08.2012	<u>Primärer Endpunkt:</u> 6MWT (Gehstrecke in Metern) <u>Sekundäre Endpunkte:</u> 3MSCT (Stufen/min) Kreatinin-norma- lisierte KS-Spiegel im Urin jeweils Änderung vom Ausgangswert bis Woche 24

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

Studie	BMN 110 2.0 mg/kg/ Woche	BMN 110 2.0 mg/kg qow	Placebo	ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in- Phase etc.
MOR-004	BMN 110 2.0 mg/kg KG/ Woche verdünnt in 0.9% NaCl ( $\leq 250$ ml) i.v. (Infusion) über etwa 4 Stunden (abhängig vom Verdünnungsvolumen) <sup>1</sup> im Abstand von mindestens 4 d (7 $\pm$ 3 d)	BMN 110 2.0 mg/kg KG verdünnt in 0.9% NaCl jede zweite Woche alternierend mit Placebo, analog BMN 110 2.0 mg/kg/Woche	Placebo-Lösung einmal pro Woche (Vehikel verdünnt in 0.9% NaCl), analog BMN 110 2.0 mg/kg/Woche	% Prämedikation <sup>2</sup> vor den Infusionen (>5% der Patienten; Placebo / BMN 110 2.0 mg/kg qow / pro Woche): Gesamt: 100.0 / 100.0 / 98.3 Anilide: 45.8 / 28.8 / 56.9 Andere Antihistamine zur systemischen Anwendung: 47.5 / 50.8 / 48.3 Glukokortikoide: 11.9 / 22.0 / 36.2 H2-Rezeptorantagonisten: 6.8 / 23.7 / 24.1 Piperazinderivate: 27.1 / 25.4 / 24.1 Aminoalkylether: 10.2 / 15.3 / 22.4 Substituierte Alkylamine: 10.2 / 16.9 / 15.5 Diphenylmethanderivate: 5.1 / 5.1 / 8.6 Begleitmedikationen siehe Fussnote <sup>3</sup>

<sup>1</sup> Um das Risiko von Infusionsreaktionen möglichst gering zu halten, wurde die Infusion mit niedriger Geschwindigkeit begonnen und die Infusionsgeschwindigkeit graduell alle 15 min bis zur maximalen Infusionsgeschwindigkeit gesteigert. Nach Verdünnung in 100 ml Kochsalzlösung betrug die anfängliche Infusionsrate 3 ml/h und wurde alle 15 min um 6 ml/h bis zum Maximum von 36 ml/h gesteigert. Die anfängliche Infusionsrate betrug 6 ml/h nach Verdünnung in 250 ml Kochsalzlösung und wurde alle 15 min um 12 ml/h auf den Maximalwert von 72 ml/h gesteigert. Das Infusionsschema konnte modifiziert werden (z.B. längere Infusion), um das Risiko von Infusionsreaktionen bei einzelnen Patienten zu minimieren.

<sup>2</sup> Antihistaminika (bevorzugt nicht sedierend) mussten vor jeder Infusion verabreicht werden. Abhängig von den Risikofaktoren (z.B. bekannte IARs oder Allergien) konnten zusätzlicher Medikamente wie H2-Rezeptorantagonisten, Montekulast-Natrium oder Steroide vor den Infusionen gegeben werden. Die Gabe von Antipyretika vor den Infusionen lag im Ermessen des behandelnden Arztes.

<sup>3</sup> Die häufigsten Begleitmedikationen (>30% in einer der Gruppen) waren Anilide, andere Antihistaminika zur systemischen Anwendung, Glukokortikoide, Propionsäurederivative, Aminoalkylether und Piperazinderivate. Der Anteil der Patienten, die Aminoalkylether (13.6%, 23.7%, 32.8%) und H2-Rezeptorantagonisten (6.8%, 25.4%, 29.3%) als Begleitmedikation anwendeten, war in der Placebogruppe niedriger als in der BMN 110 2.0 mg/kg/qow bzw. BMN 110 2.0 mg/kg/Woche Gruppe. Demgegenüber wandten mehr Patienten der Placebogruppe als der BMN 110-Gruppen Amide (15.3%, 5.1%, 8.6%) und Benzodiazepine (15.3%, 3.4%, 6.9%) als Begleitmedikation an.

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

Tabelle 4-11: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel

Studie Gruppe	N	Alter (Jahre) Mittelwert (SD)	Geschlecht w / m (%)	Altersgruppen in Jahren (%)			Rasse (%)			
				5-11 J.	12-18 J.	≥19 J.	Asiatisch	Schwarz, Afroame- rikaner	Weiß	Andere
MOR-004										
Placebo	59	15.0 (11.30)	54.2 / 45.8	50.8	25.4	23.7	18.6	0	74.6	6.8
BMN 110 qow	59	15.3 (10.79)	42.4 / 57.6	52.5	27.1	20.3	25.4	3.4	59.3	11.9
BMN 110 pro Woche	58	13.1 (8.10)	55.2 / 44.8	55.2	27.6	17.2	24.1	3.4	62.1	10.3

Studie Gruppe	N	6MWT Gehstrecke (m) Mittelwert (SD)	Stratifizierte 6MWT Gehstrecke in Metern (%)		Benutzung einer Gehhilfe während des 6MWT (%)	3MSCT (Stufen/min) Mittelwert (SD)	uKS (µg/ml) Mittelwert (SD)
			≤200 m	>200 m			
MOR-004							
Placebo	59	211.9 (69.88)	39.0	61.0	18.6%	30.0 (14.05)	25.7 (15.09)
BMN 110 qow	59	205.7 (81.19)	40.7	59.3	27.1%	27.1 (15.80)	28.6 (21.17)
BMN 110 pro Woche	58	203.9 (76.32)	39.7	60.3	15.5%	29.6 (16.44)	26.9 (14.11)

*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.*

Bei der Zulassungsstudie MOR-004 handelt es sich um eine doppelblinde, randomisierte, multizentrische, Placebo-kontrollierte Parallelgruppen-Studie über 24 Wochen bei Patienten mit MPS IVA. Das primäre Ziel der Studie war der Vergleich der Wirksamkeit von BMN 110 in zwei Dosierungen (2.0 mg/kg KG einmal wöchentlich oder alle zwei Wochen) und Placebo hinsichtlich der Ausdauer im 6MWT (Gehstrecke) nach 24 Wochen Behandlungsdauer. Die vorliegende Bewertung bezieht sich auf die Gesamtpopulation der eingeschlossenen Patienten.

Patienten im Alter von mindestens fünf Jahren mit der klinischen Diagnose MPS IVA und einer verminderten GALNS-Aktivität in Fibroblasten oder Leukozyten oder einer bestätigten Diagnose mittels genetischer Tests konnten an der Studie teilnehmen. Patienten mit einer Gehstrecke zwischen 30 und 325 m im 6MWT vor Beginn der Behandlung wurden in die Studie eingeschlossen. Die Studie umfasste eine bis zu 3-wöchige Screening-/Baseline-Phase und eine 24-wöchige doppelblinde Behandlungsphase. Die Studiendauer war auf 24 Wochen limitiert, da größere chirurgische Eingriffe während der Studie nicht möglich waren, und daher aus ethischen Gründen eine längere Studiendauer nicht vertretbar war.

Die Patienten wurden randomisiert in einem Verhältnis von 1:1:1 der Behandlung mit BMN 110 2.0 mg/kg KG einmal wöchentlich oder jede zweite Woche (alternierend mit Placebo, jeweils verdünnt in 0.9% NaCl) oder Placebo einmal wöchentlich zugeteilt. Die Prüfpräparate wurden jeweils über etwa 4 h unter enger Überwachung der Patienten infundiert. Die Infusionsrate wurde sukzessive alle 15 min erhöht, um das Risiko möglicher Überempfindlichkeits- oder Infusionsreaktionen zu minimieren. Der primäre Endpunkt war die Änderung der Gehstrecke (in Metern) im 6MWT von der Baseline-Visite bis Woche 24. Sekundäre Endpunkte waren die Ausdauer im 3MSC-Test (Stufen/min) und die normalisierte KS-Ausscheidung im Urin. Um die Variabilität der Ausdauertests (6MWT und 3MSC) möglichst niedrig zu halten, wurde ein Qualitätssicherungsprogramm, welches Trainings vor Ort und eine Zertifizierung der Zentren einschloss, implementiert. Die primäre Wirksamkeitsanalyse bezog sich auf die ITT-Population. Um die Robustheit der Primäranalyse zu bestätigen, wurden mehrere Sensitivitätsanalysen ausgeführt. Protokollverletzungen wurden während der Doppelblindphase entsprechend einer im Studienprotokoll festgelegten Richtlinie zu Protokollverletzungen als geringfügig oder bedeutend eingestuft.

Die Sicherheitsvariablen umfassten u.a. die Dokumentation aller UEs, Standard-Labortests, Vitalzeichen, physische Untersuchungen, EKGs und Immunogenitätstest, um BMN 110-spezifische Antikörper zu detektieren. Hypersensitivitätsreaktionen und mit der Infusion assoziierte Reaktionen (IAR, UEs während und innerhalb eines Tages nach der Infusion unabhängig vom Kausalzusammenhang), die nicht auf allergische Reaktionen beschränkt waren, waren als UEs besonderen Interesses vordefiniert. Im Falle schwerer Infusions-

reaktionen während der Studie stand ein unabhängiges Expertengremium („Allergic Reaction Review Board“, ARRB) zur Verfügung, deren Mitglieder nicht direkt an der Studie beteiligt waren. Ein unabhängiges „Data Monitoring Committee“ (DMC) überwachte und bewertete in regelmäßigen Abständen den Studienverlauf sowie die während der Studie aufgetretenen UEs auf Basis von Berichten eines unabhängigen Statistikers und konnte falls erforderlich Empfehlungen hinsichtlich des weiteren Verlaufs der Studie abgeben.

Insgesamt 177 Patienten wurden den Behandlungsgruppen randomisiert zuteilt (ITT: n=176, n=59 Placebo und BMN 110 2.0 mg/kg KG jede 2. Woche, n=58 BMN 110 2.0 mg/kg KG/Woche). Ein Patient der Placebogruppe, bei dem die Diagnose MPS IVA nicht bestätigt werden konnte und der keine Infusion erhalten hatte, wurde in Übereinstimmung mit dem statistischen Analyseplan aus allen Analysen ausgeschlossen. Somit entsprach die ITT- der Sicherheitspopulation. Insgesamt 175 der 176 Patienten beendeten die Studie (99.4%) und nur ein Patient (0.6%) brach die Studie ab (Widerruf der Einwilligungserklärung).

Die Patientengruppen waren hinsichtlich des Alters (5-57 Jahre), der Verteilung der Altersgruppen und der Geschlechtsverteilung zu Beginn der Studie vergleichbar (Tabelle 4-11). In allen Behandlungsgruppen war etwa die Hälfte der Patienten zwischen fünf und 11 Jahre alt. Die meisten Patienten waren weiß, und in der Placebogruppe (74.6%) waren etwas mehr weiße Patienten vertreten als in der BMN 110 2.0 mg/kg KG qow (59.3%) und der BMN 110 2.0 mg/kg KG/Woche Gruppe (62.1%). Der Anteil der hispanischen Patienten und Latinos war in der BMN 110 2.0 mg/kg/qow Gruppe (27.1%) etwas höher als in den Placebo (22.0%) und BMN 110 2.0 mg/kg/Wochen (15.5%) Gruppen. Die Mehrzahl der Patienten in der Placebo- und den BMN 110-Gruppen stammte aus Europa (45.8%, 35.6%, 43.1%) und Nordamerika (27.1%, 27.1%, 25.9%). Die Gehstrecke im 6MWT zu Beginn der Studie unterschied sich zwischen den Gruppen nicht nennenswert und zeigte einen geringen Vorteil zugunsten der Placebogruppe (siehe Tabelle 4-11). Der Anteil der Patienten, die  $\leq 200$  m bzw. mehr als 200 m gehen konnten, war in den Gruppen vergleichbar. Der Anteil der Patienten, die zu Beginn der Studie im 6MWT eine Gehhilfe benutzten, war in der BMN 110 2.0 mg/kg/qow Gruppe (27.1%) etwas höher als in den Placebo (18.6%) und BMN 110 2.0 mg/kg/Woche (15.5%) Gruppen. In der Analyse des MPS HAQ (Subscore zur Mobilität) gaben zum Ausgangspunkt deutlich mehr Patienten oder deren Betreuer an, dass die Patienten einen Rollstuhl benutzen: Placebo 37.3%, BMN 110 qow 57.6%, BMN 110 jede Woche 51.7% (siehe Tabelle 4-46). Dieses Ergebnis deutet an, dass die Patienten der BMN 110-Gruppen in ihrer Mobilität mehr eingeschränkt waren als die Patienten der Placebogruppe. Der Anteil der Patienten/Betreuer, die zu Beginn der Studie angaben, eine Gehilfe zu benutzen (MPS HAQ), betrug 30.5% in der Placebogruppe, 37.3% in der BMN 110 qow Gruppe und 29.3% bei den Patienten, die BMN 110 jede Woche erhielten. Die Ergebnisse des 3MSC-Tests sowie des uKS waren zwischen den Behandlungsgruppen zu Beginn der Studie vergleichbar.

Die Zahl und Verteilung medizinischer Vorbefunde waren in den Gruppen insgesamt ähnlich und spiegelten die hohe Morbidität vom MPS IVA wider. Verglichen mit der Placebo- und der BMN 110 2.0 mg/kg/qow Gruppe war die Inzidenz immunologischer Vorerkrankungen

(16.9%, 27.1%, 5.2%), hepatobiliärer Vorerkrankungen (8.5%, 11.9%, 3.4%), und psychiatrischer Vorerkrankungen (20.3%, 15.3%, 3.4%) in der BMN 110 2.0 mg/kg/Woche Gruppe etwas niedriger.

Die mittlere Zahl der Infusionen und die mittlere Zahl der nicht abgeschlossenen Infusionen waren zwischen den Gruppen vergleichbar und in allen Gruppen war die mittlere Dosiscompliance vergleichbar hoch (96.8% bis 99.2%). Die Zahl der nicht wie geplant verabreichten Infusionen (und der Prozentsatz der nicht verabreichten relativ zu den geplanten Infusion) war niedrig, allerdings in den BMN 110 2.0 mg/kg/qow (n=16, 1.1%) und 2.0 mg/kg/Woche Gruppen (n=25, 1.8%) geringfügig höher als in der Placebogruppe (n=9, 0.6%). Ein gleichsinniger Trend ergab sich hinsichtlich des Anteils der Patienten, bei denen die Infusionen nicht wie geplant verabreicht werden konnten (13.6%, 18.6% bzw. 31.0%). Die Gründe, warum Infusionen nicht verabreicht wurden, waren sehr unterschiedlich (z.B. Begleiterkrankungen, UEs, kein Venenzugang oder persönliche Gründe wie Transportprobleme) und es ergab sich kein klares Muster. Alle Patienten, die eine Infusion nicht wie geplant erhalten hatten, erhielten die nachfolgenden Infusionen wie vorgesehen, außer es handelte sich um die letzte Infusion.

Da die Anzahl der in die Studie eingeschlossenen Patienten angesichts der Seltenheit der Erkrankung als hoch einzuschätzen ist, wird das Patientenkollektiv als für die Zielpopulation repräsentativ eingestuft. Ein großer Teil der eingeschlossenen Patienten stammte aus Europa inklusive Deutschland, sodass angenommen wird, dass die Ergebnisse auf den deutschen Versorgungskontext übertragbar sind. Zudem ergaben die Subgruppenanalysen stratifiziert nach Herkunft (Region) keine Hinweise auf signifikante Interaktionen (siehe Abschnitt 4.3.1.3.10).

#### 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 4-12: 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			
MOR-004	ja	ja	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential wird auf Studienebene wird unter Berücksichtigung der in Tabelle 4-12 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Bei der Studie MOR-004 handelt es sich um eine multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte internationale Parallelgruppenstudien der Phase 3 bei Patienten mit MPS IVA und einer Gehstrecke zwischen 30 und 325 m im 6MWT. Eine adäquate Erzeugung der Randomisierungssequenz liegt vor. Die Gruppenzuteilung erfolgte verdeckt mit Hilfe von Randomisierungslisten, die durch unabhängige Dritte erstellt wurden. Sowohl die Patienten als auch die Prüfer und das medizinische Personal waren bis zur Endauswertung der Studie verblindet (siehe auch Abschnitt 4.3.1.3.1). Das ITT-Prinzip wurde adäquat umgesetzt und die Imputationsstrategien für fehlende Werte waren angemessen. Nur ein Patient, bei dem die Diagnose MPS IVA nicht bestätigt werden konnte, wurde protokollgemäß aus der ITT-Population ausgeschlossen. Wie bei einer sehr seltenen Erkrankung zu erwarten, ist die Fallzahl relativ niedrig. Die Fallzahlschätzung beruhte allerdings auf konservativen Annahmen und unter Berücksichtigung der Seltenheit der Erkrankung ist die Zahl der eingeschlossenen Patienten als hoch 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 4-13: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	6MWT	3MSCT	uKS*	MVV FEV etc.	MPS HAQ	Anthropo- metrie	Gesamte UEs und SUEs	Überempfind- lichkeitsreak- tionen und IAR
MOR-004	ja	ja	ja	ja	ja	ja	ja	ja

\*Nicht direkt patientenrelevant (pharmakodynamisch/biochemischer Marker)

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

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

Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

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

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

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

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

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 0). 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 4-14: Operationalisierung der Gehstrecke im 6MWT

Studie	Operationalisierung
MOR-004	Der primäre Endpunkt war die Änderung der Gehstrecke im 6MWT vom Ausgangspunkt bis Woche 24 in der BMN 110 2.0 mg/kg/Woche und 2.0 mg/kg/qow Gruppe jeweils verglichen mit Placebo. Der Test wurde entsprechend der ATS-Richtlinie ausgeführt (16). Der 6MWT

wurde bei der Screening-Untersuchung, nach 12 und 24 Wochen Behandlung oder im Falle eines vorzeitigen Abbruchs der Studie erhoben. Patienten, die bei der Screening-Visite im 6MWT scheiterten ( $<30$  m), durften nicht erneut getestet werden und durften an der Studie nicht teilnehmen. Zu jedem Erhebungszeitpunkt wurde der Test zweimal an verschiedenen Tagen ausgeführt. In Woche 12 und 24, mussten die Tests innerhalb von fünf Tagen vor der Infusion stattfinden. Die beiden 6MWT-Tests sollten möglichst zur selben Tageszeit stattfinden und die Ausdauerests (6MWT und 3MSCT) durften nicht an ein und demselben Tag erfolgen. Die Ausdauerests sollten möglichst in der folgenden Reihenfolge stattfinden: 6MWT, 3MSCT, 6MWT und 3MSCT. Patienten, die 30 min oder länger gereist waren, sollten nicht an diesem Tag getestet werden. Die Patienten wurden instruiert, innerhalb von 6 min so weit wie möglich zu gehen, wobei der Wortlaut der Instruktionen vor und während des Tests, um den Patienten anzuleiten und zu motivieren, vorgegeben war. Der Patient konnte bei laufender Stoppuhr eine Pause im Stehen einlegen (gegebenenfalls angelehnt an eine Wand), falls sich der Patient allerdings setzen musste galt der Test als beendet. Die Gründe für einen vorzeitigen Abbruch des Tests waren zu dokumentieren. Die Tests fanden in einem zertifizierten 30 m langen, ebenen Innenraumflur mit harter Oberfläche bei angemessener Beleuchtung ohne sichtbare Uhr statt. In der ATS-Richtlinie zum 6MWT wird eine Korridorlänge von 30 m empfohlen. Die Gehstrecke war mit Markierungen am Startpunkt, alle 3 m sowie bei 30 m gekennzeichnet und vor Beginn des Tests sollten am Start- sowie Wendepunkt (30 m) Kegel positioniert werden. Die Tests sollten möglichst in Anwesenheit von zwei Klinikangestellten stattfinden, was allerdings keine Bedingung darstellte. Die Zeit wurde mittels einer Accusplit-Stoppuhr gemessen, die vor jedem Einsatz kalibriert werden sollte, und die Zahl der Runden (Kreuzen des Startpunkts) wurde mittels eines manuellen Rundenzählers oder anderer geeigneter Dokumentationsmöglichkeiten erfasst. Falls der Patient in der Lage war, mehr als 325 m zu gehen, sollte der Patient nicht gestoppt werden. Modifikationen des Tests (z.B. im Falle einer Sehstörung) waren nur nach Rücksprache mit dem Monitor zulässig. Falls der Patient nicht in der Lage war, 6 min zu gehen, sollten die Zeit und die Gehstrecke dokumentiert werden. Im Falle eines Sturzes sollte die Stoppuhr weiterlaufen es sei denn der Patient konnte den Test nicht beenden. Die Gehstrecke sollte auf das nächste Zehntel eines Meters gerundet werden.

Gehhilfen waren erlaubt, mussten aber während der Studie konsistent eingesetzt werden, unabhängig davon, ob während eines spezifischen Tests eine Gehhilfe benötigt wurde oder nicht. Während der Tests war nur eine Gehhilfe erlaubt. Gehhilfen sollten nach der Screening-Visite möglichst nicht eingeführt werden. Falls (neue) Gehhilfen während der Studie benötigt wurden, sollten während der Tests möglichst keine Gehhilfen eingesetzt werden unter der Voraussetzung, dass die Sicherheit der Patienten gewährleistet war. Die Örtlichkeit der Tests sowie die verwendete Stoppuhr durften während der Studie nicht geändert werden. Falls ein Patient an dem geplanten Testtag zu müde war, konnte der Test auf einen anderen Tag verschoben werden. Relative Kontraindikationen für den 6MWT waren ein Ruhepuls  $>120$  bpm, ein systolischer Blutdruck  $>180$  mmHg oder ein diastolischer Blutdruck  $>100$  mmHg. Der Puls und die Sauerstoffsättigung wurden vor dem Test und unmittelbar sowie 2 min nach Beendigung des Tests überwacht. Der Test sollte u.a. bei Vorliegen von Schmerzen in der Brust, nicht tolerabler Dyspnoe, Beinkrämpfen, Schwanken, Diaphoresis oder blassem/fahlem Erscheinungsbild unmittelbar abgebrochen werden.

Biomarin hat ein Qualitätssicherungsprogramm etabliert, um die Variabilität zu minimieren und sicher zu stellen, dass die beteiligten Personen in jedem Studienzentrum ein angemessenes und einheitliches Training erhalten haben und zur Ausführung der Ausdauerests (6MWT und 3MSC) zertifiziert sind. Das Programm schloss ein erstes Training während eines Prüfarzt-treffens sowie ein Qualitätssicherungsprogramm in den Studienzentren ein, um die Testausführung zu überprüfen und das für die Tests verantwortliche Personal zu zertifizieren. In den meisten Studienzentren fand das vor Ort Training vor Einschluss des ersten Patienten und erneut nach Einschluss von etwa 2-3 Patienten statt. Im Falle eines Wechsels des Personals oder auf Wunsch des Prüfarztes waren ein weiteres Training und eine Re-Zertifizierung vorgesehen.

In die statistische Analyse wurde jeweils der Mittelwert beider Tests pro Erhebungszeitpunkt einbezogen oder eines Einzeltests, wenn nur ein Testergebnis vorlag. Die Änderung der Gehstrecke (in Metern) im 6MWT in Woche 24 verglichen mit Baseline wurde mittels ANCOVA mit den Faktoren Behandlung, Altersklasse (5-11, 12-18,  $\geq 19$  Jahre) und 6MWT-Kategorien bei der Baseline-Visite ( $\leq 200$  m und  $>200$  m) analysiert (primäre Analyse). Um

einen Typ I-Fehler von 0.05 einzuhalten, wurde die Multiplizität nach Hochberg adjustiert. Das Ergebnis der Studie wurde als positiv bewertet, wenn der Vergleich beider BMN 110-Dosisgruppen mit Placebo einen p-Wert von 0.05 oder der Vergleich einer der BMN 110-Dosisgruppen mit Placebo einen p-Wert von 0.025 ergab (t-Test im ANCOVA-Modell).

In die Wirksamkeitsanalyse wurden alle randomisierten Patienten, die mindestens eine Dosis der Prüfpräparate erhalten hatten, eingeschlossen (ITT). In den Auswertungen der Ausdauerests (6MWT und 3MSCT) wurde ein Score von 0 eingegeben, wenn der Patient gestorben war oder physisch nicht in der Lage war, an dem Test teilzunehmen. In der primären Analyse wurde multiple Imputation eingesetzt, wenn Werte aus anderen Gründen fehlten. Es wurde angenommen, dass die Werte zufällig fehlten (MAR), da erwartet wurde, dass nur wenige Patienten vorzeitig aus der Studie ausscheiden und nur wenige Daten zeitweise fehlen. Folgende kontinuierliche Variablen wurden berücksichtigt:

- 6MWT: Alter (Baseline) und 6MWT in Woche 0, 12 und 24
- 3MSC-Test: Alter (Baseline und 6MWT in Woche 0 und 3MSCT in Woche 0, 12 und 24
- KS im Urin: Alter (Baseline) und 6MWT in Woche 0 und KS im Urin in Woche 0, 2, 4, 8, 12, 16, 20 und 24
- RFTs: Alter (Baseline) und 6MWT in Woche 0 und RFT in Woche 0 und 24

Für die Analyse wurde das Programm SAS PROC MI eingesetzt. Die Imputationen wurden für jede Behandlungsgruppe separat unter Verwendung des randomisiert ausgewählten „seed“ 38867 ausgeführt. In der Analyse wurden die adjustierten Freiheitsgrade äquivalent zur Zahl der Patienten in der Analyse abzüglich der Freiheitsgrade der Kovariaten (inklusive Achsenabschnitt) eingesetzt.

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

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

Tabelle 4-15: Bewertung des Verzerrungspotenzials für 6MWT 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
MOR-004	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt 6MWT wird unter Berücksichtigung der in Tabelle 4-15 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 Prüferärzte, das medizinische Personal und die Patienten waren bis zur finalen Schließung der klinischen Datenbank verblindet. Die höhere Inzidenz von Überempfindlichkeitsreaktionen und IARs in den BMN 110-Gruppen, die zu einer Unterbrechung der Infusionen führten, könnte zu einer Entblindung einzelner Patienten seitens der Behandler geführt haben, worauf auch die in der Placebogruppe verglichen mit den BMN 110-Gruppen (2.0 mg/kg jede zweite Woche bzw. pro Woche) unterschiedliche Häufigkeit, mit der z.B. Glukokortikoide (11.9%, 22.0% bzw. 36.2%) und H2-Rezeptorantagonisten (6.8%, 23.7% bzw. 24.1%) vor den Infusionen verabreicht wurden, hindeuten könnte. Allerdings unterschied sich die Gesamtzahl der IARs im Gruppenvergleich nicht (Tabelle 4-62). Da der 6MWT unter standardisierten Bedingungen nach einem zuvor festgelegten Qualifizierungsprogramm, das eine Zertifizierung der Zentren voraussetzte, ausgeführt wurde, wird nicht angenommen, dass eine partielle Entblindung falls zutreffend die Ergebnisse relevant verzerrt haben könnte. Eine adäquate Umsetzung des ITT Prinzips fand statt. Alle Patienten, die mindestens eine Dosis der Prüfpräparate erhalten hatten, wurden in die ITT-Analyse eingeschlossen. Nur ein Patient, der nicht behandelt worden war und bei dem die Diagnose MPS IVA nicht bestätigt werden konnte, wurde protokollgemäß aus der ITT-Population ausgeschlossen. Fehlende Werte wurden mittels multipler Imputation ersetzt und die Robustheit der Primäranalyse durch verschiedene Sensitivitätsanalysen bestätigt. Hinweise auf eine ergebnisgesteuerte Berichterstattung oder das Vorhandensein von sonstigen Faktoren, die das Verzerrungspotential beeinflussen könnten, lagen 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 4-16: Ergebnisse für 6MWT (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik)

<b>MOR 004</b>	<b>n</b>	<b>6MWT Baseline (m) Mittelwert (SD)</b>	<b>6MWT 24 Wochen (m) Mittelwert (SD)</b>	<b>Änderung von Baseline bis Woche 24 (m) Mittelwert (SD)</b>	<b>Änderung von Baseline bis Woche 24 (%) Mittelwert (SD)</b>
Placebo	59	211.9 (69.88)	225.4 (83.22)	13.5 (50.63)	8.7 (28.83)
BMN 110 qow	59	205.7 (81.19)	219.9 (87.60)	14.2 (40.82)	10.2 (26.42)
BMN 110 pro Woche	58	203.9 (76.32)	240.0 (86.61)	36.0 (58.11)	23.8 (44.43)

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

Tabelle 4-17: Ergebnisse für 6MWT (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (Primär- und Sensitivitätsanalysen)

<b>MOR-004</b>	<b>N</b>	<b>6MWT Baseline (m) LSM (95% KI)</b>	<b>6MWT 24 Wochen (m) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test<sup>3</sup>) BMN 110 vs. Placebo</b>
<b>ANCOVA (ITT, Primäranalyse)<sup>1</sup></b>						
Placebo	59	213.4 (194.8, 232.0)	209.5 (193.1, 226.0)	13.6 (0.6, 26.5)	-	-
BMN 110 qow	59	205.7 (187.2, 224.3)	205.4 (188.9, 221.9)	14.1 (1.1, 27.1)	0.5 (-17.8, 18.9)	0.9542
BMN 110 pro Woche	58	202.3 (183.6, 221.0)	223.6 (206.8, 240.5)	36.0 (22.9, 49.1)	22.5 (4.0, 40.9)	0.0174
<b>ANCOVA mit Messwiederholung (ITT)<sup>2</sup></b>						
Placebo	59	213.4 (194.8, 232.0)	225.1 (209.4, 240.9)	13.4 (0.6, 26.2)	-	-
BMN 110 qow	59	205.7 (187.2, 224.3)	221.2 (205.4, 237.1)	14.2 (1.3, 27.0)	0.8 (-17.3, 18.9)	0.9329
BMN 110 pro Woche	58	202.3 (183.6, 221.0)	240.0 (224.0, 255.9)	36.5 (23.5, 49.5)	23.1 (4.9, 41.3)	0.0132
<b>ANCOVA: Gepoolter mittlerer Placebowert (ITT)<sup>1</sup></b>						
Placebo	59	213.4 (194.8, 232.0)	225.1 (209.1, 241.1)	13.6 (0.6, 26.5)	-	-
BMN 110 qow	59	205.7 (187.2, 224.3)	221.6 (205.6, 237.6)	14.8 (1.8, 27.7)	1.2 (-17.1, 19.5)	0.8975
BMN 110 pro Woche	58	202.3 (183.6, 221.0)	239.1 (223.0, 255.2)	35.9 (22.9, 49.0)	22.4 (4.0, 40.8)	0.0176
<b>ANCOVA: LOCF-Imputation (ITT)<sup>1</sup></b>						
Placebo	59	213.4 (194.8, 232.0)	225.1 (209.0, 241.2)	13.6 (0.6, 26.6)	-	-
BMN 110 qow	59	205.7 (187.2, 224.3)	220.8 (204.8, 236.9)	14.0 (1.1, 27.0)	0.4 (-17.9, 18.8)	0.9625
BMN 110 pro Woche	58	202.3 (183.6, 221.0)	239.2 (223.0, 255.4)	36.0 (22.9, 49.1)	22.4 (4.0, 40.9)	0.0174

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

<b>MOR-004</b>	<b>N</b>	<b>6MWT Baseline (m) LSM (95% KI)</b>	<b>6MWT 24 Wochen (m) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test<sup>3</sup>) BMN 110 vs. Placebo</b>
<b>ANCOVA mit Behandlung als einziger Kovariate (ITT)</b>						
Placebo	59	211.9 (192.4, 231.4)	225.4 (203.3, 247.4)	13.5 (0.6, 26.4)	-	-
BMN 110 qow	59	205.7 (186.2, 225.2)	219.9 (197.8, 241.9)	14.2 (1.3, 27.1)	0.7 (-17.6, 19.0)	0.9394
BMN 110 pro Woche	58	203.9 (184.3, 223.6)	240.0 (217.7, 262.2)	36.0 (23.0, 49.1)	22.5 (4.2, 40.9)	0.0166
<b>ANCOVA: Nur Erster 6MWT (ITT)<sup>1</sup></b>						
Placebo	59	215.3 (195.8, 234.8)	210.3 (193.3, 227.3)	12.5 (-1.8, 26.8)	-	-
BMN 110 qow	59	204.6 (185.2, 224.0)	208.5 (191.5, 225.6)	18.5 (4.2, 32.7)	6.0 (-14.2, 26.2)	0.5607
BMN 110 pro Woche	58	204.8 (185.2, 224.4)	225.6 (208.2, 243.0)	35.5 (21.1, 49.9)	23.0 (2.7, 43.3)	0.0266
<b>ANCOVA: Nur zweiter 6MWT (ITT)<sup>1</sup></b>						
Placebo	59	211.6 (192.6, 230.5)	208.8 (191.4, 226.1)	14.7 (-0.4, 29.7)	-	-
BMN 110 qow	59	206.9 (188.0, 225.8)	202.9 (184.9, 220.8)	10.5 (-5.3, 26.2)	-4.2 (-26.0, 17.6)	0.7033
BMN 110 pro Woche	58	199.8 (180.7, 218.9)	221.8 (204.0, 239.5)	36.7 (21.5, 51.9)	22.0 (0.6, 43.4)	0.0435
<b>ANCOVA: Bester 6MWT (ITT)<sup>1</sup></b>						
Placebo	59	226.6 (207.2, 246.1)	222.9 (205.5, 240.2)	14.0 (0.1, 27.9)	-	-
BMN 110 qow	59	220.5 (201.0, 239.9)	216.8 (199.4, 234.3)	11.1 (-2.8, 25.0)	-2.9 (-22.6, 16.7)	0.7686
BMN 110 pro Woche	58	214.0 (194.3, 233.6)	234.1 (216.3, 251.9)	35.1 (21.0, 49.1)	21.1 (1.3, 40.9)	0.0371

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

MOR-004	N	6MWT Baseline (m) LSM (95% KI)	6MWT 24 Wochen (m) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (m) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (m) BMN 110 vs. Placebo LSM (95% KI)	p-Wert (t-Test <sup>3</sup> ) BMN 110 vs. Placebo
<b>ANCOVA: 6MWT als kontinuierliche Baseline-Kovariate (ITT)<sup>1</sup></b>						
Placebo	59	213.4 (194.8, 232.0)	220.3 (207.1, 233.6)	13.9 (1.0, 26.9)	-	-
BMN 110 qow	59	205.7 (187.2, 224.3)	220.5 (207.1, 233.8)	14.1 (1.1, 27.0)	0.1 (-18.2, 18.5)	0.9878
BMN 110 pro Woche	58	202.3 (183.6, 221.0)	242.1 (228.4, 255.7)	35.7 (22.6, 48.8)	21.7 (3.3, 40.2)	0.0212
<b>ANCOVA: Ausschluss eines Ausreißers (ITT)<sup>1</sup></b>						
Placebo	58	195.1 (184.5, 205.6)	206.2 (190.3, 222.2)	11.2 (-1.5, 23.8)	-	-
BMN 110 qow	59	190.6 (180.1, 201.1)	205.9 (190.0, 221.9)	15.4 (2.7, 28.0)	4.2 (-13.3, 21.7)	0.6362
BMN 110 pro Woche	58	186.9 (176.2, 197.7)	224.3 (208.1, 240.5)	37.3 (24.5, 50.2)	26.2 (8.6, 43.8)	0.0038
<b>ANCOVA: PP-Analyse<sup>1</sup></b>						
Placebo	55	209.6 (191.0, 228.3)	224.5 (208.1, 240.9)	16.0 (2.8, 29.2)	-	-
BMN 110 qow	55	206.4 (187.7, 225.1)	219.1 (202.6, 235.5)	12.1 (-1.1, 25.3)	-3.9 (-22.6, 14.8)	0.6819
BMN 110 pro Woche	52	201.6 (182.3, 220.8)	241.6 (224.7, 258.5)	39.5 (25.9, 53.2)	23.6 (4.6, 42.6)	0.0154
<b>Observed cases<sup>1</sup></b>						
Placebo	59	213.4 (194.8, 232.0)	225.5 (209.6, 241.4)	13.5 (0.5, 26.5)	-	-
BMN 110 qow	59	205.7 (187.2, 224.3)	223.0 (206.9, 239.0)	14.7 (1.6, 27.8)	1.2 (-17.3, 19.7)	0.8984
BMN 110 pro Woche	58	202.3 (183.6, 221.0)	241.6 (225.4, 257.8)	36.7 (23.4, 49.9)	23.2 (4.6, 41.7)	0.0149

<sup>1</sup> ANCOVA (Änderung von Baseline bis Woche 24) mit den Kovariaten Behandlung, Altersgruppen und 6MWT-Kategorien zu Beginn der Studie.<sup>2</sup> ANCOVA mit den Faktoren Behandlung, Visite, Behandlung-Visite, Altersgruppen und 6MWT-Kategorien zu Beginn der Studie<sup>3</sup> p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell. Nach Hochberg-Adjustierung, um der Multiplizität Rechnung zu tragen, wurden die Ergebnisse als signifikant betrachtet, wenn beide p-Werte <0.05 oder ein p-Wert <0.025 lagen.

Tabelle 4-18: Ergebnisse für 6MWT (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel: Einfluss zusätzlicher Kovariaten (ITT-Population)

Kovariaten Statistik	Kovariaten-Effect	BMN 110 2.0 mg/kg/qow <sup>1</sup> vs. Placebo	BMN110 2.0 mg/kg/Woche <sup>1</sup> vs. Placebo
Interaktion Altersgruppe (12-18 Jahre) und Baseline 6MWT-Kategorie			
Schätzer (SE)	-18.7 (18.78)	0.51 (9.30)	22.52 (9.36)
p-Wert	0.3206	0.9565	0.0172
Interaktion Altersgruppe (≥19 Jahre) und Baseline 6MWT-Kategorie			
Schätzer (SE)	-22.5 (20.69)		
p-Wert	0.2775		
Geschlecht			
Schätzer (SE)	2.22 (7.78)	0.78 (9.36)	22.41 (9.37)
p-Wert	0.7757	0.9335	0.0179
Baseline 3MSC (Stufen/min)			
Schätzer (SE)	0.51 (0.33)	1.94 (9.30)	22.83 (9.31)
p-Wert	0.1230	0.8352	0.0152
Baseline MVV			
Schätzer (SE)	-0.09 (0.20)	0.41 (9.31)	21.96 (9.43)
p-Wert	0.6668	0.9650	0.0210
Baseline FVC			
Schätzer (SE)	1.12 (6.12)	0.57 (9.32)	22.68 (9.46)
p-Wert	0.8553	0.9514	0.0176
Niedriges uKS			
Schätzer (SE)	-12.3 (9.22)	0.75 (9.27)	22.65 (9.33)
p-Wert	0.1836	0.9355	0.0162
Baseline uKS			
Schätzer (SE)	-0.52 (0.29)	1.43 (9.24)	22.07( 9.29)
p-Wert	0.0730	0.8776	0.0186
Klinisch signifikant abnormales EKG (Baseline)			
Schätzer (SE)	-24.6 (51.71)	0.55 (9.31)	22.93 (9.42)
p-Wert	0.6349	0.9528	0.0160
Klinisch signifikant abnormales Echokardiogramm (Baseline)			
Schätzer (SE)	5.64 (23.71)	0.34 (9.35)	22.16 (9.46)
p-Wert	0.8121	0.9714	0.0203

Kovariaten Statistik	Kovariaten-Effect	BMN 110 2.0 mg/kg/qow <sup>1</sup> vs. Placebo	BMN110 2.0 mg/kg/Woche <sup>1</sup> vs. Placebo
Score Selbstpflege-Domäne			
Schätzer (SE)	-2.29 (1.56)	1.30 (9.27)	22.82 (9.32)
p-Wert	0.1427	0.8887	0.0154
Score Mobilitätsdomäne			
Schätzer (SE)	-2.03 (1.46)	2.24 (9.34)	22.83 (9.32)
p-Wert	0.1676	0.8111	0.0154
Normalisierte, standardisierte Größe (z-Score) entsprechend CDC			
Schätzer (SE)	0.80 (1.74)	2.58 (9.36)	22.23 (9.37)
p-Wert	0.6461	0.7834	0.0189
Standgröße			
Schätzer (SE)	0.31 (0.28)	2.58 (9.33)	22.76 (9.34)
p-Wert	0.2747	0.7825	0.0160

<sup>1</sup> Schätzung des Behandlungseffekts (BMN 110 vs. Placebo) auf den 6MWT und den p-Wert, adjustiert für eine zusätzliche Kovariate.

Zum Zeitpunkt Baseline lagen für alle Patienten Ergebnisse beider 6MWTs vor. Nach 24 Wochen war ein Patient (1.7%) der BMN 110 2.0 mg/kg/qow Gruppe nicht in der Lage, an Test 1 teilzunehmen und drei Patienten (5.1%) derselben Gruppe konnten nicht an Test 2 teilnehmen (n=1 Fehlen beider Tests).

Die mittlere Gehstrecke ( $\pm$  SD; ITT) lag zu Beginn der Studie in der Placebo-Gruppe bzw. bei den Patienten, die BMN 110 qow oder einmal wöchentlich erhielten, bei 211.9 (69.9) m, 205.7 (81.2) m bzw. 203.9 (76.3) m. Die mittlere Änderung ( $\pm$  SD) der Gehstrecke im 6MWT von Baseline bis Woche 24 betrug in der ITT-Population 13.5 ( $\pm$ 50.6), 14.9 ( $\pm$ 40.8) und 36.5 ( $\pm$ 58.5) m nach Placebo, BMN 110 2.0 mg/kg/qow bzw. BMN 110 2.0 mg/kg/Woche (entsprechend 8.7 $\pm$ 28.8%, 10.6 $\pm$ 26.5% und 23.9 $\pm$ 44.8%). Der primäre Wirksamkeitsendpunkt der Studie wurde erreicht und BMN 110 2.0 mg/kg/Woche war hinsichtlich der Änderung der Gehstrecke in Woche 24 der Placebo-Behandlung signifikant überlegen ( $p=0.0174$  nach Hochberg-Adjustierung für Multiplizität). Der geschätzte Behandlungseffekt (95% KI) in Woche 24 betrug 22.5 m (4.0, 40.9;  $p=0.0174$ ) in der BMN 110 2.0 mg/kg/Woche Gruppe und 0.5 m (-17.8, 18.9;  $p=0.9542$ ) in der BMN 110 2.0 mg/kg/qow Gruppe verglichen mit Placebo (ITT; ANCOVA unter Berücksichtigung der Kovariaten Altersgruppe und 6MWT-Kategorie zu Beginn der Studie). Wie aus Abbildung 2 ersichtlich, war bereits in Woche 12 ein Unterschied zwischen der BMN 110 2.0 mg/kg/Woche Gruppe und der Placebo-Gruppe erkennbar, der bis Woche 24 weiter zunahm.

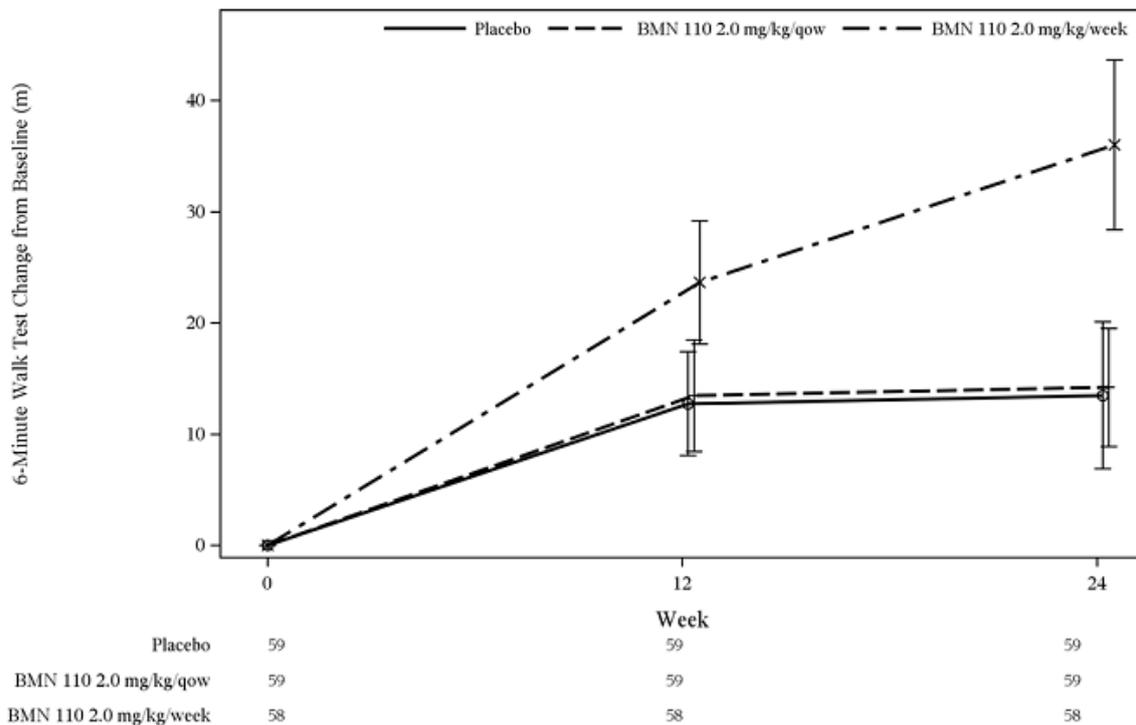


Abbildung 2: Änderung der Gehstrecke im 6MWT über die Zeit bis Woche 24 (ITT-Population)

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

<Abbildung Meta-Analyse>

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

Nicht zutreffend

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

#### 4.3.1.3.2 Responderanalyse 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 0). 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 4-19: Operationalisierung der Responder-Analyse des 6MWT

Studie	Operationalisierung
MOR-004	<p>Responderanalysen und eine Analyse mittels kontinuierlicher Verteilungsfunktion (KVF) waren prospektiv als unterstützende Analysen für den primären Endpunkt geplant. Eine quantitative Auswertung der KVF mittels Wilcoxon-Mann-Whitney Test und Umwandlung der Mann-Whitney Effektgröße in relative Risiken wurde <i>post hoc</i> für die Nutzenbewertung durchgeführt.</p> <p>Da ein validierter MCID für die heterogene Patientenpopulation nicht existierte, wurde eine Delphie-Studie initiiert, um eine Responder-Definition bzw. Schwellenwerte klinisch relevanter Änderungen des 6MWT zu ermitteln. Neun Kliniker, die Erfahrungen in der Behandlung lysosomaler Speicherkrankheiten und speziell des Morquio A-Syndroms hatten, nahmen an einer modifizierten Delphie-Studie teil. Die Studie bestand aus einem konventionellen anonymen Teil mit zwei Feedback-Runden und einem finalen Konsenustreffen der Teilnehmer. Das Delphi-Konsortium (1 Teilnehmer aus privaten Gründen verhindert) gab folgende Konsensempfehlungen hinsichtlich einer Responder-Definition (Schwellenwert) ausgedrückt als prozentuale Verbesserung des 6MWT vom Ausgangswert bis Woche 24 der Behandlung ab:</p> <ul style="list-style-type: none"> <li>– 15% bei einer basalen Gehstrecke von 30-325 m (für alle eingeschlossenen Patienten) oder alternativ:</li> <li>– 20% bei einer basalen Gehstrecke von 30-100 m</li> <li>– 15% bei einer basalen Gehstrecke von 100-200 m</li> <li>– 10% bei einer basalen Gehstrecke von 200-325 m</li> </ul> <p>Der Anteil der Patienten, deren prozentuale Änderung der Gehstrecke gemessen mittels 6MWT in Woche 24 oberhalb der Empfehlung des Delphi-Gremiums lag, wurde mittels logistischer Regression nach Adjustierung für Altersgruppen und Baseline-6MWT-Kategorie ausgewertet.</p> <p>Zusätzlich wurde der Prozentsatz der Patienten, deren Gehstrecke sich im 6MWT in Woche 24 verglichen mit dem Baseline-Wert um „X“ Meter oder mehr verbessert hatte, für die drei Behandlungsarme in Abhängigkeit der verschiedenen „X“ als kumulative Verteilungsfunktion (KVF) dargestellt. In diesem Dossier werden für diesen Endpunkt jedoch der Einfachheit halber nur die relevanten Daten für die Behandlung mit Placebo oder mit BMN 110 einmal pro Woche dargestellt. Der Prozentsatz derjenigen Patienten, die in Woche 24 mit „X“ Metern oder mehr auf die Therapie angesprochen hatten, wurde gegen eine kontinuierliche Skala möglicher Verbesserungen aufgetragen. Dieses Vorgehen erlaubte einen Vergleich der Behandlungsgruppen unter Anwendung verschiedener Response-Kriterien, sodass mehrere Responder-Definitionen simultan bewertet werden konnten. Dieser Ansatz wurde - bei einer seltenen Erkrankung wie MPS IVA - gegenüber einer kategorialen Responder-Definition aufgrund der kleinen Gruppengröße und der heterogenen Patientenpopulation bevorzugt, da ein möglicher Verlust statistischer Power den Nachweis eines Behandlungseffekts verhindern könnte, und ein einziges Responsekriterium nicht sinnvoll auf die heterogene Patientenpopulation anwendbar ist.</p> <p>Für die quantitative Auswertung der Unterschiede in den KVF's (Dichotomisierung des Endpunktparameters) wurde der robuste Wilcoxon-Mann-Whitney Test herangezogen, um die Mann-Whitney Effektgröße, sowie ihr dazugehöriges Konfidenzintervall (robust procedure) darzustellen (41). Die Berechnungen wurden mit dem validierten Programm TESTIMATE (idv-Data Analysis and Study Planning) durchgeführt. Das Konfidenzintervall basierte auf der Varianzschätzung nach Wei L.J. und Lachin J.M. (42–44). Die Umwandlung der Mann-Whitney Effektgröße in andere Effektgrößen (POR, RR, etc.) erfolgte nach der „Substitution method“ (45), und basierte auf folgendem Zusammenhang:</p> $MW = \frac{\beta}{(\beta-1)^2} (\beta-1-\ln(\beta)) \text{ [POR} = \beta; \text{ unter Annahme einer logistischen Verteilung].}$ <p>Die Odds Ratio wurde mittels iterativer Approximation ermittelt.</p> <p>Relative Risiken wurden aus den Odds Ratios wie folgt abgeleitet:</p> $RR(CER) = \frac{\beta}{1-CER+\beta-CER}$ <p>Eine Übersicht der verwendeten Formeln wurde vor Kurzem in der wissenschaftlichen Literatur</p>

publiziert (41,46).
---------------------

*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 4-20: Bewertung des Verzerrungspotenzials für Responder-Analyse 6MWT 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
MOR-004	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Angesichts der Tatsache, dass es sich um eine doppelblinde, randomisierte Studie handelt, deren Verzerrungspotential als niedrig eingestuft wurde, und sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung ergeben, wurde das Verzerrungspotential für diesen Endpunkt als niedrig eingestuft.

*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.*

Auf Grund der extrem seltenen Erkrankung ist für den 6MWT noch kein validiertes Response-Kriterium für Morquio A Patienten etabliert, auch wenn diese Studie die größte Enzymersatzstudie für lysosomale Speicherkrankheiten darstellt. Aus diesem Grund wurden *post hoc* Analysen unter Einsatz kategorialer Response-Kriterien durchgeführt, die von klinischen Experten in einer Delphi Studie vorgeschlagen wurden (Tabelle 4-21). Die relativen Response-Raten der BMN 110- und Placebo-Gruppe (basierend auf ORs und den Delphi Kriterien) waren zwar groß, erreichten allerdings nicht das Signifikanzniveau, was auf die relativ geringe Studiengröße und den Powerverlust durch die Dichotomisierung zurückzuführen ist. Bei den Responder-Kriterien der Delphi-Studie handelt es sich nicht um patientenbasierte Daten, sodass deren Relevanz für die Patienten formell unklar ist. Es sollte jedoch betont werden, dass es ethisch und praktisch unwahrscheinlich ist, *a priori* (vor

Durchführung der pivotalen Studie) ein Responsekriterium für Arzneimittel zur Behandlung von seltenen Leiden validieren zu können, insbesondere wenn das Arzneimittel als erste medikamentöse Intervention in einer neuen Indikation eingesetzt wird. Auch auf Grund der vielen potentiellen Effektmodifikatoren, die einen Einfluss auf die relevante Gehstreckenveränderung nehmen könnten, war eine *a priori* Definition eines individuellen minimal relevanten Unterschiedes für die zu erwartende heterogene Patientenpopulation mit Morquio A Syndrom nicht sinnvoll möglich.

Tabelle 4-21: 6MWT Responder-Analyse basierend auf den Delphi-Kriterien (ITT, „observed cases“, Post-hoc-Analyse)

Gruppe	n/N (%)	BMN vs. Placebo OR <sup>1</sup> (95% CI)	p-Wert
<b>&gt;15% relative Verbesserung bis Woche 24 (Ausgangsgehstrecke 30-325 m)</b>			
Placebo	18/59 (30.5)	-	-
BMN qow	22/58 (37.9)	1.44 (0.65, 3.18)	0.3703
BMN pro Woche	26/57 (45.6)	2.14 (0.97, 4.72)	0.0603
<b>Relative Verbesserung abhängig von Ausgangsgehstrecke<sup>2</sup></b>			
Placebo	20/59 (33.9)	-	-
BMN qow	24/58 (41.4)	1.42 (0.66, 3.07)	0.3719
BMN pro Woche	27/57 (47.4)	1.93 (0.89, 4.18)	0.0950

<sup>1</sup> Die OR entspricht dem Verhältnis der Odds-Rate von BMN 110 gegenüber Placebo. Die OR wurde mittels eines Logit-Modells adjustiert für Altersgruppen und 6MWT Baseline-Kategorie berechnet.

<sup>2</sup> Responder-Kriterien: >20% Änderung bei einer Ausgangsgehstrecke zwischen 30 und <100 m, 15% bei einer Ausgangsgehstrecke von 100 bis <200 m und 10% bei einer Ausgangsgehstrecke von 200 bis 325 m.

Abbildung 4 zeigt die Resultate einer Responderanalyse unter Anwendung eines breiteren Bandes von Response-Kriterien (von  $\geq 0$  m bis  $\geq 100$  m Verbesserung der 6-Minuten-Gehstrecke gegenüber Baseline). Der Anteil der Patienten, deren Gehstrecke sich bis Woche 24 verglichen mit dem Ausgangswert deutlich besserte, war unter Behandlung mit BMN 110 einmal pro Woche für jedes beliebige Response-Kriterium größer, als unter Behandlung mit Placebo. Für ein Response-Kriterium von  $\geq 60$  m Verbesserung ergab sich nach den Kriterien von GRADE (47) ein großer Effektunterschied auf Basis des relativen Risikos zwischen der Behandlung mit Placebo oder mit BMN 110 einmal pro Woche ( $RR > 2$  bzw.  $< 0.5$ ). Das relative Risiko der beiden Behandlungsgruppen (Placebo versus BMN 110 einmal pro Woche) war trotz der geringen Studiengröße für ein Responsekriterium von  $\geq 60$  m statistisch signifikant ( $RR=0.458$ ;  $KI: 0.226-0.926$ ; Responderrate Placebo: 9/59; Responderrate BMN 110 einmal pro Woche: 19/57). Obwohl eine *a priori* definierte minimal klinisch relevante Veränderung nicht definiert war, ist eine Verbesserung der Gehstrecke von  $\geq 60$  m in dieser Patientenpopulation mit Sicherheit als klinisch relevant einzustufen.

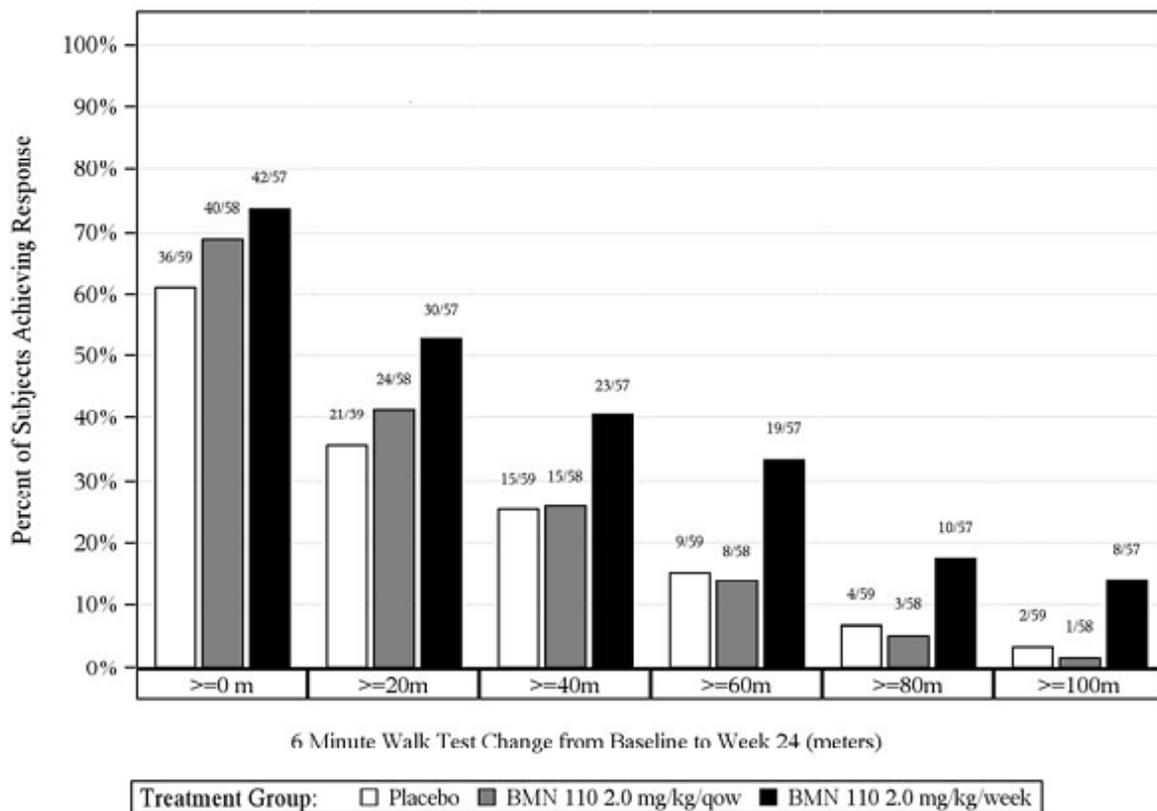


Abbildung 4: Responder-Analyse des 6MWTs kategorisiert nach der Änderung der Gehstrecke von Baseline bis Woche 24 („observed cases“)

Klassische Responderanalysen, wie die oben beschriebenen, leiden jedoch in vielen Fällen unter signifikanten Nachteilen, insbesondere wenn stetige Daten in Kategorien umgewandelt werden (48,49):

- durch die Dichotomisierung wird ein erheblicher Informationsverlust relativ zu einer Auswertung auf der Originalskala in Kauf genommen.
- Ein Response-Kriterium auf einer kontinuierlichen Skala ist häufig – zumindest für eine heterogene Patientenpopulation - klinisch nicht eindeutig sondern arbiträr.
- Insbesondere bei unterschiedlichen Varianzen in den zu vergleichenden Behandlungsgruppen kann der Unterschied der Responderraten in Abhängigkeit des Response-Kriteriums potentiell beträchtlich variieren, oder sogar zu gegenteiligen Aussagen führen.
- Eine beobachtete Response lässt sich nicht kausal auf die zugrunde liegende Intervention zurückführen (z. B.: regression-to-the-mean, Trainingseffekte).

- Weder die gängigen Parallelgruppen- noch Cross-Over-Studien erlauben es, überhaupt Aussagen zur *individuellen* Response zu machen. Selbst eine Verschlechterung unter Therapie könnte eine Response darstellen, da das Ausmass der Verschlechterung unter Kontrollbedingungen ja noch grösser sein könnte (chronisch progressive Erkrankungen).

Der Versuch einer Quantifizierung eines klinisch relevanten Zusatznutzens mittels klassischer Responderanalysen ist somit insbesondere bei Studien für Arzneimittel zur Behandlung von seltenen Leiden und heterogenen Patientenpopulationen kritisch zu betrachten.

Für den Fall von „sonstigen“ Zielgrößen, für die keine Responderauswertungen mit validierten Responsekriterien - und mit daraus ableitbaren relativen Risiken - vorliegen, wurde vom IQWiG eine approximative Methodik vorgeschlagen, um eine Ausmaßfeststellung durchzuführen (Referenz 101 in (7)). Eine im Ansatz ähnliche, robuste und effiziente statistische Analysemethode wurde daher vom pU für diesen Endpunkt eingesetzt. Diese Methodik erlaubte es, einen Zusatznutzen für die mit BMN 110 2.0 mg/kg/Woche behandelte Gesamtpopulation gegenüber der gesamten Placebogruppe zu quantifizieren.

Die Analyse basierte auf einer Auswertung der kumulativen Verteilungsfunktion (KVF) für die individuellen Veränderungen der Gehstrecken aller Patienten im 6MWT. Die KVF (Abbildung 5) zeigte für das Kontinuum aller theoretisch möglichen Responder-Definitionen eine klare Trennung zwischen der Placebo-Gruppe und der Patientengruppe, die mit BMN 110 2.0 mg/kg/Woche behandelt wurde. Die Unterschiede zwischen den beiden Behandlungsgruppen entsprachen annähernd einer Parallelverschiebung der Placeboverteilung im Vergleich zur Verteilung der Patienten unter Behandlung mit Elosulfase alfa. Beide Funktionen waren näherungsweise normalverteilt. Unabhängig vom Wert der Gehstreckenveränderung (x-Achse, Abbildung 5) war der kumulative Prozentsatz der Patienten, die einen beliebigen Wert einer Verbesserung erreichten, unter Behandlung mit BMN 2.0 mg/kg/Woche höher als unter Behandlung mit Placebo (y-Achse, Abbildung 5). Es war somit klar erkennbar, dass alle Patienten unter Behandlung von BMN 110 gegenüber der Placebo-Intervention profitierten.

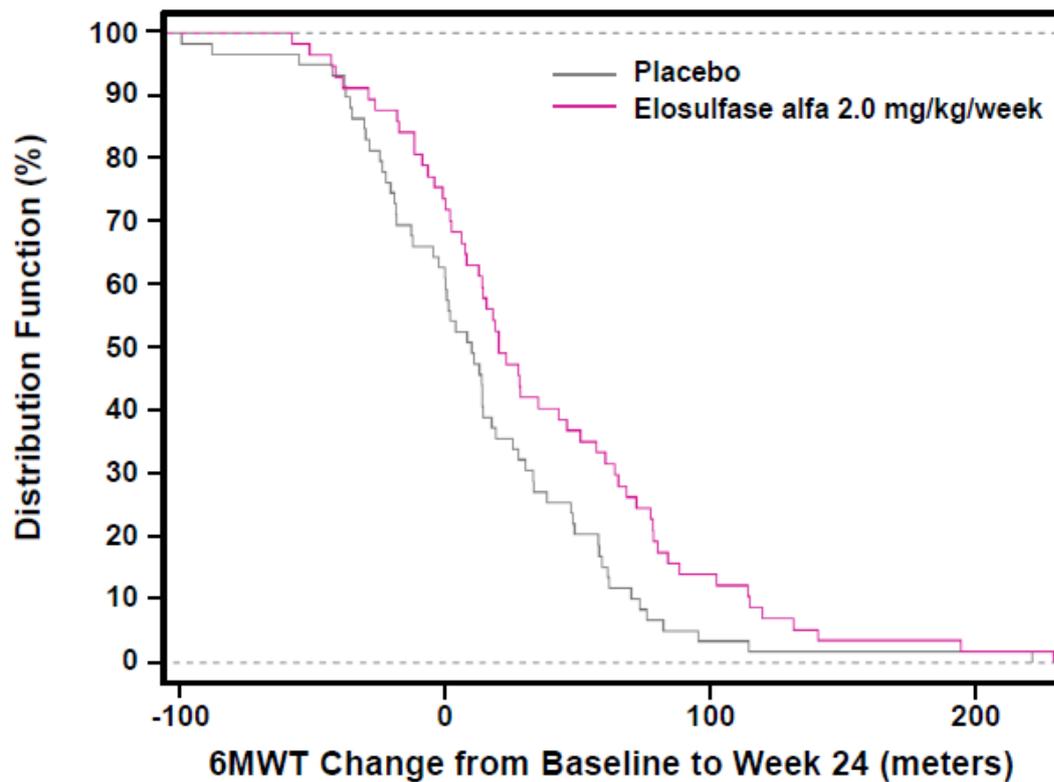


Abbildung 5: Responder-Analyse des 6MWTs: KVF der Veränderung von Baseline bis Woche 24 (ITT-Population)

Für jeden Punkt der Kurve entspricht der Wert auf der y-Achse dem Anteil der Patienten, deren Werte (Gehstrecke im 6MWT von Baseline bis Woche 24) größer oder gleich dem korrespondierenden Wert auf der x-Achse an diesem Punkt waren.

Als quantitatives Maß für den Unterschied der Gehstreckenverbesserung zwischen beiden Behandlungsgruppen wurde die Wilcoxon-Mann-Whitney Effektgröße (MW-AUC) herangezogen, da sie robust, nicht-parametrisch, und daher modellunabhängig ist.

Grafisch lässt sich die Mann Whitney AUC als ‚Percentage/Percentage Plot‘ darstellen, in dem die Responderaten aus der kontinuierlichen Verteilungsfunktion (Abbildung 5) für die Placebogruppe den Responderaten der Verumgruppe gegenübergestellt werden. In dem hypothetischen Fall von identen Responderaten in der Placebo- und der Verumgruppe ergäbe sich in dieser Darstellung eine Diagonale von links unten nach rechts oben; die Fläche unter dieser Kurve (AUC) ergäbe einen dimensionslosen Wert von 0.5. Eine Überlegenheit von Verum gegenüber Placebo wird durch eine Fläche (Mann Whitney AUC) von  $> 0.5$  repräsentiert (Abbildung 6).

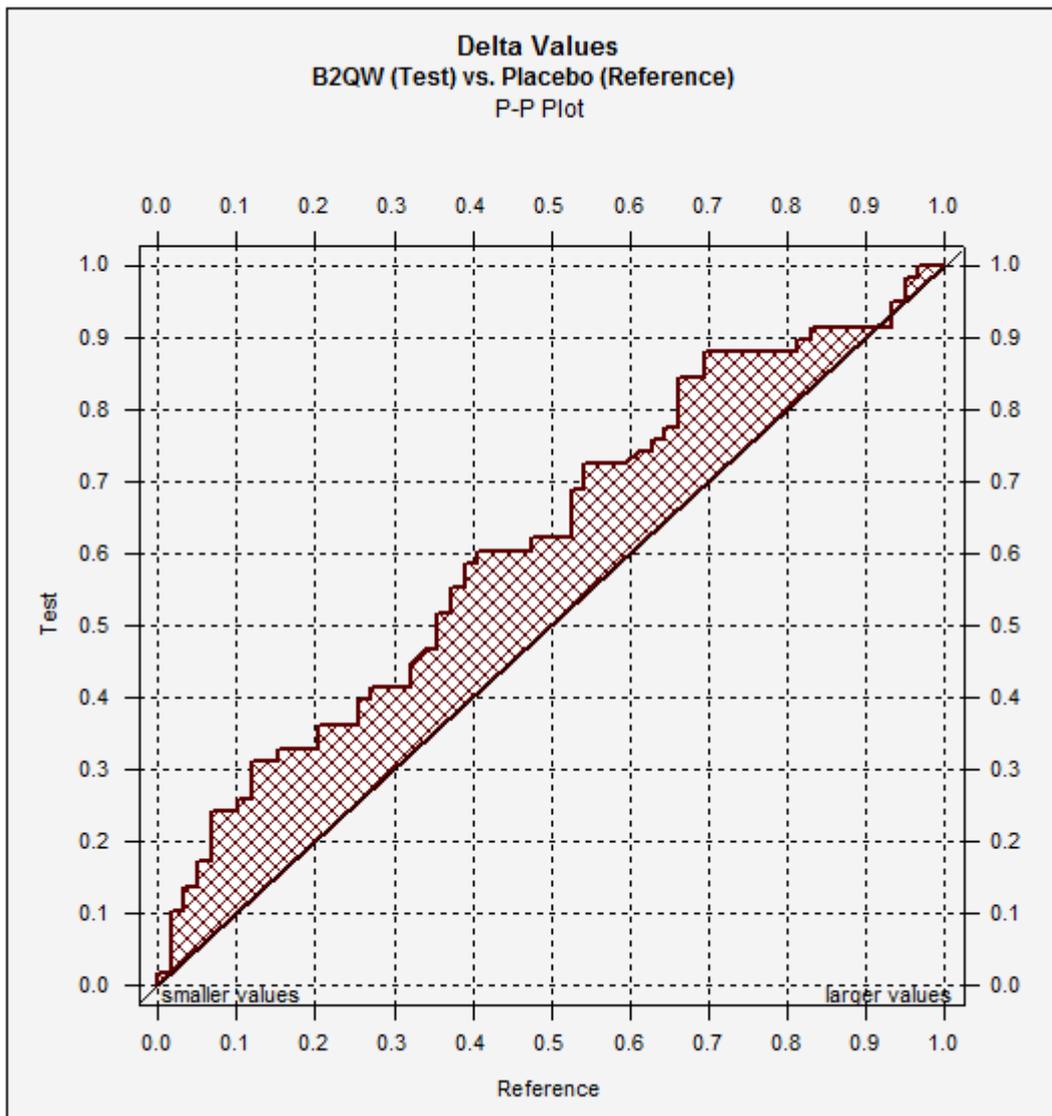


Abbildung 6: Percentage/Percentage Plot für die Gesamtpopulation (BMN 110 2.0 mg/kg/Woche versus Placebo)

Die Auswertung der Mann-Whitney Statistik zeigte eine signifikante Überlegenheit der Gehstreckenänderungen für die gesamte mit BMN 110 2.0 mg/kg/Woche behandelte Population gegenüber der mit Placebo behandelten Population; MW: 0.6166 (95% KI: 0.5123 – 0.7209) (Tabelle 4-22).

Tabelle 4-22: Wilcoxon-Mann-Whitney-U-Test (Gesamtpopulation)

Statistik	Gruppe	
	BMN 110 2.0 mg/kg/Woche	Placebo
N	58	59
Mean Rank	65.879	52.237
Hodges-Lehmann Schätzer	18.400	
$P(X<Y) + 0.5 P(X=Y)$	0.6166	
95% KI	0.5123 – 0.7209	
S.E. (Wei-Lachin)	0.0532	
Rang Summe	3821.0	
p-Wert (exakt)	0.0294	

Die Mann-Whitney Effektgröße zeichnet sich weiterhin dadurch aus, dass sie in andere bekannte und häufig verwendete Effektgrößenmaße umwandelbar ist (POR, RR, NNT, etc.) (45,46). Die Umwandlung der MW Effektgröße in ein OR ergab einen Wert von 2.0368 (95% KI: 1.0766 - 4.1090) (Odds Verbesserung in der Gruppe BMN 2.0 mg/kg/Woche zu Odds Verbesserung in der Gruppe Placebo) (s. Tabelle 4-25). Unter Annahme eines Event-Risikos in der Kontrollgruppe (control event risk, CER) konnten die relativen Risiken aus der Odds Ratio abgeleitet werden. In Tabelle 4-23 sind die abgeleiteten relativen Risiken für die gegebenen Kontrollgruppen-Risiken (von 5% bis 40%) dargestellt.

Für die in Tabelle 4-23 angegebenen CERs ließen sich auch aus den individuellen Patientendaten die entsprechenden Gehstreckenveränderungen ableiten (s. Abbildung 5). Diese von den CERs abhängigen Gehstreckenveränderungen entsprechen formell den Responsekriterien für die dazugehörigen relativen Risiken (RRs). Da die von den CERs abgeleiteten relativen Risiken alle deutlich signifikant waren, und da in dem in Tabelle 4-23 angegebenen Kontinuum der Gehstreckenveränderungen mit Sicherheit klinisch relevante Veränderungen repräsentiert waren (von 14 m bis  $\geq 95$  m), ist der Unterschied der Gehstreckenveränderungen zwischen den beiden Behandlungsgruppen als klinisch relevant zu betrachten.

Es zeigte sich somit auch auf Basis der aus der Mann-Whitney Effektgröße abgeleiteten relativen Risiken eine signifikante und klinisch relevante Überlegenheit in Bezug auf die Verbesserung der Gehstrecken für die mit BMN 110 2.0 mg/kg/Woche behandelten Patientenpopulation gegenüber der Placebogruppe. Nach § 5 (7) Abs. 2 AM NutzenV ist die

Überlegenheit der Behandlung mit Elosulfase alfa als eine therapierelevante Abschwächung schwerwiegender Symptome gegenüber der Vergleichsintervention zu betrachten, und somit als ein beträchtlicher Zusatznutzen.

Tabelle 4-23: Ableitung des relativen Risikos aus der OR (Gesamtpopulation)

		<b>RR derived from OR</b> OR (= $\beta$ ): 2.0368 (95% CI: 1.0766 - 4.1090)			<b>Invertiertes Relatives Risiko</b> OR (= $\beta$ ): 0.4910 (95% CI: 0.2434 - 0.9289)		
CER	Interpolierte Gehstreckenveränderung (Meter)	RR (CER)	RR (CER) unteres Limit	RR (CER) oberes Limit	RR	RR unteres Limit	RR oberes Limit
0.0500	$\geq 95.1$	1.9364	1.0725	3.5562	0.5164	0.2812	0.9324
0.1000	$\geq 73$	1.8455	1.0684	3.1345	0.5419	0.3190	0.9360
0.2000	$\geq 57.5$	1.6870	1.0604	2.5336	0.5928	0.3947	0.9431
0.3000	$\geq 33$	1.5536	1.0524	2.1260	0.6437	0.4704	0.9502
0.4000	$\geq 14.1$	1.4397	1.0446	1.8314	0.6946	0.5460	0.9573

*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.*

Um potentielle Effektmodifikatoren zu identifizieren, und um den Anforderungen hinsichtlich der geforderten Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext nachzukommen, wurde untersucht, ob rassenabhängige physiologisch bedingte Unterschiede zu Unterschieden in der Wirksamkeit von Elosulfase alfa führen könnten (50,51).

Auf Grund der geringen Studiengröße war eine retrospektive, ausführliche statistische Analyse (mittels Heterogenitätstests) von potentiellen Effektmodifikatoren nicht sinnvoll. Der pU beschränkte sich daher darauf, eventuelle Effektunterschiede zwischen der weißen Patientenpopulation (50) und der gesamten Studienpopulation zu untersuchen, da die Patientengruppe, die mit BMN 110 2.0 mg/kg/Woche behandelt worden war, nur 62.1% weiße Patienten enthielt (s. Tabelle 4-11). Zur Auswertung wurden daher alle weißen Patienten in der Placebogruppe (n=44) und in der mit BMN 110 2.0 mg/kg/Woche behandelten Patientenpopulation (n=36) herangezogen. Die KVF dieser beiden Populationen ist in Abbildung 7 dargestellt.

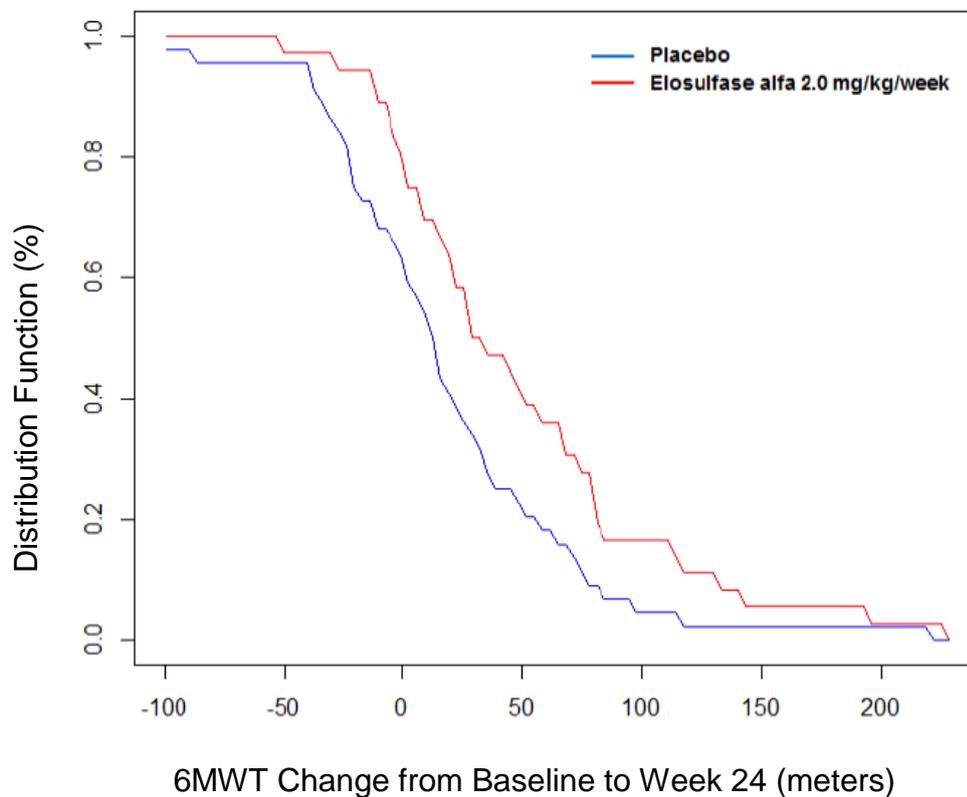


Abbildung 7: Responder-Analyse des 6MWTs: KVF der Veränderung von Baseline bis Woche 24 (Weiße-Population (50))

Die grafische Darstellung der kontinuierlichen Verteilungsfunktion als ‚Percentage/Percentage Plot‘ ist in Abbildung 8 zu finden.

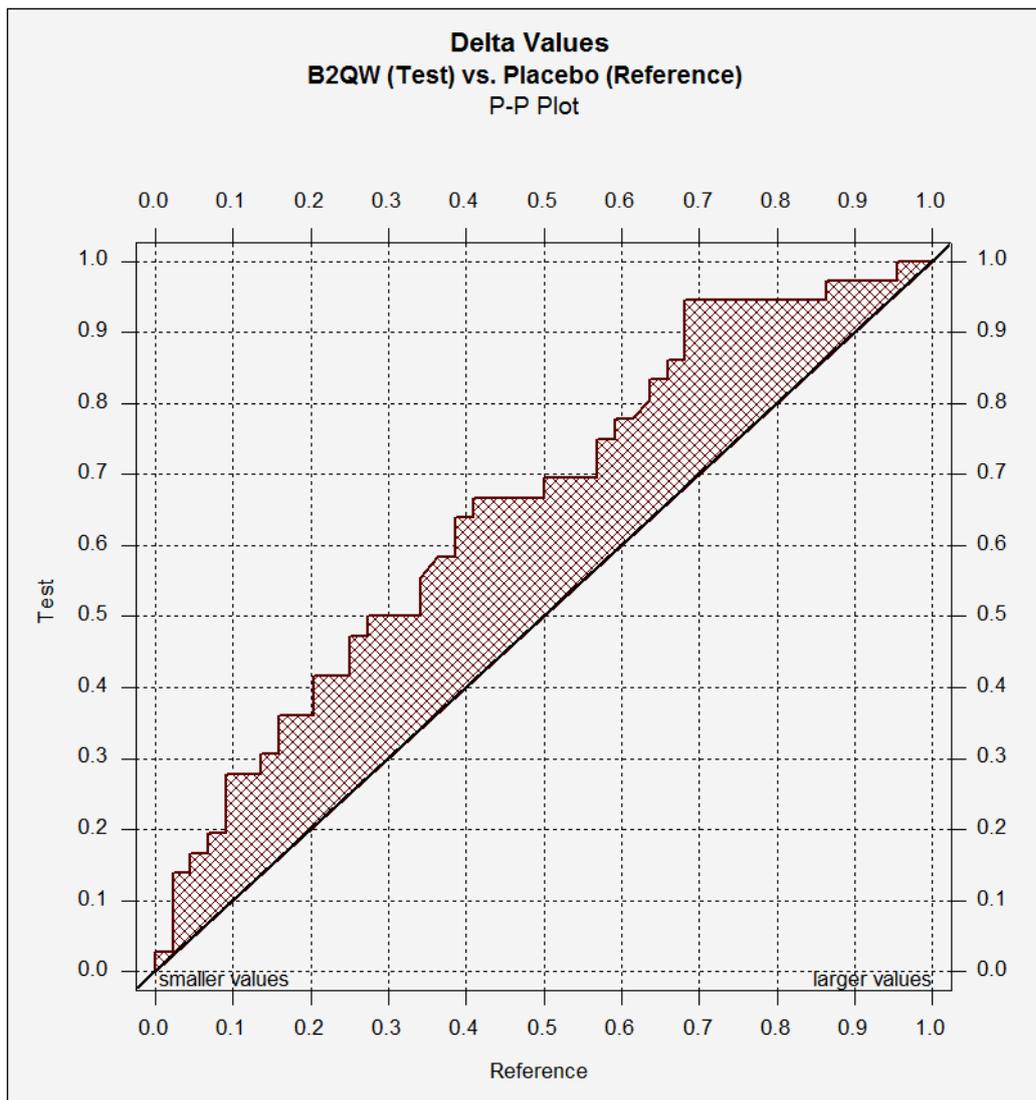


Abbildung 8: Percentage/Percentage Plot für die weiße Population (BMN 110 2.0 mg/kg/Woche versus Placebo)

Die Auswertung der Mann-Whitney Statistik zeigte eine signifikante Verbesserung der Gehstrecken in der mit BMN 2.0 mg/kg/Woche behandelten weißen Population gegenüber der mit Placebo behandelten weißen Population, die numerisch größer war als der Unterschied in der Gesamtpopulation; MW: 0.6572 (95% KI: 0.5328– 0.7816) (Tabelle 4-24, Tabelle 4-25).

Tabelle 4-24: Wilcoxon-Mann-Whitney-U-Test (weiße Population (50))

Statistik	Gruppe	
	BMN 110 2.0 mg/kg/Woche	Placebo
N	36	44
Mean Rank	47.417	34.841
Hodges-Lehmann Schätzer	26.200	
$P(X < Y) + 0.5 P(X = Y)$	0.6572	
95% KI	0.5328 – 0.7816	
S.E. (Wei-Lachin)	0.0635	
Rang Summe	1707.0	
p-Wert (exakt)	0.0156	

Wie für die Gesamtpopulation wurde die Mann-Whitney Effektgröße in ein Odds Ratio umgewandelt. Der Effektgrößenunterschied zwischen den beiden Behandlungsgruppen war auf Basis der aus der Mann-Whitney Effektgröße abgeleiteten Odds Ratio beträchtlich; OR = 2.6456 (95% KI: 1.2178 - 6.6101) (s. Tabelle 4-25).

Tabelle 4-25: Effektgrößenunterschiede der Gehstreckenveränderungen von Baseline zu Woche 24 (BMN 2.0 mg/kg/Woche versus Placebo)

	Effektgrößenmaß	
	MW Statistik	OR
<b>Gesamtpopulation</b>	0.6166 (95% KI: 0.5123 - 0.7209)	2.0368 (95% KI: 1.0766 - 4.1090)
<b>Weißer Population (50)</b>	0.6572 (95% KI: 0.5328 - 0.7816)	2.6456 (95% KI: 1.2178 - 6.6101)

In Tabelle 4-26 sind die aus der Odds Ratio abgeleiteten relativen Risiken für die entsprechenden Kontrollgruppenrisiken und die dazugehörigen Gehstreckenveränderungen (von 19 m bis  $\geq 95$  m) der weißen Patientenpopulation angegeben. Wie für die Gesamtpopulation kann aus der signifikanten und klinisch relevanten Überlegenheit der

Intervention mit BMN 110 2.0 mg/kg/Woche auf Basis der relativen Risiken ein beträchtlicher Zusatznutzen für das zu bewertende Arzneimittel abgeleitet werden.

Tabelle 4-26: Ableitung des relativen Risikos aus der OR (weiße Population (50))

		RR derived from OR OR (= $\beta$ ): 2.6456 (95% CI: 1.2178 - 6.6101)			Invertiertes Relatives Risiko OR (= $\beta$ ): 0.3780 (95% CI: 0.1513 - 0.8212)		
CER	Interpolierte Gehstreckenveränderung (Meter)	RR (CER)	RR (CER) unteres Limit	RR (CER) oberes Limit	RR	RR unteres Limit	RR oberes Limit
0.0500	≥ 95.1	2.4445	1.2047	5.1621	0.4091	0.1937	0.8301
0.1000	≥ 75.6	2.2718	1.1918	4.2345	0.4402	0.2362	0.8390
0.2000	≥ 57.7	1.9905	1.1670	3.1150	0.5024	0.3210	0.8569
0.3000	≥ 33	1.7712	1.1431	2.4637	0.5646	0.4059	0.8748
0.4000	≥ 19	1.5954	1.1202	2.0376	0.6268	0.4908	0.8927

Die Methodik des IQWiG wird für die Quantifizierung des Zusatznutzens für Arzneimittel zur Behandlung von seltenen Leiden nicht herangezogen. Hier sei jedoch festgehalten, dass das obere Band des Konfidenzintervalls sämtlicher relativen Risiken unter dem Schwellenwert von 0.9 lag (Tabelle 4-26), und daher auch nach IQWiG Methodik für die weiße Population ein beträchtlicher Zusatznutzen festzustellen ist.

<Abbildung Meta-Analyse>

Abbildung 9: Meta-Analyse für <Endpunkt xxx> aus RCT; <zu bewertendes Arzneimittel> versus <Vergleichstherapie>

Nicht zutreffend

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.

#### 4.3.1.3.3 Ausdauer im 3MSCT – 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 0). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

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

Tabelle 4-27: Operationalisierung des Endpunkts „3MSCT“

Studie	Operationalisierung
MOR-004	<p>Die Patienten unterzogen sich während der Screening-Visite, in Woche 12 und 24 sowie im Falle eines vorzeitigen Abbruchs innerhalb von einer Woche einem 3MSCT (sekundärer Endpunkt). Um möglichen Schwierigkeiten im Ablauf in Woche 12 vorzubeugen, konnten die Patienten den 3MSCT-Test innerhalb von 5 d vor der Infusion in Woche 11 oder 13 absolvieren. Gleichermaßen konnte der 3MSCT-Test innerhalb von 5 d vor der Infusion in Woche 23 stattfinden. Der Test wurde zu jedem Untersuchungszeitpunkt zweimal an verschiedenen Tagen ausgeführt. Die beiden 3MSCT sollten möglichst zur selben Tageszeit stattfinden und die Ausdauertests (6MWT und 3MSCT) durften nicht an ein und demselben Tag erfolgen. Die Ausdauertests sollten möglichst in der folgenden Reihenfolge stattfinden: 6MWT, 3MSCT, 6MWT und 3MSCT. Patienten, die 30 min oder länger gereist waren, sollten nicht an diesem Tag getestet werden. Die Patienten wurden entsprechend eines vorgegebenen Skriptes aufgefordert, über 3 min eine Treppe mit Geländer, welches zur Unterstützung genutzt werden konnte, hoch zu gehen. Der Wortlaut der Motivationen seitens des Klinikpersonals während des Tests war ebenfalls vorgegeben. Die Treppe sollte angemessen beleuchtet sein und mindestens 127 Stufen haben, das Fehlen einer Treppe mit 127 Stufen stellte allerdings keinen Ausschlussgrund dar. In letzteren Fall sollte die Treppe mindestens so viele Stufen haben, dass der Patient deren Ende nicht innerhalb von weniger als 1 min erreicht. Die Dauer des Tests wurde mit einer Stoppuhr erfasst, die vor jedem Test kalibriert werden sollte. Die Tests sollten im Verlauf der Studie möglichst von denselben Personen betreut werden. Eine Änderung der Örtlichkeit oder ein Wechsel der Stoppuhr waren im Verlauf der Studie nicht zulässig. Eine physische Unterstützung des Patienten während des Tests war nicht erlaubt. Die Patienten konnten während des Tests im Stehen eine Pause einlegen (angelehnt an eine Wand oder das Geländer), durften sich aber nicht hinsetzen, in welchem Fall der Test beendet war. Die nach einer und zwei Minuten erreichten Stufen wurden jeweils mit einem Bohnensack gekennzeichnet. Modifikationen des Tests (z.B. im Falle einer Sehstörung) waren nur nach Rücksprache mit dem Monitor zulässig. Im Falle der Patient erreichte das Ende der Treppe in weniger als 3 min, sollte die Zeit bis zum Erreichen des Endes der Treppe dokumentiert werden. Der Puls und die Sauerstoffsättigung wurden vor dem Test und unmittelbar sowie 2 min nach Beendigung des Tests überwacht.</p> <p>Das in Tabelle 4-14 für den 6MWT-Test beschriebene Trainings- und Qualitätssicherungsprogramm wurde für den 3MSC-Test ebenfalls implementiert.</p> <p>In die Wirksamkeitsanalyse wurden alle randomisierten Patienten eingeschlossen (ITT). Es wurde die Zahl der Stufen, die der Patient pro Minute erklommen hatte, ausgewertet:</p> $\text{Zahl der Stufen pro Minute} = \text{Zahl der Stufen} / \text{Zeit}$ <p>Es wurde jeweils der Mittelwert aus beiden Tests zu jeden Untersuchungszeitpunkt gebildet oder das Ergebnis eines Einzeltests eingegeben, wenn nur ein Testresultat vorlag. Die Änderung der Zahl der pro Minute erklommenen Stufen in Woche 24 verglichen mit dem Ausgangswert wurde mittels ANCOVA-Model mit der Kovariate 3MSC zur Basisvisite (Stufen/min) und den Faktoren Behandlung, 6MWT-Stratifizierung bei der Eingangsvisite (<math>\leq 200</math> m, <math>&gt; 200</math> m) sowie Altersgruppen (5-11, 12-18, <math>\geq 19</math> Jahre) ausgewertet. Die BMN 110-Gruppen wurden jeweils mit Placebo verglichen und die p-Werte mittels t-Test im ANCOVA-Modell bestimmt. Zur Adjustierung der Multiplizität wurde das Verfahren nach Hochberg verwendet, um einen Typ I-Fehler von 0.05 zu erhalten. Bei Patienten, die während der Studie verstorben waren oder physisch nicht in der Lage waren, an dem Test teilzunehmen, wurde für fehlende Werte 0 eingegeben. In der primären Analyse wurde multiple Imputation eingesetzt, wenn Werte aus anderen Gründen fehlten (siehe Kapitel 4.3.1.3.1). Der Parameter Änderung der Zahl der erklommenen Stufe vom Ausgangswert bis Woche 24 wurde analog der Primäranalyse ausgewertet.</p>

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

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

Tabelle 4-28: Bewertung des Verzerrungspotenzials für den Endpunkt „3MSCT“ 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
MOR-004 (9)	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt 3MSCT wird unter Berücksichtigung der in Tabelle 4-28 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Die Begründung der Einschätzung entspricht derjenigen des 6MWT (siehe Abschnitt 4.3.1.3.1).

*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 4-29: Ergebnisse für den Endpunkt „3MSCT“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik)

MOR 004	n <sup>1</sup>	3MSCT Baseline (Stufen/min) Mittelwert (SD)	3MSCT 24 Wochen (Stufen/min) Mittelwert (SD)	Änderung von Baseline bis Woche 24 (Stufen/min) Mittelwert (SD)	Änderung von Baseline bis Woche 24 (%) Mittelwert (SD)
Placebo	59	30.0 (14.05)	33.6 (18.36)	3.6 (8.51)	11.4 (24.05)
BMN 110 qow	59	27.1 (15.80)	30.4 (17.77)	3.2 (10.29)	18.8 (49.25)
BMN 110 pro Woche	58	29.6 (16.44)	34.3 (18.70)	4.7 (7.99)	25.7 (46.14)

<sup>1</sup> Prozentuale Änderung bis Woche 24: n=56 Placebo, n=58 BMN 110 qow, n=55 BMN 110 pro Woche

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

Tabelle 4-30: Ergebnisse für den Endpunkt „3MSCT“ (mittlere Änderung der pro Minute erklommenen Stufen von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (Primär- und Sensitivitätsanalysen)

MOR-004	N	3MSC Baseline (Stufen/min) LSM (95% KI)	3MSC 24 Wochen (Stufen/min) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (Stufen/min) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (Stufen/min) BMN 110 vs. Placebo LSM (95% KI)	p-Wert (t-Test <sup>3</sup> ) BMN 110 vs. Placebo
<b>ANCOVA (ITT, Primäranalyse)<sup>1</sup></b>						
Placebo	59	30.1 (27.0, 33.1)	32.2 (29.8, 34.5)	3.6 (1.3, 6.0)	-	-
BMN 110 qow	59	27.3 (24.3, 30.3)	31.7 (29.3, 34.1)	3.2 (0.9, 5.5)	-0.5 (-3.7, 2.8)	0.7783
BMN 110 pro Woche	58	29.3 (26.3, 32.4)	33.3 (30.9, 35.7)	4.8 (2.5, 7.1)	1.1 (-2.1, 4.4)	0.4935
<b>ANCOVA mit Messwiederholung (ITT)<sup>2</sup></b>						
Placebo	59	30.4 (26.5, 34.2)	32.6 (30.3, 34.8)	3.6 (1.3, 5.8)	-	-
BMN 110 qow	59	27.1 (23.3, 30.9)	32.3 (30.0, 34.6)	3.3 (1.1, 5.6)	-0.2 (-3.5, 3.0)	0.8881
BMN 110 pro Woche	58	29.2 (25.4, 33.0)	33.7 (31.4, 36.0)	4.7 (2.4, 7.0)	1.2 (-2.1, 4.4)	0.4782
<b>ANCOVA: Gepoolter mittlerer Placebowert (ITT)<sup>1</sup></b>						
Placebo	59	30.1 (27.0, 33.1)	32.5 (30.3, 34.8)	3.7 (1.4, 5.9)	-	-
BMN 110 qow	59	27.3 (24.3, 30.3)	32.3 (30.0, 34.6)	3.4 (1.1, 5.7)	-0.3 (-3.5, 3.0)	0.8719
BMN 110 pro Woche	58	29.3 (26.3, 32.4)	33.6 (31.3, 35.9)	4.7 (2.4, 7.0)	1.1 (-2.2, 4.3)	0.5138
<b>ANCOVA: LOCF-Imputation (ITT)<sup>1</sup></b>						
Placebo	59	30.1 (27.0, 33.1)	32.5 (30.2, 34.8)	3.6 (1.3, 5.9)	-	-
BMN 110 qow	59	27.3 (24.3, 30.3)	32.2 (29.9, 34.5)	3.3 (1.0, 5.6)	-0.4 (-3.7, 2.9)	0.8116
BMN 110 pro Woche	58	29.3 (26.3, 32.4)	33.6 (31.3, 35.9)	4.7 (2.4, 7.0)	1.1 (-2.2, 4.3)	0.5135

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

<b>MOR-004</b>	<b>N</b>	<b>3MSC Baseline (Stufen/min) LSM (95% KI)</b>	<b>3MSC 24 Wochen (Stufen/min) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (Stufen/min) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (Stufen/min) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test<sup>3</sup>) BMN 110 vs. Placebo</b>
<b>ANOVA: Behandlung als einziger Term (ITT)</b>						
Placebo	59	30.0 (26.0, 34.0)	33.6 (29.0, 38.3)	3.6 (1.3, 6.0)	-	-
BMN 110 qow	59	27.1 (23.2, 31.1)	30.4 (25.7, 35.1)	3.2 (0.9, 5.5)	-0.4 (-3.7, 2.8)	0.7950
BMN 110 pro Woche	58	29.6 (25.6, 33.6)	34.3 (29.6, 39.0)	4.7 (2.4, 7.1)	1.1 (-2.2, 4.4)	0.5116
<b>ANCOVA: Nur Erster 3MSC (ITT)<sup>1</sup></b>						
Placebo	59	30.3 (27.3, 33.3)	31.8 (29.2, 34.4)	3.6 (1.1, 6.2)	-	-
BMN 110 qow	59	27.4 (24.4, 30.4)	31.5 (28.8, 34.1)	3.3 (0.8, 5.9)	-0.3 (-3.9, 3.3)	0.8695
BMN 110 pro Woche	58	28.1 (25.1, 31.2)	34.3 (31.6, 37.0)	6.1 (3.5, 8.7)	2.5 (-1.1, 6.1)	0.1773
<b>ANCOVA: Nur zweiter 3MSC (ITT)<sup>1</sup></b>						
Placebo	59	29.8 (26.6, 33.1)	32.4 (29.8, 35.0)	3.8 (1.3, 6.3)	-	-
BMN 110 qow	59	27.2 (24.0, 30.5)	31.4 (28.8, 34.1)	2.9 (0.3, 5.4)	-1.0 (-4.5, 2.6)	0.5952
BMN 110 pro Woche	58	30.5 (27.3, 33.8)	32.1 (29.5, 34.8)	3.6 (1.0, 6.1)	-0.3 (-3.8, 3.3)	0.8902
<b>ANCOVA: Bester 3MSC (ITT)<sup>1</sup></b>						
Placebo	59	32.1 (28.8, 35.3)	34.3 (31.8, 36.8)	3.9 (1.4, 6.4)	-	-
BMN 110 qow	59	29.1 (25.8, 32.3)	33.6 (31.0, 36.2)	3.2 (0.7, 5.7)	-0.7 (-4.2, 2.8)	0.6996
BMN 110 pro Woche	58	31.9 (28.6, 35.2)	34.3 (31.7, 36.9)	3.9 (1.4, 6.4)	-0.0 (-3.5, 3.5)	0.9989

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

MOR-004	N	3MSC Baseline (Stufen/min) LSM (95% KI)	3MSC 24 Wochen (Stufen/min) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (Stufen/min) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (Stufen/min) BMN 110 vs. Placebo LSM (95% KI)	p-Wert (t-Test <sup>3</sup> ) BMN 110 vs. Placebo
<b>ANCOVA: 6MWT als kontinuierliche Baseline-Kovariate (ITT)<sup>1</sup></b>						
Placebo	59	30.1 (27.0, 33.1)	32.4 (30.0, 34.8)	3.6 (1.3, 5.9)	-	-
BMN 110 qow	59	27.3 (24.3, 30.3)	31.9 (29.5, 34.4)	3.1 (0.8, 5.5)	-0.4 (-3.7, 2.8)	0.7901
BMN 110 pro Woche	58	29.3 (26.3, 32.4)	33.7 (31.2, 36.2)	4.9 (2.5, 7.2)	1.3 (-2.0, 4.6)	0.4377
<b>ANCOVA: Interaktion Behandlung und Ausgangswert im 3MSC (ITT)</b>						
Placebo	59	30.1 (27.0, 33.1)	32.1 (29.7, 34.4)	3.5 (1.2, 5.8)	-	-
BMN 110 qow	59	27.3 (24.3, 30.3)	31.6 (29.2, 34.0)	3.0 (0.7, 5.3)	-0.5 (-3.7, 2.8) <sup>3</sup>	0.7731
BMN 110 pro Woche	58	29.3 (26.3, 32.4)	33.4 (31.0, 35.8)	4.8 (2.5, 7.1)	1.3 (-1.9, 4.6) <sup>3</sup>	0.4262
BMN 110 qow					Interaktion Be- handlung / Base- line: -0.2 (-0.5, -0.0)	0.0305
BMN 110 pro Woche					-0.1 (-0.4, 0.1)	0.2312
<b>ANCOVA: Ausschluss eines Ausreißers (ITT)<sup>1</sup></b>						
Placebo	59	27.3 (24.2, 30.4)	32.1 (29.9, 34.4)	3.3 (1.0, 5.5)	-	-
BMN 110 qow	58	24.7 (21.5, 27.8)	31.0 (28.7, 33.3)	2.1 (-0.2, 4.4)	-1.2 (-4.2, 1.9)	0.4544
BMN 110 pro Woche	58	26.6 (23.4, 29.7)	33.3 (31.0, 35.6)	4.4 (2.1, 6.7)	1.2 (-1.9, 4.2)	0.4566

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

MOR-004	N	3MSC Baseline (Stufen/min) LSM (95% KI)	3MSC 24 Wochen (Stufen/min) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (Stufen/min) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (Stufen/min) BMN 110 vs. Placebo LSM (95% KI)	p-Wert (t-Test <sup>3</sup> ) BMN 110 vs. Placebo
<b>ANCOVA: PP-Analyse<sup>1</sup></b>						
Placebo	55	30.4 (27.4, 33.5)	32.6 (30.2, 34.9)	3.7 (1.3, 6.0)	-	-
BMN 110 qow	55	27.1 (24.1, 30.2)	32.0 (29.7, 34.4)	3.1 (0.8, 5.5)	-0.5 (-3.9, 2.8)	0.7547
BMN 110 pro Woche	52	29.2 (26.1, 32.4)	33.7 (31.2, 36.1)	4.7 (2.3, 7.2)	1.1 (-2.3, 4.5)	0.5262
<b>Observed cases<sup>1</sup></b>						
Placebo	59	30.1 (27.0, 33.1)	32.7 (30.4, 35.0)	3.7 (1.4, 6.0)	-	-
BMN 110 qow	59	27.3 (24.3, 30.3)	32.5 (30.2, 34.8)	3.4 (1.1, 5.7)	-0.2 (-3.5, 3.0)	0.8853
BMN 110 pro Woche	58	29.3 (26.3, 32.4)	33.9 (31.5, 36.2)	4.8 (2.5, 7.2)	1.2 (-2.1, 4.4)	0.4885

<sup>1</sup> ANCOVA (Änderung von Baseline bis Woche 24) mit den Kovariaten Behandlung, Altersgruppen, 6MWT-Kategorien und 3MSC (kontinuierlich) zu Beginn der Studie.

<sup>2</sup> ANCOVA (Änderung von Baseline bis Woche 24) mit den Kovariaten Behandlung, Visite, Behandlung - Visite, Altersgruppen, 6MWT-Kategorien und kontinuierlicher 3MSC zu Beginn der Studie.

<sup>3</sup> p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell. Nach Hochberg-Adjustierung, um der Multiplizität Rechnung zu tragen, wurden die Ergebnisse als signifikant betrachtet, wenn beide p-Werte <0.05 oder ein p-Wert <0.025 lagen.

Tabelle 4-31: Ergebnisse für den Endpunkt „3MSCT“ (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel: Einfluss zusätzlicher Kovariaten (ITT-Population)

Kovariaten Statistik	Kovariaten-Effect	BMN 110 2.0 mg/kg/qow <sup>1</sup> vs. Placebo	BMN110 2.0 mg/kg/Woche <sup>1</sup> vs. Placebo
Interaktion Altersgruppe (12-18 Jahre) und Baseline 6MWT-Kategorie Schätzer (SE) p-Wert	1.02 (3.36) 0.7617	-0.47 (1.67) 0.7794	1.14 (1.67) 0.4970
Interaktion Altersgruppe (≥19 Jahre) und Baseline 6MWT-Kategorie Schätzer (SE) p-Wert	0.72 (3.73) 0.8476		
Geschlecht Schätzer (SE) p-Wert	-1.88 (1.38) 0.1739	-0.69 (1.66) 0.6774	1.18 (1.66) 0.4781
Baseline 3MSC (Stufen/min) Schätzer (SE) p-Wert	-0.04 (0.06) 0.4553	-0.47 (1.66) 0.7783	1.14 (1.66) 0.4935
Baseline MVV Schätzer (SE) p-Wert	0.02 (0.04) 0.6215	-0.47 (1.66) 0.7784	1.23 (1.68) 0.4637
Baseline FVC Schätzer (SE) p-Wert	0.72 (1.12) 0.5187	-0.47 (1.66) 0.7787	1.28 (1.68) 0.4467
Niedriges uKS Schätzer (SE) p-Wert	-2.39 (1.65) 0.1507	-0.46 (1.65) 0.7828	1.17 (1.66) 0.4811
Baseline uKS Schätzer (SE) p-Wert	0.00 (0.05) 0.9540	-0.47 (1.67) 0.7776	1.14 (1.67) 0.4939
Klinisch signifikant abnormales EKG (Baseline) Schätzer (SE) p-Wert	-11.0 (9.24) 0.2334	-0.43 (1.66) 0.7934	1.36 (1.67) 0.4180
Klinisch signifikant abnormales Echokardiogramm (Baseline) Schätzer (SE) p-Wert	-3.43 (4.21) 0.4165	-0.34 (1.67) 0.8369	1.32 (1.68) 0.4315

Kovariaten Statistik	Kovariaten-Effect	BMN 110 2.0 mg/kg/qow <sup>1</sup> vs. Placebo	BMN110 2.0 mg/kg/Woche <sup>1</sup> vs. Placebo
Score Selbstpflege-Domäne			
Schätzer (SE)	-0.68 (0.30)	-0.39 (1.64)	1.21 (1.64)
p-Wert	0.0228	0.8114	0.4635
Score Mobilitätsdomäne			
Schätzer (SE)	-0.90 (0.29)	0.03 (1.62)	1.23 (1.62)
p-Wert	0.0019	0.9845	0.4473
Normalisierte, standardisierte Größe (z-Score) entsprechend CDC			
Schätzer (SE)	0.37 (0.32)	-0.02 (1.72)	1.45 (1.71)
p-Wert	0.2571	0.9919	0.3992
Standgröße			
Schätzer (SE)	0.05 (0.05)	-0.09 (1.72)	1.41 (1.72)
p-Wert	0.3379	0.9587	0.4112

<sup>1</sup> Schätzer des Behandlungseffekts (BMN 110 vs. Placebo) auf den 3MSCT und den p-Wert, adjustiert für jeweils eine zusätzliche Kovariate.

Die Mehrzahl der Patienten, war in der Lage, an beiden Tests zu allen Untersuchungszeitpunkten teilzunehmen. Drei Patienten der Placebogruppe (5.1%) konnten an beiden Tests zur Eingangsvisite und in Woche 24 nicht teilnehmen. In der BMN 110 qow Gruppe lagen für jeweils einen Patienten (1.7%) zur Eingangsvisite bzw. in Woche 24 keine Testergebnisse vor. Drei der Patienten, die einmal wöchentlich mit BMN 110 behandelt wurden, führten die Tests zur Eingangsvisite nicht aus (5.2%) und von diesem nahmen zwei Patienten (3.4%) nicht an den Tests in Woche 24 teil (der dritte Patient nahm an dem Test nach vorzeitiger Beendigung nicht teil). Abgesehen von einem Patienten (anderer Grund, Woche 24 BMN 110 jede zweite Woche) waren alle Patienten körperlich in der Lage, die Treppe zu ersteigen. Nur wenige Patienten schlossen den 3MSCT nicht ab und die Zahl dieser Patienten war in den Gruppen vergleichbar.

Zu Beginn der Studie war die Ausdauer im 3MSCT (Stufen/min, Mittelwert  $\pm$  SD) in der Gruppe, die BMN 110 alle zwei Wochen erhielten, geringfügig niedriger ( $27.1 \pm 2.01$ ) als in den Vergleichsgruppen ( $30.0 \pm 2.01$  Placebo,  $29.6 \pm 2.03$  BMN 110 wöchentlich). Verglichen mit Placebo, besserte sich innerhalb von 24 Wochen die Ausdauer der Patienten der BMN 110 2.0 mg/kg/Woche Gruppe im 3MSCT numerisch allerdings nicht signifikant. Die mittlere absolute Änderung der Zahl der in Woche 24 erklommenen Stufen ( $\pm$  SD) lag bei  $3.6 \pm 8.5$  Stufen/min unter Placebo ( $11.4 \pm 24.0\%$ ),  $3.4 \pm 10.2$  Stufen/min unter BMN 110 2.0 mg/kg/qow ( $19.7 \pm 49.2\%$ ) und  $4.8 \pm 8.1$  Stufen/min unter BMN 110 2.0 mg/kg/Woche ( $25.7 \pm 46.1\%$ ). Der geschätzte Behandlungseffekt (LSM, 95% KI) in Woche 24 betrug verglichen mit Placebo 1.1 Stufen/min (-2.1, 4.4;  $p=0.4935$ ) bei einmal wöchentlicher Behandlung mit BMN 110 und -0.5 Stufen/min (-3.7, 2.8;  $p=0.7783$ ) bei Behandlung alle

zwei Wochen (ITT, ANCOVA mit den Kovariaten Altersgruppen, 6MWT-Kategorie und kontinuierlicher 3MSCT zu Beginn der Studie). In Woche 12 war die Änderung der Zahl der erklommenen Stufen in den BMN 110-Gruppen numerisch höher als in der Placebo-Gruppe (Placebo: 2.9 Stufen/min [10.1%]; BMN 110 2.0 mg/kg/qow: 3.6 Stufen/min [22.7%]; BMN 110 2.0 mg/kg/Woche: 3.6 Stufen/min [16.6%]) und bei den Patienten, die einmal wöchentlich mit BMN 110 behandelt wurden, besserte sich die Ausdauer bis Woche 24 weiter.

Da nur einzelne Werte fehlten, war das Ergebnis der Primäranalyse und der „observed case“-Analyse (Änderung in Woche 24 vs. Placebo, LSM, 95% KI: -0.2, -3.5, 3.0,  $p=0.8853$  BMN 110 qow, 1.2, -2.1, 4.4,  $p=0.4885$  BMN 110 pro Woche) sowie der PP-Analyse ähnlich. Die Sensitivitätsanalysen bestätigten das Ergebnis der Primäranalyse und die Ergebnisse waren gleichgerichtet. In keiner der Analysen zeigten sich signifikante Unterschiede zwischen BMN 110 und Placebo (Tabelle 4-30). Die Patienten, die BMN 110 einmal wöchentlich erhielten, zeigten eine deutlichere Besserung im 3MSCT, wenn nur der erste Test in die Analyse einbezogen wurde (6.5 vs. 3.6 Stufen/min im zweiten Test). Mittels Streudiagramm der studentisierten Residuen wurde ein Ausreißer in der BMN 110 qow Gruppe identifiziert. Bei diesem Patienten hatte sich die Ausdauer in Woche 24 um 45.8 Stufen/min gesteigert. Der Ausschluss des Ausreißers änderte das Ergebnis der Primäranalyse nicht. Auch der Einschluss zusätzlicher Kovariaten änderte das Ergebnis der Primäranalyse nicht (Tabelle 4-31).

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

Nicht zutreffend.

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.

#### **4.3.1.3.4 KS im Urin – 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 0). 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 4-32: Operationalisierung des Endpunkts „uKS“

Studie	Operationalisierung
MOR-004	<p>Der Kreatinin-normalisierte KS-Gehalt im ersten Morgenurin wurde quantitativ bestimmt. Urinproben zur Bestimmung des KS wurden während der Eingangsuntersuchung (Mittelwert aus zwei Messungen an 2 d), in Woche 2 und 4 und danach alle 4 Wochen sowie innerhalb einer Woche nach einem möglichen Studienabbruch (jeweils Einzelproben) entnommen. Der normalisierte KS Gehalt wurde aus dem Quotienten aus KS im Urin und Kreatinin im Urin gebildet. Die uKS-Gehalte wurden in einem zentralen Labor in den USA bestimmt.</p> <p>Da der uKS-Gehalt bei Morquio-Patienten stark vom Alter abhängt, wurde die prozentuale Änderung in Woche 24 in der primären Wirksamkeitsanalyse ausgewertet. Die absolute und prozentuale Änderung des normalisierten uKS in Woche 24 verglichen mit Baseline wurde mittels eines ANCOVA-Modells mit der kontinuierlichen Kovariate normalisiertes uKS bei der Baseline-Visite und den Faktoren Behandlung, 6MWT-Stratifizierung (<math>\leq 200</math> m, <math>&gt; 200</math> m) und Altersgruppen (5-11, 12-18, <math>\geq 19</math> Jahre) analysiert. Die BMN 110-Gruppen wurden jeweils mit Placebo verglichen und signifikante Gruppenunterschiede mittels t-Test im ANCOVA-Modell evaluiert.</p> <p>Um der Multiplizität Rechnung zu tragen wurde hinsichtlich der sekundären Endpunkte stufenweise vorgegangen: Die Ergebnisse des 3MSCT wurden zuerst ausgewertet und die Ergebnisse des normalisierten uKS durften nur dann als signifikant bezeichnet werden, wenn das Ergebnis des 3MSCT signifikant war.</p> <p>Der Endpunkt uKS ist verwendet jedoch einen Surrogatparameter, aus diesem Grund ist dieser Endpunkt nicht direkt bewertungsrelevant.</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 4-33: Bewertung des Verzerrungspotenzials für uKS 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
MOR-004 (9)	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt uKS wird unter Berücksichtigung der in Tabelle 4-33 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Die Gründe für diese Einschätzung sind in Absatz 4.3.1.3.1 dargestellt. Da es sich bei dem Endpunkt uKS um einen objektiven Parameter handelt, kann ausgeschlossen werden, dass eine potentielle Entblindung einzelner

Patienten einen Einfluss auf das Ergebnis gehabt haben könnte. Im Protokoll war ein stufenweises Vorgehen spezifiziert, nach dem die uKS-Ergebnisse nur dann als signifikant bezeichnet werden durften, wenn die Ergebnisse des 3MSCT signifikant waren, was nicht der Fall war. Es wird allerdings nicht angenommen, dass dies die sehr deutlichen und hoch signifikanten Unterschiede des uKS zwischen BMN 110 und Placebo verzerrt haben könnte.

*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 4-34: Ergebnisse für den Endpunkt „uKS“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik)

<b>MOR 004</b>	<b>n</b>	<b>uKS Baseline (µg/mg) Mittelwert (SD)</b>	<b>uKS 2 Wochen (µg/mg) Mittelwert (SD)</b>	<b>uKS 24 Wochen (µg/mg) Mittelwert (SD)</b>	<b>Änderung von Baseline bis Woche 24 (µg/mg) Mittelwert (SD)</b>	<b>Änderung von Baseline bis Woche 24 (%) Mittelwert (SD)</b>
Placebo	59	26.1 (15.43)	26.5 (15.18)	23.5 (13.55)	-2.7 (7.84)	-3.6 (27.41)
BMN 110 qow	59	28.6 (21.17)	17.8 (10.54)	16.5 (10.18)	-12.1 (16.03)	-35.3 (20.74)
BMN 110 pro Woche	58	26.9 (14.11)	17.2 (9.72)	14.5 (8.35)	-12.4 (9.73)	-43.7 (22.29)

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

Tabelle 4-35: Ergebnisse für den Endpunkt „uKS“ (normalisiert, prozentuale Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (Primär- und Sensitivitätsanalysen)

MOR-004	N	uKS Baseline (µg/mg) LSM (95% KI)	uKS 24 Wochen (µg/mg) LSM (95% KI)	Änderung von Baseline bis Woche 24 (%) LSM (95% KI)	Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)	p-Wert (t-Test <sup>3</sup> ) BMN 110 vs. Placebo
<b>ANCOVA (ITT, Primäranalyse)<sup>1</sup></b>						
Placebo	59	26.9 (23.5, 30.3)	22.4 (20.6, 24.3)	-3.9 (-9.9, 2.2)	-	-
BMN 110 qow	59	28.6 (25.2, 32.0)	14.4 (12.6, 16.2)	-34.1 (-39.8, -28.4)	-30.2 (-38.5, -22.0)	<0.0001
BMN 110 pro Woche	58	26.2 (22.7, 29.6)	12.7 (10.7, 14.6)	-44.6 (-50.8, -38.4)	-40.7 (-49.0, -32.4)	<0.0001
<b>ANCOVA mit Messwiederholung (ITT)<sup>2</sup></b>						
Placebo	59	26.5 (23.0, 29.9)	23.8 (22.1, 25.5)	-4.8 (-10.1, 0.5)	-	-
BMN 110 qow	59	28.6 (25.2, 32.0)	15.9 (14.2, 17.6)	-34.3 (-39.5, -29.1)	-29.5 (-37.0, -22.1)	<0.0001
BMN 110 pro Woche	58	26.1 (22.7, 29.6)	14.2 (12.4, 15.9)	-45.0 (-50.3, -39.6)	-40.2 (-47.7, -32.7)	<0.0001
<b>ANCOVA: Gepoolter mittlerer Placebowert (ITT)<sup>1</sup></b>						
Placebo	59	26.4 (23.0, 29.8)	24.1 (22.3, 26.0)	-3.6 (-9.1, 2.0)	-	-
BMN 110 qow	59	28.6 (25.2, 32.0)	16.2 (14.4, 18.1)	-32.9 (-38.5, -27.3)	-29.3 (-37.2, -21.4)	<0.0001
BMN 110 pro Woche	58	26.2 (22.7, 29.6)	14.9 (13.0, 16.7)	-43.1 (-48.7, -37.5)	-39.5 (-47.4, -31.6)	<0.0001
<b>ANCOVA: LOCF-Imputation (ITT)<sup>1</sup></b>						
Placebo	59	22.27 (18.77, 25.78)	22.03 (20.23, 23.83)	-9.18 (-14.6, -3.72)	-	-
BMN 110 qow	59	24.28 (20.78, 27.78)	14.35 (12.58, 16.12)	-38.1 (-43.4, -32.7)	-28.9 (-36.3, -21.5)	<0.0001
BMN 110 pro Woche	58	21.83 (18.27, 25.39)	12.40 (10.57, 14.23)	-49.3 (-54.9, -43.7)	-40.1 (-47.5, -32.7)	<0.0001

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

<b>MOR-004</b>	<b>N</b>	<b>uKS Baseline (µg/mg) LSM (95% KI)</b>	<b>uKS 24 Wochen (µg/mg) LSM (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (%) LSM (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test<sup>3</sup>) BMN 110 vs. Placebo</b>
<b>ANOVA mit Behandlung als einziger Kovariate (ITT)</b>						
Placebo	59	26.1 (21.7, 30.6)	23.5 (20.7, 26.3)	-3.6 (-10.2, 3.1)	-	-
BMN 110 qow	59	28.6 ( , )	16.5 (13.7, 19.3)	-35.3 (-41.7, -29.0)	-31.8 (-40.9, -22.6)	<.0001
BMN 110 pro Woche	58	26.9 ( , )	14.5 (11.6, 17.4)	-43.7 (-50.4, -36.9)	-40.1 (-49.3, -30.9)	<.0001
<b>ANCOVA: 6MWT als kontinuierliche Baseline-Kovariate (ITT)<sup>1</sup></b>						
Placebo	59	26.9 (23.5, 30.3)	22.8 (20.9, 24.6)	-4.1 (-10.2, 2.0)	-	-
BMN 110 qow	59	28.6 (25.2, 32.0)	14.9 (13.0, 16.8)	-34.1 (-39.8, -28.5)	-30.0 (-38.3, -21.7)	<0.0001
BMN 110 pro Woche	58	26.2 (22.7, 29.6)	13.3 (11.2, 15.3)	-44.3 (-50.5, -38.1)	-40.2 (-48.6, -31.9)	<0.0001
<b>ANCOVA: PP-Analyse<sup>1</sup></b>						
Placebo	55	26.9 (23.7, 30.1)	23.1 (21.4, 24.8)	-3.2 (-9.5, 3.1)	-	-
BMN 110 qow	55	28.1 (24.9, 31.2)	15.3 (13.6, 17.0)	-35.1 (-41.0, -29.3)	-31.9 (-40.5, -23.4)	<0.0001
BMN 110 pro Woche	52	25.7 (22.5, 29.0)	13.2 (11.4, 15.0)	-44.0 (-50.2, -37.9)	-40.8 (-49.6, -32.1)	<0.0001
<b>Observed cases<sup>1</sup></b>						
Placebo	59	26.5 (23.1, 30.0)	24.3 (22.6, 26.1)	-4.5 (-9.8, 0.8)	-	-
BMN 110 qow	59	28.5 (25.1, 31.9)	16.0 (14.3, 17.8)	-34.1 (-39.3, -28.9)	-29.6 (-37.1, -22.2)	<0.0001
BMN 110 pro Woche	58	26.1 (22.7, 29.5)	14.1 (12.3, 15.9)	-46.2 (-51.6, -40.9)	-41.8 (-49.3, -34.2)	<0.0001

<sup>1</sup> ANCOVA (Änderung von Baseline bis Woche 24) mit den Kovariaten Behandlung, Altersgruppen und 6MWT-Kategorien zu Beginn der Studie.<sup>2</sup> ANCOVA mit den Faktoren Behandlung, Visite, Behandlung-Visite, Altersgruppen und 6MWT-Kategorien zu Beginn der Studie<sup>3</sup> p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell. Nach Hochberg-Adjustierung, um der Multiplizität Rechnung zu tragen, wurden die Ergebnisse als signifikant betrachtet, wenn beide p-Werte <0.05 oder ein p-Wert <0.025 lagen.

Tabelle 4-36: Ergebnisse für den Endpunkt „uKS“ (normalisiert, prozentuale Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel: Einfluss zusätzlicher Kovariaten (ITT-Population)

Kovariaten Statistik	Kovariaten-Effect	BMN 110 2.0 mg/kg qow <sup>1</sup> vs. Placebo	BMN110 2.0 mg/kg/Woche <sup>1</sup> vs. Placebo
Interaktion Altersgruppe (12-18 Jahre) und Baseline 6MWT-Kategorie			
Schätzer (SE)	-0.02 (0.08)	-0.30 (0.04)	-0.41 (0.04)
p-Wert	0.8296	<0.0001	<0.0001
Interaktion Altersgruppe (≥19 Jahre) und Baseline 6MWT-Kategorie			
Schätzer (SE)	0.02 (0.10)		
p-Wert	0.8502		
Geschlecht			
Schätzer (SE)	-2.88 (3.55)	-30.6 (4.18)	-40.6 (4.20)
p-Wert	0.4184	<0.0001	<0.0001
Baseline 3MSC (Stufen/min)			
Schätzer (SE)	0.11 (0.15)	-30.0 (4.21)	-40.6 (4.20)
p-Wert	0.4798	<0.0001	<0.0001
Baseline MVV			
Schätzer (SE)	0.00 (0.11)	-30.2 (4.20)	-40.7 (4.25)
p-Wert	0.9953	<0.0001	<0.0001
Baseline FVC			
Schätzer (SE)	1.68 (3.06)	-30.3 (4.19)	-40.3 (4.24)
p-Wert	0.5837	<0.0001	<0.0001
Niedriges uKS			
Schätzer (SE)	-8.92 (4.96)	-30.4 (4.16)	-40.4 (4.17)
p-Wert	0.0744	<0.0001	<0.0001
Baseline uKS			
Schätzer (SE)	-0.86 (0.13)	-30.2 (4.19)	-40.7 (4.20)
p-Wert	<0.0001	<0.0001	<0.0001
Klinisch signifikant abnormales EKG (Baseline)			
Schätzer (SE)	-25.0 (22.77)	-30.2 (4.18)	-40.2 (4.23)
p-Wert	0.2733	<0.0001	<0.0001
Klinisch signifikant abnormales Echokardiogramm (Baseline)			
Schätzer (SE)	-13.4 (10.34)	-29.8 (4.19)	-40.0 (4.24)
p-Wert	0.1976	<0.0001	<0.0001

Kovariaten Statistik	Kovariaten-Effect	BMN 110 2.0 mg/kg qow <sup>1</sup> vs. Placebo	BMN110 2.0 mg/kg/Woche <sup>1</sup> vs. Placebo
Score Selbstpflege-Domäne			
Schätzer (SE)	0.07 (0.73)	-30.3 (4.20)	-40.7 (4.21)
p-Wert	0.9227	<0.0001	<0.001
Score Mobilitätsdomäne			
Schätzer (SE)	-0.28 (0.66)	-30.0 (4.24)	-40.6 (4.21)
p-Wert	0.6674	<0.0001	<0.0001
Normalisierte, standardisierte Größe (z-Score) entsprechend CDC			
Schätzer (SE)	0.84 (0.87)	-28.7 (4.27)	-38.7 (4.26)
p-Wert	0.3356	<0.0001	<0.0001
Standgröße			
Schätzer (SE)	0.10 (0.15)	-28.9 (4.27)	-38.8 (4.27)
p-Wert	0.5045	<0.0001	<0.0001

<sup>1</sup> Schätzer des Behandlungseffekts (BMN 110 vs. Placebo) auf den uKS und den p-Wert, adjustiert für jeweils eine zusätzliche Kovariate.

Von allen Patienten konnten zu jedem der vorgegebenen Messzeitpunkte die uKS-Werte bestimmt werden. Die mittleren absoluten und mittleren prozentualen Abnahmen des normalisierten uKS vom Ausgangswert bis Woche 24 waren in beiden BMN 110-Behandlungsgruppen ausgeprägter als unter Placebo. Die mittlere Veränderungen ( $\pm$  SD) in der ITT-Population von Baseline bis Woche 24 betrugen -2.8  $\mu$ g/mg ( $\pm$ 8.0) für Placebo, -12.2  $\mu$ g/mg ( $\pm$ 16.3) für BMN 110 2,0 mg/kg/qow und -12.6  $\mu$ g/mg ( $\pm$ 9.5) für BMN 110 2,0 mg/kg/Woche. Die mittlere prozentualen Veränderungen ( $\pm$  SD) von Baseline bis Woche 24 waren -4.4% ( $\pm$ 27,0) für Placebo, -35.2% ( $\pm$ 20.7) für BMN 110 2,0 mg/kg/qow und -45.1% ( $\pm$ 19.9) für BMN 110 2,0 mg/kg/Woche. Der geschätzte Behandlungseffekt in Woche 24 verglichen mit Placebo betrug -40.7% (95% KI -49.0, -32.4;  $p$ <0.0001) in der BMN 110 2.0 mg/kg/Woche Gruppe und -30.2% (95% KI -38.5, -22.0;  $p$ <0.0001) in der BMN 110 2.0 mg/kg/qow Gruppe (ITT; ANCOVA Modell adjustiert nach den Baseline-Kovariaten Altersgruppen, 6MWT-Kategorie und normalisiertes uKS). Die Ergebnisse der Sensitivitätsanalysen bestätigen die Robustheit des Ergebnisses der Primäranalyse und die Unterschiede zugunsten von BMN 110 verglichen mit Placebo waren in jedem Fall hoch signifikant (Tabelle 4-35). Zusätzliche Sensitivitätsanalysen, bei denen verschiedene Kovariaten einzeln in das ANCOVA-Modell eingegeben wurden, zeigten, dass das Ausmaß des Behandlungseffektes mit jeder zusätzlichen Kovariate der primären Analyse vergleichbar war (Tabelle 4-36).

Die uKS-Spiegel sanken in beiden BMN 110-Gruppen bereits innerhalb von zwei Wochen deutlich ab und die Werte blieben in beiden Gruppen über den Verlauf der Studie weitgehend

stabil, mit Ausnahme eines temporären Anstiegs in der BMN 110 qow Gruppe in Woche 12 (Abbildung 11). Dies war auf einen Ausreißer zurückzuführen.

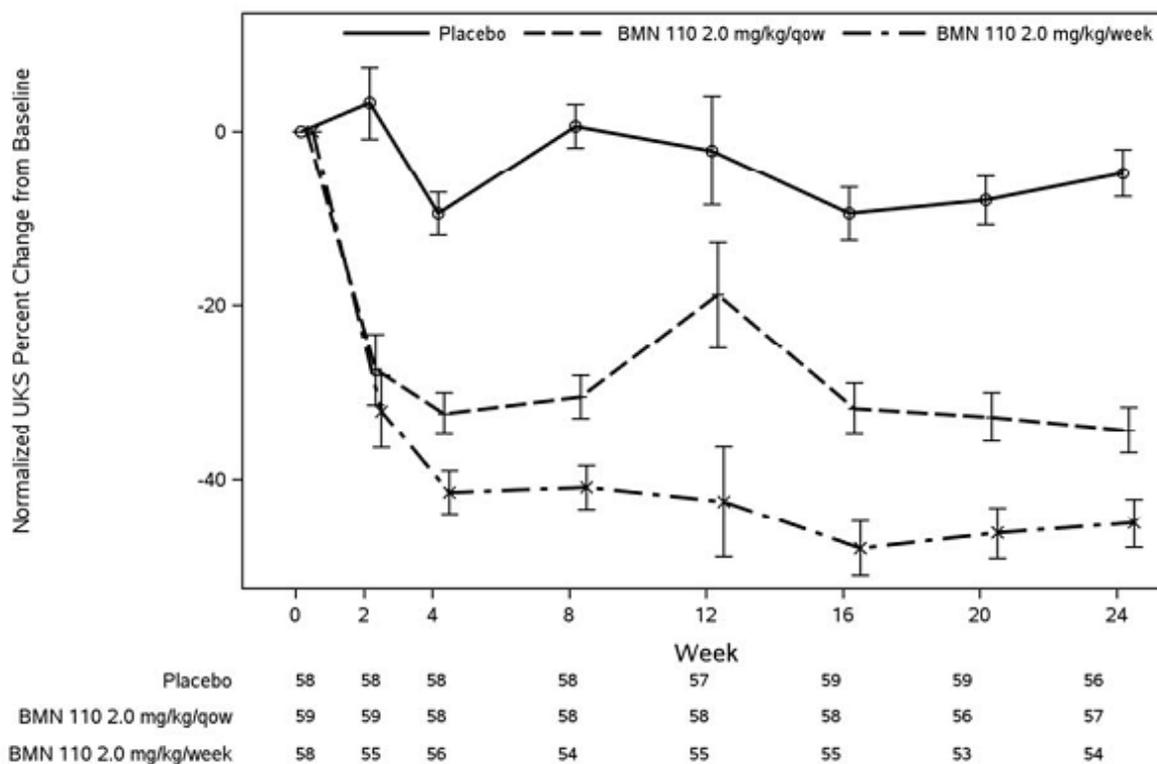


Abbildung 11: Modell-basierte ANCOVA mit Messwiederholung der prozentualen Veränderung des normalisierten uKS (ITT)

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

Nicht zutreffend.

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.

#### 4.3.1.3.5 Respiratorische Funktion – 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 0). 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 4-37: Operationalisierung des Endpunktes „respiratorische Funktion“

Studie	Operationalisierung
MOR-004	<p>Die respiratorische Funktion (FET, FEV1, FIVC, FVC und MVV) wurde in Übereinstimmung mit den Standards der ATS (16) während der Eingangsuntersuchung, in Woche 24 sowie innerhalb einer Woche nach einem möglichen Studienabbruch gemessen. Der MVV war als zusätzlicher unterstützender Wirksamkeitsparameter und die übrigen respiratorischen Parameter als tertiäre Endpunkte im Studienprotokoll spezifiziert. Die respiratorischen Parameter waren folgendermaßen definiert:</p> <p>MVV ("Maximum Voluntary Ventilation"): Volumen der ausgeatmeteten Luft in l/min während wiederholter maximaler respiratorischer Anstrengungen. Der Patient wurde angehalten, über nicht weniger als 12 sec und nicht länger als 15 sec so tief und so schnell wie möglich zu atmen.</p> <p>FVC ("Forced Vital Capacity"): Maximales Luftvolumen (l), welches nach maximaler Einatmung auf einmal forciert ausgeatmet werden kann</p> <p>FEV1 ("Forced Expiratory Volume in 1 Second"): Maximales innerhalb der ersten Sekunde nach maximaler Einatmung forciert ausgeatmetes Luftvolumen (l)</p> <p>FIVC ("Forced Inspiratory Vital Capacity"): Maximales Luftvolumen (l), welches nach maximaler Ausatmung auf einmal forciert eingeatmet werden kann</p> <p>FET ("Forced ExpiratoryTime"): Zeit, die benötigt wird, um einen tiefen Atemzug komplett auszuatmen (d.h. FET100% ist die Zeit, die benötigt wird, um FVC zu erreichen).</p> <p>Die prozentuale Änderung des Ausgangs-MVV bis Woche 24 wurde mittels ANCOVA mit dem Ausgangs-MVV als kontinuierliche Variable und den Faktoren Behandlungsgruppe, Altersgruppen (5-11, 12-18, <math>\geq 19</math> Jahre) und 6MWT-Stratifizierung (<math>\leq 200</math> m, <math>&gt;200</math> m) analysiert. Es wurden die BMN 110-Gruppen mit Placebo verglichen und Gruppenunterschiede mittels t-Test analysiert. Für die übrigen respiratorischen Parameter wurde die prozentuale Änderung bis Woche 24 in einem der Primäranalyse ähnlichen ANCOVA-Modell ausgewertet. Der einzige Unterschied bestand darin, dass eine zusätzliche Kovariate (Ausgangswert als kontinuierliche Variable) berücksichtigt wurde.</p> <p>Bei Patienten, die gestorben waren oder die physisch nicht in der Lage waren, an den Respirationstests teilzunehmen, wurde unabhängig vom Behandlungsarm oder der Studienwoche der schlechteste im Datensatz vorhandene Wert eingegeben. In der primären Analyse wurde multiple Imputation eingesetzt, wenn Werte aus anderen Gründen fehlten. Da nach der Eingangsuntersuchung bis Woche 24 keine Messungen vorgesehen waren, wurde auf die LOCF-Imputation und die Analyse mit Messwiederholung verzichtet.</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 4-38: Bewertung des Verzerrungspotenzials für den Endpunkt „respiratorische Funktion“ 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
MOR-004 (9)	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt respiratorische Funktion wird unter Berücksichtigung der in Tabelle 4-38 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. Die Gründe für diese Einschätzung sind in Abschnitt 4.3.1.3.1 dargestellt.

*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 4-39: Ergebnisse für den Endpunkt „respiratorische Funktion“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik)

MOR 004	n	Baseline Mittelwert (SD)	24 Wochen Mittelwert (SD)	Absolute Änderung von Baseline bis Woche 24 Mittelwert (SD)	Änderung von Baseline bis Woche 24 (%) Mittelwert (SD)
<b>MVV (l/min)</b>					
Placebo	59	33.9 (25.74)	34.5 (27.65)	0.5 (8.13)	2.1 (19.52)
BMN 110 qow	59	32.1 (19.25)	33.8 (20.68)	1.7 (5.62)	6.6 (22.47)
BMN 110 pro Woche	58	27.5 (16.09)	29.2 (14.19)	1.7 (7.53)	14.5 (37.76)
<b>FVC (l)</b>					
Placebo	59	1.2 (0.85)	1.2 (0.86)	-0.0 (0.14)	1.2 (13.75)
BMN 110 qow	59	1.1 (0.65)	1.1 (0.67)	0.0 (0.10)	4.4 (11.70)
BMN 110 pro Woche	58	0.9 (0.50)	1.0 (0.50)	0.0 (0.09)	5.1 (11.71)

<b>MOR 004</b>	<b>n</b>	<b>Baseline Mittelwert (SD)</b>	<b>24 Wochen Mittelwert (SD)</b>	<b>Absolute Änderung von Baseline bis Woche 24 Mittelwert (SD)</b>	<b>Änderung von Baseline bis Woche 24 (%) Mittelwert (SD)</b>
<b>FEV1 (l)</b>					
Placebo	59	1.0 (0.69)	1.0 (0.69)	-0.0 (0.13)	2.5 (16.80)
BMN 110 qow	59	0.9 (0.52)	1.0 (0.55)	0.0 (0.10)	3.3 (13.35)
BMN 110 pro Woche	58	0.8 (0.42)	0.9 (0.41)	0.0 (0.08)	5.4 (11.54)
<b>FVC (l)</b>					
Placebo	59	1.0 (0.78)	1.3 (2.12)	0.3 (1.58)	13.9 (50.03)
BMN 110 qow	59	1.0 (0.60)	1.0 (0.60)	-0.0 (0.12)	-2.1 (19.08)
BMN 110 pro Woche	58	0.8 (0.49)	0.8 (0.43)	-0.0 (0.26)	7.0 (31.66)
<b>FET (sec)</b>					
Placebo	59	4.3 (2.95)	4.3 (2.40)	-0.0 (2.72)	21.8 (101.07)
BMN 110 qow	59	3.9 (2.22)	4.2 (2.40)	0.3 (1.28)	21.3 (69.02)
BMN 110 pro Woche	58	5.7 (7.65)	5.0 (4.99)	-0.7 (5.73)	54.4 (240.15)

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

Tabelle 4-40: Ergebnisse für den Endpunkt „MVV“ (mittleres Volumen der ausgeatmeteten Luft in l/min während wiederholter maximaler respiratorischer Anstrengungen von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (Primär- und Sensitivitätsanalysen)

<b>MOR-004</b>	<b>N</b>	<b>MVV Baseline (l/min) LSM (95% KI)</b>	<b>MVV 24 Wochen (l/min) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (%) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test<sup>1</sup>) BMN 110 vs. Placebo</b>
<b>ANCOVA (ITT, Primäranalyse)<sup>1</sup></b>						
Placebo	59	33.4 (28.1, 38.6)	31.5 (29.1, 33.9)	4.1 (-5.4, 13.5)	-	-
BMN 110 qow	59	32.2 (27.1, 37.3)	32.6 (30.4, 34.7)	7.5 (-1.1, 16.0)	3.4 (-9.9, 16.6)	0.6111
BMN 110 pro Woche	58	28.0 (22.8, 33.2)	32.5 (30.4, 34.6)	14.4 (4.6, 24.2)	10.3 (-1.8, 22.4)	0.0943
<b>ANCOVA Gepoolter mittlerer Placebowert (ITT)<sup>1</sup></b>						
Placebo	59	33.9 (28.9, 38.9)	31.7 (29.6, 33.7)	-0.1 (-6.6, 6.4)	-	-
BMN 110 qow	59	32.8 (27.8, 37.8)	32.9 (30.8, 34.9)	4.1 (-2.4, 10.5)	4.2 (-5.0, 13.3)	0.3685
BMN 110 pro Woche	58	29.2 (24.2, 34.2)	32.2 (30.1, 34.3)	5.5 (-1.1, 12.0)	5.6 (-3.7, 14.8)	0.2352
<b>ANOVA: Behandlung als einziger Term (ITT)</b>						
Placebo	59	33.9 (28.4, 39.5)	34.5 (28.3, 40.7)	3.3 (-6.1, 12.6)	-	-
BMN 110 qow	59	32.1 (26.7, 37.5)	33.8 (28.1, 39.4)	7.2 (-1.4, 15.9)	4.0 (-9.3, 17.2)	0.5523
BMN 110 pro Woche	58	27.5 (22.0, 33.0)	29.2 (23.4, 34.9)	15.5 (5.4, 25.5)	12.2 (-0.1, 24.5)	0.0515
<b>ANCOVA: 6MWT als kontinuierliche Baseline-Kovariate (ITT)</b>						
Placebo	59	33.4 (28.1, 38.6)	31.5 (29.0, 34.0)	4.4 (-4.9, 13.8)	-	-
BMN 110 qow	59	32.2 (27.1, 37.3)	32.6 (30.4, 34.7)	7.4 (-1.1, 15.9)	3.0 (-10.1, 16.1)	0.6518
BMN 110 pro Woche	58	28.0 (22.8, 33.2)	32.5 (30.3, 34.7)	14.1 (4.4, 23.8)	9.7 (-2.3, 21.7)	0.1120

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

MOR-004	N	MVV Baseline (l/min) LSM (95% KI)	MVV 24 Wochen (l/min) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (%) LSM (95% KI)	Absolute Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)	p-Wert (t-Test <sup>1</sup> ) BMN 110 vs. Placebo
<b>ANCOVA: Ausschluss eines Ausreißers (ITT)</b>						
Placebo	59	35.7 (30.4, 41.1)	31.7 (29.3, 34.1)	2.9 (-4.9, 10.7)	-	-
BMN 110 qow	59	34.5 (29.3, 39.7)	32.7 (30.6, 34.9)	6.3 (-0.9, 13.6)	3.5 (-7.9, 14.8)	0.5391 <sup>2</sup>
BMN 110 pro Woche	57	30.7 (25.3, 36.1)	32.5 (30.5, 34.6)	9.9 (2.5, 17.3)	7.0 (-2.8, 16.8)	0.1591 <sup>2</sup>
<b>ANCOVA: PP-Analyse</b>						
Placebo	55	33.1 (27.9, 38.4)	30.7 (28.5, 33.0)	4.2 (-5.6, 14.1)	-	-
BMN 110 qow	55	31.9 (26.9, 37.0)	32.2 (30.3, 34.1)	7.3 (-1.6, 16.2)	3.1 (-10.8, 17.0)	0.6577
BMN 110 pro Woche	52	27.1 (21.9, 32.4)	32.5 (30.7, 34.4)	15.4 (4.8, 25.9)	11.1 (-1.6, 23.9)	0.0871
<b>Observed cases</b>						
Placebo	59	33.5 (27.9, 39.2)	32.5 (30.3, 34.6)	3.0 (-3.5, 9.6)	-	-
BMN 110 qow	59	32.2 (26.7, 37.8)	33.4 (31.3, 35.5)	6.2 (-0.1, 12.6)	3.2 (-5.9, 12.3)	0.4904
BMN 110 pro Woche	58	28.5 (22.9, 34.1)	33.2 (31.0, 35.4)	10.1 (3.5, 16.8)	7.1 (-2.3, 16.4)	0.1362

ANCOVA (Änderung von Baseline bis Woche 24) mit den Kovariaten Behandlung, Altersgruppen, 6MWT-Kategorien und MVV (kontinuierlich) zu Beginn der Studie.

<sup>1</sup> p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell. Nach Hochberg-Adjustierung, um der Multiplizität Rechnung zu tragen, wurden die Ergebnisse als signifikant betrachtet, wenn beide p-Werte <0.05 oder ein p-Wert <0.025 lagen.

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

Tabelle 4-41: Ergebnisse für andere respiratorische Endpunkte aus RCT mit dem zu bewertenden Arzneimittel (ITT)

<b>MOR-004</b>	<b>N</b>	<b>Baseline LSM (95% KI)</b>	<b>24 Wochen LSM (95% KI)</b>	<b>Prozentuale Änderung von Baseline bis Woche 24 LSM (95% KI)</b>	<b>Prozentuale Änderung von Baseline bis Woche 24 BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test<sup>1</sup>) BMN 110 vs. Placebo</b>
<b>FVC (l)</b>						
Placebo	59	1.1 (1.0, 1.3)	1.1 (1.0, 1.1)	1.6 (-3.1, 6.3)	-	-
BMN 110 qow	59	1.1 (0.9, 1.3)	1.1 (1.1, 1.1)	4.6 (1.0, 8.1)	3.0 (-2.4, 8.3)	0.2751
BMN 110 pro Woche	58	0.9 (0.8, 1.1)	1.1 (1.1, 1.1)	4.9 (1.2, 8.5)	3.3 (-3.1, 9.6)	0.3041
<b>FEV1 (l)</b>						
Placebo	59	1.0 (0.8, 1.1)	0.9 (0.9, 1.0)	3.3 (-2.6, 9.1)	-	-
BMN 110 qow	59	0.9 (0.8, 1.1)	1.0 (0.9, 1.0)	3.5 (-1.0, 7.9)	0.2 (-7.4, 7.9)	0.9546
BMN 110 pro Woche	58	0.8 (0.7, 1.0)	1.0 (0.9, 1.0)	5.1 (1.0, 9.2)	1.8 (-5.5, 9.2)	0.6129
<b>FIVC (l)</b>						
Placebo	59	1.0 (0.8, 1.2)	1.2 (0.9, 1.4)	19.2 (-15.1, 53.5)	-	-
BMN 110 qow	59	1.0 (0.9, 1.2)	0.9 (0.7, 1.1)	-1.6 (-29.2, 25.9)	-20.9 (-63.6, 21.9)	0.3344
BMN 110 pro Woche	58	0.9 (0.7, 1.0)	1.0 (0.8, 1.3)	16.6 (-44.9, 78.1)	-2.6 (-65.7, 60.5)	0.9303
<b>FET (sec)</b>						
Placebo	59	4.2 (3.0, 5.4)	4.6 (3.9, 5.3)	21.1 (-18.7, 60.8)	-	-
BMN 110 qow	59	3.9 (2.7, 5.1)	4.7 (4.0, 5.4)	17.3 (-22.2, 56.9)	-3.7 (-59.8, 52.4)	0.8962
BMN 110 pro Woche	58	5.8 (4.5, 7.2)	4.8 (4.0, 5.5)	63.6 (22.5, 104.6)	42.5 (-14.9, 99.9)	0.1457

ANCOVA der prozentualen Änderung von Baseline unter Berücksichtigung der Kovariaten Behandlung, Altersklassen, 6MWT-Kategorie und des jeweiligen Ausgangswerts als kontinuierliche Variable.

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

<sup>1</sup> p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell. Nach Hochberg-Adjustierung, um der Multiplizität Rechnung zu tragen, wurden die Ergebnisse als signifikant betrachtet, wenn beide p-Werte  $<0.05$  oder ein p-Wert  $<0.025$  lagen.

Zu Beginn der Studie waren fünf (8.5%), vier (6.8%) bzw. zwei Patienten (3.4%) der Placebo-, BMN 110 2.0 mg/kg/qow bzw. BMN 110 2.0 mg/kg/Woche Gruppe nicht in der Lage, an den respiratorischen Tests teilzunehmen. Die korrespondierenden Zahlen in Woche 24 waren drei (5.1%), ein (1.7%) bzw. ein Patient (1.7%). Die Mehrzahl der Patienten konnte aufgrund fehlender Reife, nicht an den Tests teilnehmen. Die fehlenden Respirationstests wurden durch Imputation ersetzt.

Verglichen mit Placebo besserte sich das MVV bis Woche 24 numerisch in der BMN 110 2.0 mg/kg/Woche Gruppe und minimal unter der Behandlung mit BMN 110 alle zwei Wochen (Tabelle 4-40). Die mittlere prozentuale Änderung ( $\pm$  SD; ITT) des MVV bis Woche 24 betrug  $2.4 \pm 20.7\%$  in der Placebo-Gruppe,  $6.1 \pm 23.8\%$  in der BMN 110 2.0 mg/kg/qow Gruppe und  $10.8 \pm 25.6\%$  bei den Patienten, denen BMN 110 einmal wöchentlich injiziert wurde (Tabelle 4-39). Die Variabilität zwischen den Patienten war sowohl zu Beginn der Studie als auch nach 24 Wochen hoch und war zwischen den Gruppen vergleichbar. Der geschätzte Behandlungseffekt verglichen mit Placebo betrug  $10.3\%$  (95% KI -1.8, 22.4;  $p=0.0943$ ) in der BMN 110 2.0 mg/kg/Woche Gruppe und  $3.4\%$  (95% KI -9.9, 16.6;  $p=0.6111$ ) in der BMN 110 2.0 mg/kg/qow Gruppe (ITT, ANCOVA unter Berücksichtigung der Kovariaten Altersgruppen, 6MWT-Kategorie und MVV zu Beginn der Studie). Die Ergebnisse der Sensitivitätsanalysen waren gleichgerichtet und zeigten keine signifikanten Unterschiede zwischen den Gruppen.

Die Delphi-Expertengruppe hat die prozentuale Änderung des MVV (20% vom Ausgangswert) als Kriterium für eine *post hoc* Responder-Analyse empfohlen. Basierend auf diesem Kriterium sprachen numerisch mehr Patienten der BMN 110-Gruppen (28.6%;  $p=0.0576$  einmal wöchentlich; 21.2%;  $p=0.2346$  jede zweite Woche) als der Placebo-Gruppe (12.0%) auf die Behandlung an.

Hinsichtlich der prozentualen Änderung anderer respiratorischer Parameter (FVC, FEV1, FIVC und FET) zeigten sich in beiden BMN 110-Gruppen meist positive Trends, das Ausmaß der Änderung war allerdings gering (Tabelle 4-39, Tabelle 4-41). In Woche 24 zeigten sich in beiden BMN 110-Gruppen negative FIVC-Werte verglichen mit Placebo, was einem negativen Effekt von BMN 110 auf die prozentuale Änderung des FIVC entspricht. Allerdings waren die KIs als Anzeichen hoher Variabilität breit und die geringe absolute Änderung des FIVC deutet an, dass dieses Ergebnis nicht klinisch relevant ist.

Um nennenswerte Änderungen der respiratorischen Parameter erreichen zu können, wäre unter Umständen eine längere Behandlung erforderlich.

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

Nicht zutreffend.

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.

#### **4.3.1.3.6 MPS HAQ (Selbstversorgung, Hilfe durch die Betreuer und Mobilität) – 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 0). 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 4-42: Operationalisierung des Endpunkts „MPS HAQ“

Studie	Operationalisierung
MOR-004	<p>Das “MPS Health Assessment Questionnaire” (MPS HAQ) wurde entwickelt, um bei Patienten mit MPS I zu erfassen, inwieweit der/die Patient(in) in der Lage ist, sich selbst zu versorgen (Essen/Trinken, Ankleiden, Baden, Pflege, Zähne putzen und Toilettengang), und wie mobil er/sie ist (Geschicklichkeit, Mobilität, Gehen, Treppensteigen und Grobmotorik). Das MPS HAQ stellte einen tertiären Endpunkt dar. Der MPS HAQ ist formell noch nicht validiert, es gibt Hinweise darauf, dass einige Fragen in der Domäne Selbstversorgung für die Patienten wenig relevant sind.</p> <p>Der MPS Gesundheitsfragebogen wurde während der Eingangsvisite sowie in Woche 12 und 24 und innerhalb einer Woche im Falle eines vorzeitigen Abbruchs von den Patienten oder deren Betreuern ausgefüllt (bei Patienten &lt;14 Jahre). Um die Variabilität zu minimieren, sollte der Fragebogen jeweils von derselben Person ausgefüllt werden. Zertifizierte Übersetzungen der Fragebögen lagen vor.</p> <p>Der Bogen umfasst 10 Domänen und insgesamt 52 Fragen zur Selbstversorgung und Mobilität. Es wurde abgefragt, ob die Patienten die verschiedenen Tätigkeiten ausüben konnten. Falls die Patienten dazu in der Lage waren sollten die Patienten oder Betreuer auf einer nummerierten, horizontalen Linie jeweils markieren, ob und inwieweit die Patienten beeinträchtigt sind (0 = überhaupt nicht schwierig bis 10 = extrem schwierig). Hinsichtlich der Mobilität wurde zusätzlich abgefragt, ob die Patienten einen Rollstuhl benutzen (nein, Rollstuhl, der geschoben werden muss, elektrischer Rollstuhl oder beides), falls ja, wie oft die Patienten einen Rollstuhl benutzen (weniger als die Hälfte, die Hälfte oder mehr als Hälfte der Zeit während der sich der Patient bewegt oder immer), ob sie Gehhilfen benutzen (ja / nein), falls ja welche Art von Gehhilfe (Stützapparat, Lagerungsothese oder Schienen; Krücken; Gehstock; Rollator) und wie oft die Gehhilfen benutzt werden (weniger als die Hälfte, die Hälfte oder mehr als Hälfte der Zeit während der sich der Patient bewegt oder immer). Dreizehn Fragen mit jeweils vier Antwortmöglichkeiten bezogen sich auf das Ausmaß der Unterstützung der Patienten bei täglichen Aktivitäten durch die Betreuer (Patient ist unabhängig, benötigt minimale, moderate oder komplette Unterstützung). Die Aktivitäten schlossen Essen, Pflege, Baden, Ankleiden des oberen und unteren Körperbereichs, Toilettengang, Aufstehen (von Möbeln oder der Toilette), Fahren mit dem Auto (Ein-/Aussteigen, Sicherheitsgurt anlegen und Öffnen und Schließen der Autotür), ins Bett legen / vom Bett aufstehen oder im Bett Position wechseln, eine Badewanne betreten oder verlassen, Bewegung in Innenräumen (etwa 3-4 Zimmer, Öffnen und Schließen von Türen, Tragen von Gegenständen), Bewegung draußen (etwa 15 Autolängen auf ebener Fläche) und Treppensteigen (12-15 Stufen hinauf und hinunter gehen mit oder ohne Benutzung eines Handlaufs).</p> <p>Für die Domänen Selbstversorgung, Mobilität und Unterstützung durch die Betreuer wurde jeweils ein Gesamtscore gebildet. In die Kalkulation des Gesamtscores zur Selbstversorgung wurden die Ergebnisse von 25 Fragen zu Essen/Trinken, Ankleiden, Baden, Pflege, Zähne putzen und Toilettengang einbezogen:</p>

Score Selbstversorgung = Summe der Scores der ersten 27 Fragen dividiert durch 27 minus fehlende Fragen innerhalb dieser Fragen

Aus 10 Fragen zur Mobilität, zum Gehen und Treppensteigen wurde ein Mobilitätsscore gebildet:

Mobilitätsscore = Summenscore aus den Fragen 28-32 und 35-39 dividiert durch 10 minus fehlende Fragen innerhalb dieser Fragen

Obwohl die Fragen 33 und 34 die Mobilität betrafen (Benutzung eines Rollstuhls und von Gehhilfen) wurden diese bei der Kalkulation des Mobilitätsscores nicht berücksichtigt.

Aus den Antworten auf die Fragen 40-52 wurde ein Gesamtscore gebildet, der den Grad der Unterstützung der Patienten durch die Betreuer widerspiegelt.

Ein höherer Wert zeigt größere Funktionseinschränkungen des Patienten an. Die Veränderung der Domänenscores zu Beginn der Studie bis Woche 24 wurde mittels ANCOVA analysiert. Die Ausgangswerte wurden als Kovariate und Behandlung, Altersgruppen und 6MWT-Kategorie zum Ausgangspunkt als Faktoren in das Modell eingegeben. Die Mittelwertsdifferenzen und deren 95% KIs zwischen den BMN 110-Gruppen und der Placebo-Gruppe wurden berechnet.

*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 4-43: Bewertung des Verzerrungspotenzials des Endpunkts „MPS HAQ“ (Selbstversorgung, Hilfe durch die Betreuer und Mobilität) 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
MOR-004 (9)	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt MPS HAQ (Selbstversorgung, Hilfe durch die Betreuer und Mobilität) wird unter Berücksichtigung der in Tabelle 4-43 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Die Gründe für diese Bewertung sind in Abschnitt 4.3.1.3.1 dargestellt.

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 4-44: Ergebnisse für den Endpunkt „MPS HAQ“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik)

MOR 004	n	Baseline Mittelwert (SD)	24 Wochen Mittelwert (SD)	Absolute Änderung von Baseline bis Woche 24 Mittelwert (SD)
<b>Domäne Selbstversorgung</b>				
Placebo	59	3.5 (2.33)	3.1 (2.29)	-0.4 (1.19)
BMN 110 qow	59	3.9 (2.76)	3.4 (2.57)	-0.5 (1.29)
BMN 110 pro Woche	58 <sup>1</sup>	3.7 (2.69)	3.4 (2.59)	-0.3 (0.90)
<b>Domäne Mobilität</b>				
Placebo	59	4.5 (2.72)	4.0 (2.73)	-0.5 (1.80)
BMN 110 qow	59	5.3 (2.87)	4.4 (2.72)	-1.0 (1.37)
BMN 110 pro Woche	58 <sup>1</sup>	4.6 (2.90)	3.8 (2.62)	-0.7 (1.59)
<b>Domäne Hilfe durch Betreuer</b>				
Placebo	59	26.1 (9.06)	24.9 (8.94)	-1.1 (5.79)
BMN 110 qow	59 <sup>1</sup>	28.6 (10.97)	26.3 (10.17)	-2.4 (5.84)
BMN 110 pro Woche	58 <sup>1</sup>	27.5 (11.06)	24.9 (8.62)	-2.3 (7.02)

<sup>1</sup> n=57 Woche 24

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

Tabelle 4-45: Ergebnisse des MPS HAQ Scores von Baseline bis Woche 24 aus RCT mit dem zu bewertenden Arzneimittel

<b>MOR-004</b>	<b>N</b>	<b>Baseline (Punkte) LSM ± SE (95% KI)</b>	<b>24 Wochen (Punkte) LSM ± SE (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (Punkte) LSM ± SE (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (Punkte) BMN 110 vs. Placebo LSM ± SE (95% KI)</b>	<b>p-Wert (t-test<sup>2</sup>) BMN 110 vs. Placebo</b>
<b>MPS HAQ Gesamtscore Selbstversorgung<sup>1</sup></b>						
Placebo	59	3.5±0.32 (2.9, 4.2)	3.3±0.14 (3.0, 3.6)	-0.4±0.14 (-0.7, -0.1)	-	-
BMN 110 qow	59	3.9±0.32 (3.2, 4.5)	3.2±0.14 (2.9, 3.5)	-0.5±0.14 (-0.7, -0.2)	-0.1±0.20 (-0.5, 0.3)	0.7242
BMN 110 pro Woche	58	3.7±0.33 (3.0, 4.3)	3.4±0.14 (3.1, 3.6)	-0.3±0.14 (-0.6, -0.0)	0.1±0.20 (-0.3, 0.5)	0.7367
<b>MPS HAQ Gesamtscore Mobilität<sup>1</sup></b>						
Placebo	59	4.5±0.34 (3.8, 5.2)	4.3±0.19 (3.9, 4.6)	-0.5±0.19 (-0.9, -0.2)	-	-
BMN 110 qow	59	5.3±0.34 (4.6, 6.0)	3.9±0.19 (3.6, 4.3)	-0.9±0.19 (-1.2, -0.5)	-0.3±0.27 (-0.8, 0.2)	0.2611
BMN 110 pro Woche	58	4.7±0.35 (4.0, 5.3)	4.0±0.19 (3.6, 4.4)	-0.8±0.19 (-1.2, -0.4)	-0.3±0.27 (-0.8, 0.3)	0.3355
<b>MPS HAQ Gesamtscore Unterstützung durch die Betreuer<sup>1</sup></b>						
Placebo	59	26.3±1.28 (23.8, 28.9)	25.8±0.71 (24.4, 27.2)	-1.5±0.71 (-2.9, -0.1)	-	-
BMN 110 qow	59	28.6±1.28 (26.1, 31.1)	25.4±0.72 (24.0, 26.8)	-1.9±0.72 (-3.3, -0.5)	-0.4±1.01 (-2.4, 1.6)	0.6794
BMN 110 pro Woche	58	27.3±1.29 (24.7, 29.8)	24.9±0.72 (23.5, 26.4)	-2.3±0.72 (-3.7, -0.9)	-0.9±1.01 (-2.8, 1.1)	0.3990

<sup>1</sup> ANCOVA mit den Modelltermen Behandlung, Altersgruppen, Baseline-6MWT-Kategorie, und kontinuierlicher Ausgangswert des MPS HAQ-Scores.

<sup>2</sup> p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell. Nach Hochberg-Adjustierung für Multiplizität wurde das Ergebnis als signifikant eingestuft, wenn beide p-Werte unter 0.05 oder ein p-Wert unter 0.025 lagen.

Die Änderung des MPS HAQ Subscores zur Selbstversorgung bis Woche 24 war in der Placebo- und den BMN 110-Gruppen vergleichbar (Tabelle 4-44). Die mittlere Änderung ( $\pm$  SD) des MPS HAQ-Subscores zur Selbstversorgung bis Woche 24 betrug  $-0.4 \pm 1.19$  in der Placebogruppe,  $-0.5 \pm 1.29$  in der BMN 110 qow Gruppe und  $-0.3 \pm 0.90$  bei den Patienten, die einmal pro Woche BMN 110 erhielten. Der mittlere Domänenscore ( $\pm$  SD) zur Hilfestellung durch die Betreuer nahm in den BMN 110-Gruppen ( $-2.4 \pm 5.84$  BMN 110 qow,  $-2.3 \pm 7.02$  BMN 110 jede Woche) bis Woche 24 geringfügig stärker ab als in der Placebo-Gruppe ( $-1.1 \pm 5.79$ ). Der mittlere Mobilitätscore ( $\pm$  SD) nahm in der Placebogruppe bis Woche 24 um  $-0.5 \pm 1.80$ , in der BMN 110 qow Gruppe um  $-1.0 \pm 1.37$  und bei den Patienten, die BMN 110 jede Woche erhielten um  $-0.7 \pm 1.59$  ab. Verglichen mit Placebo unterschieden sich die Behandlungseffekte nicht signifikant zwischen den Gruppen (siehe Tabelle 4-45).

Die Fragen zur Benutzung eines Rollstuhls oder von Gehhilfen wurden bei der Berechnung des Mobilitätsscores nicht berücksichtigt und separat ausgewertet (*post hoc*). Insgesamt 35 (59.3%), 23 (39.0%) bzw. 27 (46.6%) der Patienten oder deren Betreuer in der Placebo-Gruppe und den BMN 110-Gruppen (qow bzw. jede Woche) gaben zu Beginn der Studie an, keinen Rollstuhl zu benutzen. Von diesen benötigten sechs, drei bzw. zwei Patienten in Woche 24 einen Rollstuhl. Unter den Patienten, die zu Beginn der Studie einen Rollstuhl benötigen, waren zu Studienende ein, drei bzw. zwei Patienten nicht mehr auf einen Rollstuhl angewiesen. Die Zahl der Patienten, die in Woche 24 einen Rollstuhl benutzten, nahm in der Placebo-Gruppe um fünf Patienten zu (8.8%), während die Anzahl in den BMN 110-Gruppen von Studienbeginn bis Woche 24 unverändert blieb (0%). Die Zahl der Patienten, die im Verlauf der Studie bis Woche 24 eine Gehhilfe benötigten, nahm in der Placebo-Gruppe und bei den einmal wöchentlich mit BMN 110 behandelten Patienten um drei bzw. einen Patienten zu, während in der BMN 110 qow Gruppe zu Studienende zwei Patienten weniger eine Gehilfe benötigten (Tabelle 4-46).

Tabelle 4-46: Ergebnisse des MPS HAQ (Benutzung eines Rollstuhls oder von Gehilfen zu Beginn der Studie und in Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT)

<b>Benutzung Rollstuhl oder Gehilfen</b>	<b>Placebo (n=59) n (%)</b>	<b>BMN 110 qow (n=59) n (%)</b>	<b>BMN 110 pro Woche (n=58) n (%)</b>
<b>Rollstuhl</b>			
Baseline: Nein <sup>1</sup>	35 (59.3)	23 (39.0)	27 (46.6)
Woche 24: Nein	29 (82.9)	20 (87.0)	25 (92.6)
Woche 24: Ja	6 (17.1)	3 (13.0)	2 (7.4)
Baseline: Ja <sup>1</sup>	22 (37.3)	34 (57.6)	30 (51.7)
Woche 24: Nein	1 (4.5)	3 (8.8)	2 (6.7)
Woche 24: Ja	21 (95.5)	31 (91.2)	28 (93.3)
<b>Gehilfe</b>			
Baseline: Nein <sup>1</sup>	38 (64.4)	34 (57.6)	39 (67.2)
Woche 24: Nein	33 (86.8)	32 (94.1)	36 (92.3)
Woche 24: Ja	5 (13.2)	2 (5.9)	3 (7.7)
Baseline: Ja <sup>1</sup>	18 (30.5)	22 (37.3)	17 (29.3)
Woche 24: Nein	2 (11.1)	4 (18.2)	2 (11.8)
Woche 24: Ja	16 (88.9)	18 (81.8)	15 (88.2)

<sup>1</sup> Patienten mit kompletten Daten.

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

Nicht zutreffend

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.

#### 4.3.1.3.7 Endpunkt Anthropometrie – 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 0). 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 4-47: Operationalisierung des Endpunkts „Anthropometrie“

Studie	Operationalisierung
MOR-004	<p>Die anthropometrischen Endpunkte waren im Studienprotokoll als zusätzliche Parameter und im statistischen Analyseplan als tertiäre Wirksamkeitsendpunkt definiert. Die anthropometrische Messungen wurden zu Beginn der Studie und nach 24 Wochen erhoben (innerhalb einer Woche bei einem möglichen Abbruch) und schlossen die Stehhöhe, Körperlänge, Sitzhöhe und das Körpergewicht ein.</p> <p>Die normalisierte Körpergröße im Stehen (z-Score) wurde unter Verwendung der normativen Daten des CDC (2000) analysiert und die Analyse wurde auf männliche Patienten im Alter von <math>\leq 18</math> Jahren und weibliche Patienten im Alter von <math>\leq 15</math> Jahren beschränkt. Der z-Score ist ein Maß der Divergenz eines individuellen experimentellen Ergebnisses von dem wahrscheinlichsten experimentellen Ergebnis, des Mittelwerts der Datenpunkte. Ein z-Score von 0 bedeutet, dass der z-Score dem Mittelwert entspricht, bei +1 bzw. -1 liegt der Wert eine SD über bzw. unter dem Mittelwert. Der z-Score wurde für Körpergröße im Stehen gewählt, da er die durch das Alter bedingte Variabilität eliminiert und dadurch möglicherweise sensitiver ist. Darüber hinaus sind z-Scores leichter als Maß der Krankheitsprogression interpretierbar.</p> <p>Die Wachstumsrate während der Studie wurde mit der Wachstumsrate vor der Studie verglichen (bei Patienten, bei denen diese Daten innerhalb von zwei Jahren vor Studieneintritt vorlagen). Die Wachstumsrate vor Studienbeginn wurde wie folgt bestimmt:</p> <p><math>[(z\text{-Score der Baseline-Stehhöhe}) - (z\text{-Score der Stehhöhe vor der Studie, aktuellster Wert nicht } &gt; 2 \text{ Jahre vor Studieneintritt})] / \text{Zeit (in Jahren) zwischen den Messungen}]</math></p> <p>Die Wachstumsrate während der Studie wurde durch den Vergleich der z-Scores zwischen der Eingangsvisite und Woche 24 dividiert durch die Zeit (in Jahren) ermittelt. Die Änderung der Wachstumsrate wurde als Rate während der Studie minus der Rate vor der Studie kalkuliert, wodurch die Wachstumsrate auch negative Werte annehmen konnte. Diese Analyse wurde auf männliche Patienten im Alter von <math>\leq 18</math> Jahren und weibliche Patienten im Alter von <math>\leq 15</math> Jahren beschränkt.</p> <p>Bei männliche Patienten im Alter von <math>\leq 18</math> Jahren und weiblichen Patienten im Alter von <math>\leq 15</math> Jahren wurde ein der Primäranalyse ähnliches ANCOVA-Modell mit einer zusätzlichen Baseline-Kovariate eingesetzt, um die z-Scores (Körpergröße im Stehen) und die Wachstumsrate zu analysieren. Andere anthropometrische Parameter wurden mittels deskriptiver Statistik ausgewertet.</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 4-48: Bewertung des Verzerrungspotenzials für den Endpunkt „Anthropometrie“ 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
MOR-004 (9)	niedrig	ja	nein	ja	ja	hoch

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential auf Studienebene wurde als niedrig eingestuft. Die primären Analysen der z-Scores und der Wachstumsrate wurden auf jüngere Patienten beschränkt, so dass es sich um Subgruppenanalysen handelt. In die Analyse der Wachstumsrate wurden jüngere Patienten eingeschlossen, bei denen Daten zur Wachstumsrate vor Studienbeginn vorlagen, so dass die Randomisierung nicht eingehalten wurde. Daher wird das Verzerrungspotential für diesen Endpunkt als hoch eingestuft.

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

Tabelle 4-49: Ergebnisse für den Endpunkt „Anthropometrie“ (Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel (ITT, deskriptive Statistik)

MOR 004	n BL, 24 W, Änderung bis 24 W	Baseline Mittelwert (SD)	24 Wochen Mittelwert (SD)	Absolute Änderung von Baseline bis Woche 24 Mittelwert (SD)
<b>Körperlänge (cm)</b>				
Placebo	56/54/52	107.7 (16.31)	108.0 (15.94)	1.0 (1.89)
BMN 110 qow	57/57/57	107.4 (14.50)	108.9 (14.57)	1.5 (2.27)
BMN 110 pro Woche	55/54/54	104.8 (12.66)	105.2 (11.79)	1.0 (2.27)
<b>Körperlänge (cm) bei jüngeren Patienten (männlich ≤18 Jahre, weiblich ≤15 Jahre)</b>				
Placebo	38/36/34	104.3 (11.74)	104.2 (10.34)	1.3 (2.02)
BMN 110 qow	44/44/44	104.6 (11.36)	105.9 (11.37)	1.3 (1.63)
BMN 110 pro Woche	41/40/40	103.9 (11.90)	104.6 (10.68)	1.4 (2.43)

<b>MOR 004</b>	<b>n</b>	<b>Baseline</b>	<b>24 Wochen</b>	<b>Absolute Änderung von</b>
	<b>BL, 24 W,</b>	<b>Mittelwert (SD)</b>	<b>Mittelwert (SD)</b>	<b>Baseline bis Woche 24</b>
	<b>Änderung</b>			<b>Mittelwert (SD)</b>
	<b>bis 24 W</b>			
<b>Stehhöhe (cm)</b>				
Placebo	58/58/57	105.5 (16.79)	106.3 (16.83)	0.9 (1.33)
BMN 110 qow	55/55/55	104.6 (15.84)	106.1 (15.61)	1.4 (2.40)
BMN 110 pro Woche	56/57/56	101.3 (13.06)	102.7 (12.96)	1.2 (1.51)
<b>Stehhöhe (cm) bei jüngeren Patienten (männlich ≤18 Jahre, weiblich ≤15 Jahre)</b>				
Placebo	39/40/39	101.7(12.28)	102.5 (12.38)	1.1 (1.36)
BMN 110 qow	41/41/41	100.7 (12.22)	102.4 (12.04)	1.7 (2.67)
BMN 110 pro Woche	43/44/43	100.1 (12.10)	101.8 (11.94)	1.6 (1.35)
<b>Körpergröße (cm) sitzend</b>				
Placebo	58/59/58	61.7 (17.32)	65.0 (18.89)	2.9 (10.38)
BMN 110 qow	59/58/58	61.4 (16.31)	62.8 (17.20)	1.4 (6.06)
BMN 110 pro Woche	58/57/57	59.0 (14.87)	61.3 (14.98)	3.0 (9.83)
<b>Körpergröße (cm) sitzend bei jüngeren Patienten (männlich ≤18 Jahre, weiblich ≤15 Jahre)</b>				
Placebo	39/40/39	57.0 (11.60)	61.8 (16.29)	4.1 (12.51)
BMN 110 qow	45/45/45	58.6 (13.99)	60.4 (15.40)	1.8 (6.80)
BMN 110 pro Woche	44/43/43	56.6 (13.32)	58.7 (12.87)	3.1 (9.13)
<b>Körpergewicht (kg)</b>				
Placebo	58/59/58	25.4 (11.47)	26.3 (11.89)	1.1 (1.50)
BMN 110 qow	59/59/59	26.5 (11.84)	27.5 (12.30)	1.0 (1.30)
BMN 110 pro Woche	58/58/58	22.9 (10.46)	23.5 (10.10)	0.7 (1.73)
<b>Körpergewicht (kg) bei jüngeren Patienten (männlich ≤18 Jahre, weiblich ≤15 Jahre)</b>				
Placebo	39/40/39	20.6 (6.44)	21.7 (7.45)	1.2 (1.38)
BMN 110 qow	45/45/45	23.1 (8.61)	24.1 (9.33)	1.1 (1.32)
BMN 110 pro Woche	44/44/44	20.3 (8.38)	21.0 (7.76)	0.7 (1.96)
<b>Stehhöhe normalisiert (z-Score)</b>				
Placebo	58/58/57	-6.0 (2.77)	-6.1 (2.64)	-0.1 (0.29)
BMN 110 qow	55/54/54	-6.3 (2.43)	-6.2 (2.33)	-0.0 (0.49)
BMN 110 pro Woche	56/57/56	-6.4 (2.55)	-6.4 (2.53)	-0.0 (0.27)

<b>MOR 004</b>	<b>n</b> <b>BL, 24 W,</b> <b>Änderung</b> <b>bis 24 W</b>	<b>Baseline</b> <b>Mittelwert (SD)</b>	<b>24 Wochen</b> <b>Mittelwert (SD)</b>	<b>Absolute Änderung von</b> <b>Baseline bis Woche 24</b> <b>Mittelwert (SD)</b>
<b>Stehhöhe normalisiert (z-Score) bei jüngeren Patienten (männlich ≤18 Jahre, weiblich ≤15 Jahre)</b>				
Placebo	39/40/39	-5.1 (2.17)	-5.4 (2.12)	-0.2 (0.30)
BMN 110 qow	41/40/40	-5.8 (2.22)	-5.8 (2.06)	-0.1 (0.56)
BMN 110 pro Woche	43/44/43	-5.6 (2.07)	-5.7 (2.04)	-0.0 (0.29)

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

Tabelle 4-50: Ergebnisse für den Endpunkt „Anthropometrie“ von Baseline bis Woche 24 aus RCT mit dem zu bewertenden Arzneimittel (ITT, jugendliche Männer ≤18 Jahre, jugendliche Frauen ≤15 Jahre)

<b>MOR-004</b>	<b>N</b>	<b>Baseline (SD) LSM ± SE (95% KI)</b>	<b>24 Wochen (SD) LSM ± SE (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (SD) LSM ± SE (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (SD) BMN 110 vs. Placebo LSM ± SE (95% KI)</b>	<b>p-Wert (t-test<sup>2</sup>) BMN 110 vs. Placebo</b>
<b>Veränderung der normalisierten Stehhöhe (z-Score)</b>						
Placebo	40	-5.2 (-5.8, -4.5)	-5.7 (-5.8, -5.5)	-0.2 (-0.3, -0.0)	-	-
BMN 110 qow	45	-5.8 (-6.4, -5.1)	-5.6 (-5.7, -5.4)	-0.1 (-0.2, 0.1)	0.1 (-0.1, 0.3)	0.2218
BMN 110 pro Woche	44	-5.6 (-6.2, -5.0)	-5.5 (-5.6, -5.4)	-0.0 (-0.2, 0.1)	0.1 (-0.0, 0.3)	0.1149
<b>Veränderung der normalisierten Wachstumsrate (z-Score)</b>						
Placebo	40	-0.7 (-0.9, -0.4)	-0.4 (-0.7, -0.0)	0.2 (-0.2, 0.5)	-	-
BMN 110 qow	45	-0.3 (-0.6, -0.1)	-0.0 (-0.4, 0.3)	0.5 (0.2, 0.9)	0.4 (-0.1, 0.9)	0.1384
BMN 110 pro Woche	44	-0.6 (-0.8, -0.3)	0.0 (-0.3, 0.3)	0.5 (0.2, 0.9)	0.4 (-0.1, 0.9)	0.1032

ANCOVA (Änderung von Baseline) mit den Kovariaten Behandlung, Altersgruppen, 6MWT-Kategorie und z-Score bzw. Wachstumsrate als kontinuierliche Variable. p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell.

Die mittlere Änderung der absoluten anthropometrischen Messungen bis Woche 24 unterschied sich zwischen den Gruppen nicht (Tabelle 4-49).

Bei den jüngeren Patienten (männlich  $\leq 18$  Jahre, weiblich  $\leq 15$  Jahre) betragen die z-Scores zur Stehhöhe (Mittelwert  $\pm$  SD) zu Beginn der Studie  $-5.2 \pm 0.33$  (n=40) in der Placebo-Gruppe,  $-5.8 \pm 0.32$  (n=45) in der BMN 110 2.0 mg/kg/qow Gruppe und  $-5.6 \pm 0.31$  (n=44) in der BMN 110 2.0 mg/kg/Woche Gruppe (Tabelle 4-49). Somit lagen die z-Scores in allen Gruppen mindestens fünf SDs unterhalb der normalen Alters-adjustierten Stehhöhe. Die Änderung der normalisierten z-Scores der Stehhöhe (LSM  $\pm$  SE) betrug  $-0.2 \pm 0.06$  in der Placebogruppe,  $-0.1 \pm 0.06$  in der BMN 110 2.0 mg/kg/qow Gruppe und  $-0.0 \pm 0.06$  in der BMN 110 2.0 mg/kg/Woche Gruppe. Der geschätzte Behandlungseffekt hinsichtlich der z-Scores (Änderung bis Woche 24) verglichen mit Placebo war 0.1 (95% KI -0.0, 0.3; p=0.1149) bzw. (95% KI -0.1, 0.3; p=0.2218) bei den Patienten, die BMN 110 einmal wöchentlich bzw. jede zweite Woche erhalten hatten (Tabelle 4-50).

Die z-Scores der normalisierten Wachstumsrate waren zu Beginn der Studie in allen Gruppen innerhalb einer SD unter der normalen Alters-adjustierten Wachstumsrate (Tabelle 4-49). Die Änderung der z-Scores der normalisierten Wachstumsrate (LSM  $\pm$  SE) bis Woche 24 betrug  $0.2 \pm 0.18$ ,  $0.5 \pm 0.18$  bzw.  $0.5 \pm 0.16$  bei den Patienten, die Placebo und BMN 110 erhielten (jede zweite Woche bzw. einmal wöchentlich). Der geschätzte Behandlungseffekt war bei den mit BMN 110 behandelten Patienten etwas größer als bei den Patienten der Placebo-Gruppe (0.4, 95% KI -0.1, 0.9; p=0.1032 BMN 110 einmal pro Woche, 0.4, 95% KI -0.1, 0.9; p=0.1384; Tabelle 4-50).

Beide BMN 110-Dosisregimes hatten bei den jüngeren Patienten einen positiven Effekt auf die Körpergröße und die Wachstumsrate gemessen mittels z-Scores. Diese Ergebnisse deuten an, dass BMN 110 möglicherweise bei noch im Wachstum befindlichen MPS IVA-Patienten die Progression der Erkrankung hinsichtlich der verminderten Wachstumsrate und Körpergröße aufhalten kann.

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

Nicht zutreffend.

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.

#### 4.3.1.3.8 Inzidenz nicht schwerwiegender und schwerwiegender UEs – 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 0). 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 4-51: Operationalisierung des Endpunkts „Inzidenz nicht schwerwiegender und schwerwiegender UEs“

Studie	Operationalisierung
MOR-004	<p>In Übereinstimmung mit der ICH wurde jedes ungünstige oder nicht beabsichtigte Zeichen (einschließlich abnormale Laborbefunde), jedes Symptom und jede Erkrankung in zeitlichem Zusammenhang mit der Anwendung der Prüfpräparate als UE dokumentiert, unabhängig davon ob das Ereignis mit den Prüfpräparaten in Zusammenhang stand oder nicht. Als UE war auch jede zwischenzeitlich aufgetretene Erkrankung oder Verletzung, die eine Exazerbation vorbestehender Erkrankungen darstellte, definiert. Es wurden nur <i>behandlungsspezifische</i> („<i>treatment-emergent</i>“) UEs erfasst. Ein behandlungsspezifisches UE war definiert als ein UE, welches nach Behandlungsbeginn neu aufgetreten ist oder dessen Schwere sich nach Behandlungsbeginn verschlimmerte. Falls das Datum des Auftretens eines UEs nicht bekannt wurde, wurde das UE als behandlungsspezifisch eingestuft. Die UEs sollten vorzugsweise als Diagnose und nicht als einzelne mit dieser Diagnose assoziierte Symptome kodiert werden (z.B. Überempfindlichkeit statt Urtikaria, Erbrechen und Schüttelfrost). Die Prüfarzte bewerteten die klinische Relevanz von Laborveränderungen, und jede klinisch signifikante Laborveränderung sollte als UE dokumentiert werden. Klinisch signifikante Veränderungen während der körperlichen Untersuchung im Vergleich zur Ausgangsuntersuchung sollten ebenfalls als UE gemeldet werden. Zu jeder Visite wurden die Patienten befragt, ob seit der letzten Visite UEs neu aufgetreten sind oder fortbestehen.</p> <p>Schwerwiegende UEs wurden vom Zeitpunkt der schriftlichen Einverständniserklärung bis 30 Tage nach der letzten Visite und nicht schwerwiegende UEs vom Start der Infusion in Woche 0 bis zur letzten Visite erfasst. Im Falle ein nicht schwerwiegendes Ereignis war zu Studienende nicht abgeklungen (Ausgangstatus wieder hergestellt oder Stabilisierung mit der Erwartung eines chronischen Verlaufs), entschieden der Prüfarzt und der klinische Monitor gemeinsam, ob eine Nachbeobachtung erforderlich ist. Bei neu aufgetretenen oder bestehenden SUEs konnte es erforderlich sein, diese auch über den Berichtszeitraum der Studie hinaus zu beobachten.</p> <p>Die Prüfarzte bewerteten die Schwere der UEs, einen möglichen Zusammenhang mit den Prüfpräparaten (kein Zusammenhang, möglicher oder wahrscheinlicher Zusammenhang) und ob das Ereignis den Kriterien eines SUEs entsprach. Der Schweregrad der UEs wurde als mild (keine Beeinträchtigung üblicher Aktivitäten), mittelschwer (Patient bei üblichen Aktivitäten etwas beeinträchtigt) oder schwer (Patient ist zu üblichen Aktivitäten nicht fähig) klassifiziert. Ein möglicher Zusammenhang zwischen einem UE und dem Prüfpräparat wurde angenommen, wenn das Ereignis in einem zeitlichen Zusammenhang mit der Verabreichung stand und das Ereignis gleichermaßen gut durch andere Faktoren oder Gründe erklärt werden konnte. Ein wahrscheinlicher Zusammenhang lag vor, wenn das Ereignis in einem zeitlichen Zusammenhang mit der Verabreichung stand und das Ereignis eher durch die Exposition mit dem Prüfpräparat als andere Faktoren oder Gründe erklärt werden konnte.</p> <p>Ein SUE war definiert als ein Ereignis, welches:</p> <ul style="list-style-type: none"> <li>– zum Tod führt,</li> <li>– lebensbedrohlich ist, d.h. das Ereignis setzt den Patienten unmittelbar dem Risiko aus, an dem Ereignis zu sterben. Diese Definition schloss keine Ereignisse ein, die zum Tod geführt haben könnten, wenn sie in einer schwereren Form aufgetreten wären.</li> <li>– eine Hospitalisierung erfordert oder eine bereits bestehende Hospitalisierung verlängert. Eine Hospitalisierung für weniger als 24 h oder mit dem Ziel, einen i.v. Katheter zur Verabreichung der Prüfpräparate anzulegen, wurde nicht als SUE bewertet.</li> <li>– zu dauerhafter oder signifikanter Behinderung oder Invalidität führt, d.h. ein Ereignis, das zu einer bedeutsamen Störung der Fähigkeit des Patienten, übliche Lebensfunktionen auszuüben, führt.</li> </ul>

- eine kongenitale Anomalie oder einen Geburtsfehler nach sich zieht,
- anderweitig als medizinisch bedeutsam betrachtet wird.

Die UEs wurden durch den Sponsor auf Basis der in den CRFs eingegebenen UEs entsprechend MedDRA (Version 15.0) nach Organsystem und PT kodiert. In die Sicherheitsanalyse wurden alle Patienten, die BMN 110 oder Placebo während der Studie erhalten hatten, eingeschlossen. Hierbei wurde das tatsächlich erhaltene Dosisregime (höchste Dosisfrequenz) analysiert. Fehlende Informationen zu UEs wurden konservativ gehandhabt. Fehlende Daten zu UEs wurden eingegeben, wenn die vorliegenden Teilmeldungen andeuteten, dass das UE wahrscheinlich behandlungsspezifisch war.

Ein in einer Kausalitätsbewertung möglicher oder ein wahrscheinlicher Zusammenhang eines UEs mit der Behandlung wurde als *Therapie-assoziiertes* Ereignis („*drug-related*“) eingestuft.

Die Therapie-assoziierten Ereignisse sind für die Bewertung der Zielgrößenkategorie „*Nebenwirkungen*“ am besten geeignet.

*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 4-52: Bewertung des Verzerrungspotenzials für den Endpunkt „Inzidenz nicht schwerwiegender und schwerwiegender UEs“ in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
MOR-004 (9)	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt „Inzidenz nicht schwerwiegender und schwerwiegender UEs“ wird unter Berücksichtigung der in Tabelle 4-52 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Die Gründe für diese Bewertung sind in Abschnitt 4.3.1.3.1 dargestellt. Wie dort angesprochen, könnte die höhere Anzahl von Überempfindlichkeitsreaktionen oder Infusionen, die aufgrund von IAR unterbrochen wurden, zu einer Entblindung einzelner Patienten geführt haben. Allerdings wird unter Berücksichtigung der hohen Qualitätsstandards der Studie nicht vermutet, dass dies - falls zutreffend - einen relevanten Einfluss auf die Berichterstattung von UEs hatte oder die Ergebnisse relevant verzerrt haben könnte.

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 4-53: Ergebnisse für den Endpunkt „Inzidenz der gesamten nicht schwerwiegenden und der gesamten schwerwiegenden UEs“ (MedDRA PT) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation, MOR-004) (9)

UE	Placebo (N=59) n (%)	BMN 110 2.0 mg/kg qow (N=59) n (%)	BMN 110 2.0 mg/kg/Woche (N=58) n (%)
Gesamte UEs	57 (96.6%)	59 (100.0%)	56 (96.6%)
Mild	36 (61.0%)	33 (55.9%)	28 (48.3%)
Mittelschwer	20 (33.9%)	23 (39.0%)	26 (44.8%)
Schwer	1 (1.7%)	3 (5.1%)	2 (3.4%)
Therapie-assoziierte gesamte UEs *	36 (61.0%)	42 (71.2%)	42 (72.4%)
Mild	32 (54.2%)	27 (45.8%)	24 (41.4%)
Mittelschwer	4 (6.8%)	14 (23.7%)	16 (27.6%)
Schwer	0 (0.0%)	1 (1.7%)	2 (3.4%)
SUEs	2 (3.4%)	4 (6.8%)	9 (15.5%)
Therapie-assoziierte SUEs *	0 (0.0%)	1 (1.7%)	2 (3.4%)

\* („drug related“)

Insgesamt 57 (96.6%) Patienten der Placebo-Gruppe, 59 (100.0%) Patienten der BMN 110 qow Gruppe und 56 (96.6%) der Patienten, die BMN 110 einmal pro Woche erhielten, entwickelten mindestens ein behandlungsspezifisches UE. Die überwiegende Mehrzahl der UEs war milder oder mittelschwerer Ausprägung. Die Anzahl der Patienten, die ein mit der Behandlung zumindest möglicherweise in Zusammenhang stehendes UE entwickelten, war in den BMN 110-Gruppen (71.2% BMN 110 qow, 72.4% BMN 110 jede Woche) etwas höher als in der Placebo-Gruppe (61.0%). Die meisten Therapie-assoziierten UEs waren in allen Gruppen mild oder mittelschwer und klangen unter symptomatischer Therapie und/oder Unterbrechung oder Abbruch der Infusion ab. Insgesamt zwei Patienten (3.4%) der Placebo-Gruppe, vier Patienten (6.8%) der BMN 110 qow Gruppen und 9 Patienten (15.5%), denen BMN 110 einmal wöchentlich injiziert wurde, entwickelten ein SUE. Die häufigsten SUEs waren Pneumonie (n=2, 3.4% BMN 110 2.0 mg/kg/Woche) und Otitis media (n=1 in beiden BMN 110-Gruppen) und standen nicht mit der Behandlung in Zusammenhang. Alle anderen SUEs traten bei nur jeweils einem Patienten auf und entsprachen Infusionsreaktionen oder Begleiterkrankungen der MPS IVA. Demgegenüber wurde bei nur einem Patienten der BMN 110 qow Gruppe (anaphylaktische Reaktion) und zwei Patienten, denen BMN 110 einmal wöchentlich injiziert wurde (Überempfindlichkeitsreaktion und Erbrechen), ein schwerwiegendes mit der Behandlung in Zusammenhang stehendes UE dokumentiert. Diese SUEs traten während der Infusion auf und konnten durch Reduktion der Infusionsrate,

Unterbrechung oder Abbruch der Infusion kontrolliert werden. Der Unterschied in der Inzidenz der mit der Behandlung in Zusammenhang stehenden UEs zwischen der Placebo-Gruppe und den Patienten, die einmal wöchentlich mit BMN 110 behandelt wurden, war nicht signifikant (Tabelle 4-54). Keiner der Patienten verstarb während der Studie oder entwickelte ein UE, das zum Abbruch der Behandlung führte.

Tabelle 4-54: Endpunkt „Inzidenz der gesamten nicht schwerwiegenden und der gesamten schwerwiegenden UEs“: Placebo vs. BMN 110 2.9 mg/kg/Woche (9)

UE	Placebo n/total	BMN 110 jede Woche n/total
Gesamte UEs	57/59	56/58
Gesamte Therapie-assozierte UEs	36/59	42/58
Gesamte SUEs	2/59	9/58
Gesamte Therapie-assozierte SUEs	0/59	2/58

Tabelle 4-55: Ergebnisse für den Endpunkt „Inzidenz gesamter behandlungsspezifischer UEs“ (MedDRA PT) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation) (9)

UE	Placebo (N=59) n (%)	BMN 110 2.0 mg/kg qow (N=59) n (%)	BMN 110 2.0 mg/kg/Woche (N=58) n (%)
<b>Gesamte UEs stratifiziert nach SOC(≥10% in einer der Gruppen)<sup>1</sup></b>			
Infektionen und parasitäre Erkrankungen	39 (66.1%)	42 (71.2%)	39 (67.2%)
Nasopharyngitis	9 (15.3%)	12 (20.3%)	10 (17.2%)
Infektionen der oberen Atemwege	9 (15.3%)	10 (16.9%)	10 (17.2%)
Gastroenteritis	4 (6.8%)	8 (13.6%)	7 (12.1%)
Virale Infektion	1 (1.7%)	6 (10.2%)	3 (5.2%)
Otitis media	4 (6.8%)	5 (8.5%)	9 (15.5%)
Rhinitis	6 (10.2%)	4 (6.8%)	5 (8.6%)
Pharyngitis	7 (11.9%)	3 (5.1%)	4 (6.9%)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	37 (62.7%)	40 (67.8%)	38 (65.5%)
Fieber	17 (28.8%)	22 (37.3%)	25 (43.1%)
Ermüdung (Fatigue)	15 (25.4%)	8 (13.6%)	9 (15.5%)

<b>UE</b>	<b>Placebo (N=59) n (%)</b>	<b>BMN 110 2.0 mg/kg qow (N=59) n (%)</b>	<b>BMN 110 2.0 mg/kg/Woche (N=58) n (%)</b>
Schüttelfrost	1 (1.7%)	6 (10.2%)	6 (10.3%)
Erkrankungen des Gastrointestinaltrakts	41 (69.5%)	40 (67.8%)	37 (63.8%)
Erbrechen	21 (35.6%)	21 (35.6%)	26 (44.8%)
Übelkeit	12 (20.3%)	14 (23.7%)	18 (31.0%)
Diarrhöe	7 (11.9%)	12 (20.3%)	12 (20.7%)
Abdominale Schmerzen	5 (8.5%)	8 (13.6%)	14 (24.1%)
Oberbauchschmerz	5 (8.5%)	4 (6.8%)	9 (15.5%)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	33 (55.9%)	36 (61.0%)	35 (60.3%)
Husten	21 (35.6%)	17 (28.8%)	16 (27.6%)
Oropharyngealschmerz	7 (11.9%)	9 (15.3%)	12 (20.7%)
Dyspnoe	3 (5.1%)	6 (10.2%)	7 (12.1%)
Erkrankungen des Nervensystems	26 (44.1%)	31 (52.5%)	31 (53.4%)
Kopfschmerzen	21 (35.6%)	24 (40.7%)	24 (41.4%)
Schwindel („dizziness“)	3 (5.1%)	4 (6.8%)	7 (12.1%)
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	30 (50.8%)	30 (50.8%)	25 (43.1%)
Schmerzen in den Extremitäten	9 (15.3%)	14 (23.7%)	9 (15.5%)
Rückenschmerzen	6 (10.2%)	10 (16.9%)	7 (12.1%)
Arthralgie	17 (28.8%)	9 (15.3%)	10 (17.2%)
Herzerkrankungen	19 (32.2%)	13 (22.0%)	16 (27.6%)
Trikuspidalklappeninsuffizienz	3 (5.1%)	7 (11.9%)	4 (6.9%)
Tachykardie	6 (10.2%)	2 (3.4%)	3 (5.2%)
Erkrankungen der Haut und des Unterhautzellgewebes	14 (23.7%)	23 (39.0%)	16 (27.6%)
Ausschlag	5 (8.5%)	6 (10.2%)	6 (10.3%)
Untersuchungen	19 (32.2%)	20 (33.9%)	15 (25.9%)
Verminderte Sauerstoffsättigung	6 (10.2%)	7 (11.9%)	6 (10.3%)
Vaskuläre Erkrankungen	13 (22.0%)	17 (28.8%)	13 (22.4%)
Erkrankungen des Ohres und des Labyrinths	10 (16.9%)	11 (18.6%)	9 (15.5%)
Ohrenschmerzen	5 (8.5%)	8 (13.6%)	3 (5.2%)
Augenerkrankungen	5 (8.5%)	6 (10.2%)	9 (15.5%)

UE	Placebo (N=59) n (%)	BMN 110 2.0 mg/kg qow (N=59) n (%)	BMN 110 2.0 mg/kg/Woche (N=58) n (%)
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen	8 (13.6%)	15 (25.4%)	8 (13.8%)
Psychiatrische Erkrankungen	3 (5.1%)	9 (15.3%)	5 (8.6%)
Chirurgische und medizinische Eingriffe	6 (10.2%)	2 (3.4%)	5 (8.6%)
Erkrankungen des Immunsystems	1 (1.7%)	6 (10.2%)	3 (5.2%)

MedDRA Version 15.0

<sup>1</sup> Patienten mit mehr als einem UE innerhalb eines MedDRA PTs wurden nur einmal gezählt.

Die häufigsten behandlungsspezifischen UEs traten in Zusammenhang mit den Infusionen auf (Erbrechen, Fieber und Kopfschmerzen). Die häufigsten behandlungsspezifischen UEs waren Erbrechen (35.6% Placebo, 35.6% BMN 110 2.0 mg/kg/qow, 44.8% BMN 110 2.0 mg/kg/Woche), Fieber (28.8%, 37.3% bzw. 43.1%) und Kopfschmerzen (35.6%, 40.7% bzw. 41.4%). Husten (35.6%), Erbrechen (35.6%), Kopfschmerzen (35.6%), Arthralgie (28.8%) und Fieber (28.8%) waren die häufigsten UEs in der Placebogruppe. In der BMN 110 2.0 mg/kg/qow Gruppe wurden Kopfschmerzen (40.7%), Fieber (37.3%) und Erbrechen (35.6%) am häufigsten berichtet. Die häufigsten UEs in der BMN 110 2.0 mg/kg/Woche Gruppe waren Erbrechen (44.8%), Fieber (43.1%), und Kopfschmerzen (41.4%). Bei den mit BMN 110 einmal wöchentlich behandelten Patienten traten Fieber, Übelkeit und abdominale Schmerzen mindestens 10% häufiger auf als in der Placebo-Gruppe. Die Inzidenz der UEs Arthralgie (28.8% Placebo, 15.3% BMN 110 2.0 mg/kg/qow, 17.2% BMN 110 2.0 mg/kg/Woche) und Ermüdung (25.4%, 13.6% bzw. 15.5%) war in der Placebogruppe mindestens 10% höher als in den BMN 110-Gruppen.

Die Verteilung der SOCs war in den Gruppen vergleichbar. Die SOCs mit den häufigsten UEs waren Infektionen und parasitäre Erkrankungen (66.1% Placebo, 71.2% BMN 110 2.0 mg/kg/qow, 67.2% BMN 110 2.0 mg/kg/Woche), allgemeine Erkrankungen und Beschwerden am Verabreichungsort (62.7%, 67.8% bzw. 65.5%) und Erkrankungen des Gastrointestinaltrakts (69.5%, 67.8% bzw. 63.8%).

Tabelle 4-56: Ergebnisse für den Endpunkt „Inzidenz der Therapie-assoziierten UEs“ (MedDRA PT) aus RCT mit dem zu bewertenden Arzneimittel (Sicherheitspopulation) (9)

UE	Placebo (n=59) n (%)	BMN 110 2.0 mg/kg qow (n=59) n (%)	BMN 110 2.0 mg/kg/Woche (n=58) n (%)
<b>Therapie-assoziierte UEs stratifiziert nach SOC(≥5% in einer der Gruppen)<sup>1</sup></b>			
Erkrankungen des Gastrointestinaltrakts	15 (25.4%)	13 (22.0%)	27 (46.6%)
Erbrechen	4 (6.8%)	7 (11.9%)	18 (31.0%)
Übelkeit	4 (6.8%)	7 (11.9%)	14 (24.1%)

<b>UE</b>	<b>Placebo (n=59) n (%)</b>	<b>BMN 110 2.0 mg/kg qow (n=59) n (%)</b>	<b>BMN 110 2.0 mg/kg/Woche (n=58) n (%)</b>
Abdominalschmerzen	1 (1.7%)	1 (1.7%)	6 (10.3%)
Diarrhöe	3 (5.1%)	2 (3.4%)	3 (5.2%)
Oberbauchschmerzen	0	0	7 (12.1%)
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort	17 (28.8%)	21 (35.6%)	26 (44.8%)
Fieber	8 (13.6%)	8 (13.6%)	19 (32.8%)
Schüttelfrost	1 (1.7%)	4 (6.8%)	6 (10.3%)
Beschwerden im Brustraum	0	1 (1.7%)	3 (5.2%)
Ermüdung („fatigue“)	2 (3.4%)	1 (1.7%)	6 (10.3%)
Erkrankungen des Nervensystems	10 (16.9%)	13 (22.0%)	20 (34.5%)
Kopfschmerzen	9 (15.3%)	8 (13.6%)	15 (25.9%)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	1 (1.7%)	9 (15.3%)	12 (20.7%)
Dyspnoe	1 (1.7%)	6 (10.2%)	2 (3.4%)
Oropharyngealschmerz	0	0	4 (6.9%)
Reizung Rachen	0	0	3 (5.2%)
Vaskuläre Erkrankungen	5 (8.5%)	9 (15.3%)	10 (17.2%)
Hitzewallung („hot flush“)	0	4 (6.8%)	2 (3.4%)
Hypertension	3 (5.1%)	2 (3.4%)	2 (3.4%)
Gesichtsrötung („flushing“)	0	1 (1.7%)	5 (8.6%)
Hypotension	1 (1.7%)	1 (1.7%)	3 (5.2%)
Erkrankungen der Haut und des Unterhautzellgewebes	8 (13.6%)	10 (16.9%)	9 (15.5%)
Ausschlag	2 (3.4%)	3 (5.1%)	3 (5.2%)
Urtikaria	0	3 (5.1%)	3 (5.2%)
Pruritus	1 (1.7%)	2 (3.4%)	4 (6.9%)
Herzerkrankungen	5 (8.5%)	1 (1.7%)	5 (8.6%)
Tachykardie	4 (6.8%)	1 (1.7%)	2 (3.4%)
Untersuchungen	7 (11.9%)	6 (10.2%)	5 (8.6%)
Erhöhte Körpertemperatur	1 (1.7%)	0	3 (5.2%)
Erhöhte Atemfrequenz	3 (5.1%)	1 (1.7%)	1 (1.7%)
Infektionen und parasitäre Erkrankungen	0 (0.0%)	1 (1.7%)	4 (6.9%)

UE	Placebo (n=59) n (%)	BMN 110 2.0 mg/kg qow (n=59) n (%)	BMN 110 2.0 mg/kg/Woche (n=58) n (%)
Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen	3 (5.1%)	3 (5.1%)	4 (6.9%)
Schmerzen in den Extremitäten	0	3 (5.1%)	0
Arthralgie	3 (5.1%)	1 (1.7%)	2 (3.4%)
Psychiatrische Erkrankungen	0 (0.0%)	0 (0.0%)	4 (6.9%)
Agitation	0	0	3 (5.2%)
Erkrankungen des Immunsystems	0 (0.0%)	4 (6.8%)	1 (1.7%)
Überempfindlichkeit	0	3 (5.1%)	1 (1.7%)

MedDRA Version 15.0

<sup>1</sup> Patienten mit mehr als einem UE innerhalb eines MedDRA PTs wurden nur einmal gezählt.

Die häufigsten UEs, bei denen ein Zusammenhang mit der Therapie vermutet wurde, waren Fieber (13.6% Placebo, 13.6% BMN 110 2.0 mg/kg/qow, 32.8% BMN 110 2.0 mg/kg/Woche), Erbrechen (6.8%, 11.9% bzw. 31.0%), Kopfschmerzen (15.3%, 13.6% bzw. 25.9%) und Übelkeit (6.8%, 11.9% bzw. 24.1%). Die häufigsten mit BMN 110 in Zusammenhang stehenden UEs (Fieber, Erbrechen, Kopfschmerzen und Übelkeit) entsprachen den für ERTs bekannten Reaktionen. Einige der Therapie-assoziierte UEs traten in der BMN 110-Gruppe (einmal wöchentlich, BMN 110 qow nicht getestet) häufiger auf als unter Placebo (Tabelle 4-57). Unter Berücksichtigung, dass es sich bei der Mehrzahl der UEs um für ERTs bekannte, nicht schwerwiegende, temporär auftretende Infusionsreaktionen milder bis mittelschwerer Ausprägung handelte, die durch geeignete Maßnahmen kontrollierbar waren (siehe Kapitel 4.3.1.3.9), ist der Schaden von BMN 110 im Vergleich zu Placebo als gering zu bewerten. Die hohe Adhärenz der Studienpopulation zeugte von der guten Verträglichkeit der Interventionen. Die seitens des IQWiG vorgeschlagenen Kriterien zur Feststellung des Ausmaßes des Zusatznutzens oder Schadens kommen bei Arzneimitteln für seltene Leiden nicht zur Anwendung.

Tabelle 4-57: Endpunkt „Inzidenz der mit der Behandlung in Zusammenhang stehenden UEs“: Placebo vs. BMN 110 2.0 mg/kg/Woche (9)

SOC / UE (PT) <sup>1</sup>	Placebo n/total	BMN 110 jede Woche n/total
SOC Gastrointestinaltrakt	15/59	27/58
Erbrechen	4/59	18/58
Übelkeit	4/59	14/58
Oberbauchschmerzen	0/59	7/58
SOC Allgemeine Erkrankungen / Verabreichungsort	17/59	26/58
Fieber	8/59	19/58
SOC Nervensystem	10/59	20/58
Kopfschmerzen	9/59	15/58
SOC Atemwege, Brustraum, Mediastinum	1/59	12/58
SOC Vaskuläre Erkran- kungen	5/59	10/58

<sup>1</sup> SOC/UEs mit einer Inzidenz von  $\geq 10\%$ . Berücksichtigt wurden UEs, deren Inzidenz sich um mindestens 10% zwischen Placebo und der BMN 110-Gruppe unterschied.

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

Nicht zutreffend

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.

#### 4.3.1.3.9 Inzidenz von Überempfindlichkeits- und mit der Infusion assoziierten Reaktionen (IARs) – 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 0). 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 4-58: Operationalisierung des Endpunkts „Überempfindlichkeitsreaktionen und IARs“

Studie	Operationalisierung
MOR-004	<p>Hypersensitivitätsreaktionen einschließlich Anaphylaxie und weniger schwere allergische Reaktionen wurden als potentielle Risiken besonderen Interesses von BMN 110 identifiziert. Eine Anaphylaxie war entsprechend der seitens des „National Institute of Allergy and Infectious Disease“ und des „Food Allergy and Anaphylaxis Network“ vorgeschlagenen Kriterien definiert.</p> <p>Die Inzidenz von Infusionsreaktionen (IAR, „infusion-associated reactions“) wurde nach Schwere und Behandlungsgruppe dokumentiert. Alle UEs einschließlich Anaphylaxie, Überempfindlichkeit und allergische Reaktionen, die nach Start der Infusion oder innerhalb eines Tages nach Ende der Infusion auftraten, wurden als IAR eingestuft, unabhängig davon, ob das Ereignis nach Meinung des Prüfarztes mit der Behandlung in Zusammenhang stand oder nicht. Symptome einer Infusionsreaktion konnten Fieber, Schüttelfrost/Rigor, Hautsymptome (Urtikaria, Angioödem, Rash), respiratorische Symptome (Dyspnoe, Keuchen, Stridor), gastrointestinale Symptome (Übelkeit, Erbrechen, abdominale Schmerzen) und/oder kardiovaskuläre Veränderungen (Hypotension, Hypertension) einschließen. Eine IAR konnte auch dann vorliegen, wenn ein Patient z.B. innerhalb eines Tages nach der Infusion stürzte und sich Verstauchungen zuzog. Den Studienzentren wurde ein SMP („Safety Management Plan“) zur Verfügung gestellt, der die Prüfarzte dabei unterstützen sollte, Überempfindlichkeitsreaktionen und IARs zu erkennen, zu handhaben, zu klassifizieren und zu berichten. Zudem fanden monatliche Telefonkonferenzen zwischen dem klinischen Monitor und den Prüfarzten statt.</p> <p>Ein unabhängiges Komitee („Allergic Reaction Review Board“, ARRB) bestehend aus Ärzten, die nicht unmittelbar an der Studie beteiligt waren, wurde eingesetzt, um schwere und schwerwiegende Infusionsreaktionen zu begutachten. Die Daten zu Infusionsreaktionen wurden hinsichtlich des Zeitpunkts ihres Auftretens (während der Infusion), der Notwendigkeit, die Infusion zu unterbrechen oder abubrechen, der Notwendigkeit, die Studie zu unterbrechen oder abubrechen und der Notwendigkeit medizinischer Interventionen (i.v. Steroide, i.v. Antihistaminika, i.v. Lösungen oder Sauerstoff) ausgewertet. Um mögliche Überempfindlichkeitsreaktionen einschließlich Anaphylaxie zu identifizieren, wurde eine breite standardisierte MedDRA-Abfrage (Algorithmen: anaphylaktische Reaktion und Angioödem) unter Verwendung zuvor festgelegter PTs initiiert. Die Algorithmen, die verwendet wurden, um spezifische PTs zu identifizieren, unterschieden sich zwischen den beiden Suchabfragen. Es war möglich, dass eine bestimmte Gruppe von UEs vorliegen musste, um als relevantes PT identifiziert zu werden. Daher konnte ein PT wie z.B. Urtikaria durch beide Recherchen identifiziert werden, aber nicht notwendigerweise mit derselben Inzidenz und Häufigkeit. Zusätzlich wurden Infusionsreaktionen („Infusion-related reactions“) nach MedDRA auf PT-Ebene recherchiert.</p> <p>Blutproben zur Bestimmung der Immungenität wurden zum Ausgangspunkt, Baseline, in Woche 2 und 4 und danach alle 4 Wochen entnommen. Die Immungenität wurde unter Einsatz validierter Tests bei allen Patienten geprüft. Diese schlossen die gesamten Antikörper, neutralisierende Antikörper (NAk), IgG, IgE, IgM, und gesamtes IgE (nur Ausgangswert) ein. NAK, IgG und IgM wurden nicht bestimmt, wenn der Test auf gesamte Antikörper negativ war. Diese Daten werden aus rein informativen Gründen berichtet und fließen nicht in die Bewertung des Ausmaßes des Zusatznutzens ein, da die Patientenrelevanz nicht ableitbar ist.</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 4-59: Bewertung des Verzerrungspotenzials für den Endpunkt „Überempfindlichkeitsreaktionen und IARs“ 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
MOR-004 (9)	niedrig	ja	ja	ja	ja	niedrig

*Begründen Sie für jede Studie die abschließende Einschätzung.*

Das Verzerrungspotential für den Endpunkt „Inzidenz behandlungsspezifischer und schwerwiegender behandlungsspezifischer UEs“ wird unter Berücksichtigung der in Tabelle 4-59 aufgeführten Faktoren und den Bewertungsbögen zur Einschätzung von Verzerrungsaspekten im Anhang 4-F als niedrig eingeschätzt. Die Gründe für diese Einschätzung sind in Abschnitt 4.3.1.3.8 dargestellt.

*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 4-60: Ergebnisse für Endpunkt „Inzidenz von Überempfindlichkeitsreaktionen“ aus RCT mit dem zu bewertenden Arzneimittel (9)

UE	Placebo (n=59) n (%)	BMN 110 2.0 mg/kg/qow (n=59) n (%)	BMN 110 2.0 mg/kg/Woche (n=58) n (%)
Gesamte Überempfindlichkeitsreaktionen <sup>1</sup>	7 (11.9%)	16 (27.1%)	12 (20.7%)
Gesamte anaphylaktische Reaktionen SMQ <sup>1</sup>	1 (1.7%)	2 (3.4%)	3 (5.2%)
Gesichtsrotung („Flushing“)	0	1 (1.7%)	2 (3.4%)
Husten	1 (1.7%)	0	1 (1.7%)
Dyspnoe	0	1 (1.7%)	1 (1.7%)
Hypotension	0	0	1 (1.7%)
Urtikaria	0	0	1 (1.7%)
Anaphylaktische Reaktion	0	1 (1.7%)	0
Lippenschwellung	1 (1.7%)	0	0
Angioödem SMQ <sup>1</sup> (gesamt)	7 (11.9%)	14 (23.7%)	10 (17.2%)
Urtikaria	0	4 (6.8%)	4 (6.9%)
Überempfindlichkeit	1 (1.7%)	4 (6.8%)	3 (5.2%)

UE	Placebo (n=59) n (%)	BMN 110 2.0 mg/kg/qow (n=59) n (%)	BMN 110 2.0 mg/kg/Woche (n=58) n (%)
Ödem Augenlid	0	0	1 (1.7%)
Obstruktive Atemwegserkrankung	0	0	1 (1.7%)
Peripheres Ödem	2 (3.4%)	4 (6.8%)	1 (1.7%)
Engegefühl im Rachen	0	0	1 (1.7%)
„Wheezing“	1 (1.7%)	0	1 (1.7%)
Aurikularschwellung	1 (1.7%)	0	0
Lippenschwellung	1 (1.7%)	1 (1.7%)	0
Nasale Obstruktion	2 (3.4%)	1 (1.7%)	0
Ödem	1 (1.7%)	0	0
Stridor	0	1 (1.7%)	0
Mit der Behandlung in Zusammenhang stehende Überempfindlichkeitsreaktionen <sup>1,2</sup>	2 (3.4%)	9 (15.3%)	7 (12.1%)
Mit der Behandlung in Zusammenhang stehende anaphylaktische Reaktionen SMQ <sup>1</sup>	0 (0.0%)	2 (3.4%)	3 (5.2%)
Gesichtsrötung („Flushing“)	0 (0.0%)	1 (1.7%)	2 (3.4%)
Dyspnoe	0 (0.0%)	1 (1.7%)	1 (1.7%)
Hypotension	0 (0.0%)	0 (0.0%)	1 (1.7%)
Urtikaria	0 (0.0%)	0 (0.0%)	1 (1.7%)
Anaphylaktische Reaktion	0 (0.0%)	1 (1.7%)	0 (0.0%)
Mit der Behandlung in Zusammenhang stehende Angioödeme SMQ <sup>1</sup>	2 (3.4%)	7 (11.9%)	5 (8.6%)
Urtikaria	0 (0.0%)	3 (5.1%)	3 (5.2%)
Überempfindlichkeit	0 (0.0%)	3 (5.1%)	1 (1.7%)
Engegefühl im Rachen	0 (0.0%)	0 (0.0%)	1 (1.7%)
Lippenschwellung	0 (0.0%)	1 (1.7%)	0 (0.0%)
Peripheres Ödem	1 (1.7%)	0 (0.0%)	0 (0.0%)
„Wheezing“	1 (1.7%)	0 (0.0%)	0 (0.0%)

MedDRA Version 15.0

Inzidenz: Patienten, die mehr als ein UE innerhalb eines MedDRA PT entwickelten, wurden nur einmal gezählt.

<sup>1</sup> Überempfindlichkeitsreaktionen wurden mittels einer standardisierten MedDRA-Abfrage (SMQ) unter Einsatz der Algorithmen „Anaphylaktische Reaktionen“ und „Angioödem“ identifiziert.

<sup>2</sup> Die UEs wurden von den Prüfern als möglicherweise oder wahrscheinlich mit den Prüfpräparaten assoziiert klassifiziert.

Wie bei jeder Proteininfusion wurden Überempfindlichkeitsreaktionen erwartet und wurden mittels zuvor festgelegter Algorithmen identifiziert. Überempfindlichkeitsreaktionen traten in den BMN 110-Gruppen häufiger auf als in der Placebo-Gruppe auf (11.9% Placebo, 27.1% BMN 110 2.0 mg/kg/qow, 20.7% BMN 110 2.0 mg/kg/Woche). Die Inzidenz von Überempfindlichkeitsreaktionen war in der BMN 110 qow Gruppe höher als bei einmal wöchentlicher Verabreichung von BMN 110 und somit zeigte sich keine Dosisabhängigkeit. Hinsichtlich der gesamten und der gesamten mit der Behandlung in Zusammenhang stehenden Überempfindlichkeitsreaktionen einschließlich anaphylaktische Reaktionen und Angioödeme ergab sich zwischen Placebo und BMN 110 einmal pro Woche kein signifikanter Unterschied (Tabelle 4-61). Ein mit BMN 110 qow behandelter Patient entwickelte eine anaphylaktische Reaktion (SUE). Dieser Patient tolerierte nachfolgende BMN 110-Injektionen ohne Wiederauftreten der Symptome. Die häufigsten UEs (nach PT) waren Urtikaria (0 Placebo, 4 BMN 110 2.0 mg/kg/qow, 4 BMN 110 2.0 mg/kg/Woche) und Überempfindlichkeitsreaktionen (1, 4 bzw. 3). Nur bei zwei Patienten wurde die Infusion aufgrund von Überempfindlichkeitsreaktionen unterbrochen (n=1 pro BMN 110-Gruppe). Mit einer Ausnahme (eine Infusion, BMN 110 qow) erhielten alle Patienten wie geplant die nachfolgenden Infusionen. Keiner der Patienten entwickelte wiederholt eine schwere Überempfindlichkeitsreaktion. Bei nur wenigen Patienten (n=9) trat wiederholt eine milde oder mittelschwere Überempfindlichkeitsreaktion auf (n=2 Placebo, n=4 bzw. n=3 BMN 110 qow bzw. einmal wöchentlich).

Tabelle 4-61: Endpunkt „Inzidenz von Überempfindlichkeitsreaktionen“: Placebo vs. BMN 110 (9)

UE	Placebo n/total	BMN 110 qow n/total	BMN 110 jede Woche n/total
Gesamt	7/59	16/59	12/58
Gesamt anaphylaktisch	1/59	2/59	3/58
Gesamt Angio- ödem	7/59	14/59	10/58
Gesamt assoziiert	2/59	9/59	7/58
Gesamt anaphylaktisch assoziiert	0/59	2/59	3/58
Gesamt Angio- ödem assoziiert	2/59	7/59	5/58

Tabelle 4-62: Ergebnisse für Endpunkt „IARs“ aus RCT mit dem zu bewertenden Arzneimittel

UE	Placebo (N=59) n (%)	BMN 110 2.0 mg/kg qow (N=59) n (%)	BMN 110 2.0 mg/kg/Woche (N=58) n (%)
Gesamte IARs <sup>1</sup>	54 (91.5%)	56 (94.9%)	52 (89.7%)
Mild	41 (69.5%)	38 (64.4%)	29 (50.0%)
Moderat	12 (20.3%)	17 (28.8%)	21 (36.2%)
Schwer	1 (1.7%)	1 (1.7%)	2 (3.4%)
Gesamte IARs <sup>1</sup> während der Infusion	48 (81.4%)	53 (89.8%)	52 (89.7%)
Gesamte schwerwiegende IARs <sup>1</sup> während der Infusion	0 (0.0%)	2 (3.4%)	2 (3.4%)
Gesamte IARs <sup>1</sup> , die zur Unterbrechung der Infusion führten	8 (13.6%)	21 (35.6%)	18 (31.0%)
Gesamte IARs <sup>1</sup> , die zum Abbruch der Infusion führten	1 (1.7%)	4 (6.8%)	6 (10.3%)
Patienten mit IARs <sup>1</sup> , die zur Unterbrechung oder zum Abbruch der Infusion führten und eine medizinische Intervention <sup>2</sup> erforderten	0 (0.0%)	9 (15.3%)	13 (22.4%)
Infusionen, die unterbrochen oder abgebrochen wurden und eine medizinische Intervention <sup>2</sup> erforderten	0 (0.0%) (1.407 Infusionen)	9 (0.64%) (1.400 Infusionen)	17 (1.26%) (1.345 Infusionen)
Häufigste IARs <sup>1</sup> (≥10% in einer der Gruppen)			
Erbrechen	9 (15.3%)	14 (23.7%)	22 (37.9%)
Fieber	11 (18.6%)	13 (22.0%)	21 (36.2%)
Kopfschmerzen	12 (20.3%)	19 (32.2%)	19 (32.8%)
Übelkeit	8 (13.6%)	10 (16.9%)	16 (27.6%)
Oberbauchschmerzen	5 (8.5%)	1 (1.7%)	8 (13.8%)
Husten	5 (8.5%)	9 (15.3%)	8 (13.8%)
Ermüdung („fatigue“)	7 (11.9%)	2 (3.4%)	8 (13.8%)
Abdominalschmerzen	2 (3.4%)	6 (10.2%)	7 (12.1%)
Oropharyngealschmerzen	3 (5.1%)	3 (5.1%)	7 (12.1%)
Schüttelfrost	1 (1.7%)	5 (8.5%)	6 (10.3%)
Diarrhoe	7 (11.9%)	6 (10.2%)	6 (10.3%)
Arthralgie	12 (20.3%)	4 (6.8%)	5 (8.6%)
Schmerzen in den Extremitäten	6 (10.2%)	8 (13.6%)	4 (6.9%)
Dyspnoe	1 (1.7%)	6 (10.2%)	3 (5.2%)

MedDRA Version 15.0

Inzidenz: Patienten, die mehr als ein UE innerhalb eines MedDRA PT entwickelten, wurden nur einmal gezählt.

<sup>1</sup> IARs (MedDRA PT) waren als mit der Verabreichung der Prüfpräparate assoziiert eingestuft wenn sie nach Beginn der Infusion oder innerhalb eines Tages nach Ende der Infusion auftraten (unabhängig davon, ob ein Zusammenhang mit den Prüfpräparaten bestand oder nicht).

<sup>2</sup> Eine medizinische Intervention war als zumindest eine der folgenden definiert: Antihistaminika i.v., Steroide i.v., Flüssigkeiten i.v. oder Sauerstoff, welche mittels der WHO-Kodierung ermittelt wurden.

IARs wurden als mit der Behandlung in Zusammenhang stehend eingestuft, wenn sie während der Infusion oder innerhalb eines Tages nach Ende der Infusion auftraten, unabhängig davon, ob die Prüfarzte einen Zusammenhang mit den Prüfpräparaten angegeben hatten oder nicht. Bei der Mehrzahl der Patienten trat unabhängig von der Gruppenzugehörigkeit mindestens eine IAR auf (n=54, 91.5%, 291 Ereignisse Placebo, n=56, 94.9%, 393 Ereignisse BMN 110 qow, n=52, 89.7%, 511 Ereignisse BMN 110 einmal wöchentlich). Es ergab sich weder hinsichtlich der gesamten IARs, der IARs während der Infusion noch der gesamten schwerwiegenden IARs ein signifikanter Gruppenunterschied (Tabelle 4-63). Die häufigsten UEs (Erbrechen, Fieber und Kopfschmerzen) waren gleichzeitig auch die häufigsten IARs, welche in den meisten Fällen mittels symptomatischer Therapie oder durch Änderung der Infusionsrate kontrollierbar waren. Hinsichtlich der häufigsten IARs, deren Inzidenz sich um mindestens 10% zwischen Placebo und den BMN 110-Gruppen unterschied, ergab sich lediglich bei den IARs Erbrechen und Fieber ein signifikanter Unterschied ( $p < 0.05$ ) zwischen BMN 110 (einmal wöchentlich) und Placebo. Die IAR Arthralgie trat unter Behandlung mit Placebo häufiger auf als unter Behandlung mit BMN 110 qow (Tabelle 4-63). Bei den mit BMN 110 behandelten Patienten (35.6% BMN 110 2.0 mg/kg qow, 31.0% BMN 110 2.0 mg/kg/Woche) führten IARs häufiger zur Unterbrechung oder zum Abbruch der Infusion als in der Placebo-Gruppe (13.6%). Bei 1.7%, 6.8% bzw. 10.3% der Patienten, denen Placebo oder BMN 110 jede zweite Woche bzw. einmal wöchentlich verabreicht wurde, wurde die Infusion am Tag der Visite aufgrund einer IAR abgebrochen. Der Anteil der Patienten, die ein UE während der Infusion entwickelten, welches entweder eine Unterbrechung oder einen Abbruch der Infusion sowie eine medizinische Intervention (meist i.v. Antihistaminika oder i.v. Steroide) erforderte, betrug 0% in der Placebo-Gruppe, 15.3% (n=9) in der BMN 110 qow und 22.4% (n=13) bei den Patienten, die einmal wöchentlich mit BMN 110 behandelt wurden. Die Unterschiede (Unterbrechung oder Abbruch der Infusion plus medizinische Intervention) waren im Gruppenvergleich vs. Placebo signifikant ( $p < 0.05$ ; Tabelle 4-63). Nur wenige Infusionen wurden unterbrochen oder abgebrochen und erforderten gleichzeitig eine medizinische Intervention, um die IAR zu kontrollieren (0%/1.407 Infusionen Placebo, 0.64%/1.400 Infusionen in BMN 110 qow und 1.26%/1.345 Infusionen BMN 110 einmal pro Woche). Nur jeweils drei der gesamten BMN 110-Infusionen (qow bzw. einmal wöchentlich) wurden abgebrochen und keiner der Patienten brach die Studie aufgrund von UEs einschließlich IARs ab. Alle Patienten, die eine IAR entwickelt hatten, erhielten und tolerierten spätere Infusionen.

Das Sicherheitsprofil von BMN 110 entsprach demjenigen anderer ERTs. Mit Ausnahme von IARs ergab sich kein kausaler Zusammenhang zwischen der Behandlung und klinisch bedeutungsvollen Änderungen der Laborparameter und Vitalzeichen.

Tabelle 4-63: Endpunkt „Inzidenz von IAR“: Placebo vs. BMN 110 (9)

UE	Placebo n/total	BMN 110 qow n/total	BMN 110 jede Woche n/total
Gesamte IARs	54/59	56/59	52/58
Gesamte IARs während Infusion	48/59	53/59	52/58
Gesamte schwer- wiegende IARs während Infusion	0/59	2/59	2/58
Gesamte IARs (Unterbrechung/ Abbruch Infusion plus medizinische Intervention)	0/59	9/59	13/58
Häufigste IARs <sup>1</sup>			
Erbrechen	9/59	14/59	22/58
Fieber	11/59	13/59	21/58
Kopfschmerzen	12/59	19/59	19/58
Übelkeit	8/59	10/59	16/58
Arthralgia	12/59	4/59	5/58

<sup>1</sup> IARs mit einer Inzidenz von  $\geq 10\%$ . Berücksichtigt wurden IARs, deren Inzidenz sich um mindestens 10% zwischen Placebo und einer der BMN 110-Gruppen unterschied.

Alle Patienten der BMN 110-Gruppen entwickelten BMN 110-spezifische Antikörper und 98.2% bis 100% blieben während der Studie positiv. Bei den Patienten, die BMN 110 einmal pro Woche erhielten, wurden die Antikörper etwas früher detektiert als bei Behandlung alle zwei Wochen (100% positiv in Woche 4 vs. 100% positiv in Woche 16), allerdings waren die mittleren Antikörper-Titer in Woche 24 in beiden BMN 110-Gruppen vergleichbar. Bei etwa 20% der Patienten waren bereits zu Beginn der Studie Antikörper gegen BMN 110 nachweisbar (15.5% bis 25.4% in den Dosis-Gruppen). Die Antikörpertiter waren zwischen den Gruppen, bei denen zu Beginn der Studie Antikörper detektiert wurden bzw. nicht nachweisbar waren, in Woche 24 vergleichbar.

Etwa 80% der mit BMN 110 behandelten Patienten entwickelten innerhalb von 24 Wochen neutralisierende Antikörper (NAK) gegen BMN 110, wobei die Inzidenz in beiden BMN 110-Gruppen vergleichbar war. Die Rate der NAK-positiven Visiten (berechnet als Prozentsatz der NAK-positiven Visiten nach der Eingangsvisite) betrug >50% der Visiten bei 74.6% der Patienten der BMN 110 qow Gruppe und 93.1% der Patienten, denen BMN 110 einmal wöchentlich injiziert wurde.

BMN 110-spezifische IgE wurden bei 6.8% (4/59) bzw. 8.6% (5/58) der Patienten, die BMN 110 jede zweite Woche bzw. einmal wöchentlich erhalten hatten, detektiert. Bei einem Patienten (BMN 110 einmal pro Woche) wurden BMN 110-spezifische IgE zur Eingangsvisite und in Woche 8 nachgewiesen. In allen anderen Fällen wurden BMN 110-spezifische IgE im Studienverlauf bis Woche 24 jeweils nur einmal beobachtet und traten somit nur temporär auf.

Ein Zusammenhang zwischen dem Auftreten BMN 110-spezifischer Antikörper oder NAK und den Ergebnissen des 6MWT, MSCT oder MVV in Woche 24 ergab sich nicht. Die Ergebnisse zur Wirksamkeit waren zwischen den Patienten, bei denen zur Eingangsvisite Antikörper nachgewiesen bzw. nicht nachgewiesen wurden, vergleichbar. Es ergab sich kein offensichtlicher Zusammenhang zwischen der prozentualen Änderung des normalisierten uKS und den Antikörper-Titern oder dem Auftreten von NAK. Bei den Patienten, denen BMN 110 einmal wöchentlich injiziert wurde und die NAK entwickelten, war die Clearance vermindert und die Halbwertszeit von BMN 110 verlängert. Dies hatte allerdings keinen Einfluss auf die Wirksamkeit oder Sicherheit von BMN 110. Höhere Antikörper- oder NAK-Titer waren nicht mit einer höheren Inzidenz von Überempfindlichkeitsreaktionen assoziiert. Keiner der IgE-positiven Patienten entwickelte ein schweres UE. Bei keinem der drei Patienten, bei denen ein mit BMN 110-assoziiertes SUE (Hypersensitivität, Erbrechen und anaphylaktische Reaktion) aufgetreten war, wurden während der Studie IgE detektiert.

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

Nicht zutreffend.

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.

#### **4.3.1.3.10 Subgruppenanalysen – RCT**

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

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

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

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

Die vorliegende Nutzenbewertung bezieht sich auf die Gesamtpopulation der in die Studie MOR-004 eingeschlossenen Patienten.

Alle unten dargestellten Subgruppenanalysen der Studie MOR-004 (9) waren a priori geplant (6MWT-Kategorie, Altersgruppen, Rasse und Region). In den Subgruppenanalysen wurde das in der Primäranalyse verwendete ANCOVA-Modell mit einer zusätzlichen Kovariate (Interaktion nach Subgruppe) eingesetzt. Eine statistische Hypothesentestung wurde nicht formuliert bzw. durchgeführt. Der Trennpunkt  $\leq 200$  m bzw.  $> 200$  m im 6MWT wurde gewählt, da nach diesen Merkmalen bei der Randomisierung stratifiziert wurde. Die gewählten Altersgruppen entsprechen der üblichen Einteilung in Kinder (Einschlusskriterium:  $\geq 5$  Jahre), Jugendliche und Erwachsene und dienten ebenfalls zur Stratifizierung. In einer multi-zentrischen Beobachtungsstudie bei MPS IVA-Patienten (MorCAP) nahm die Gehstrecke im 6MWT mit zunehmendem Alter ab ( $232.5 \pm 140.1$  bei 5-11 Jahre alten Patienten,  $181.2 \pm 177.3$  bei 12-18 Jahre alten Patienten,  $193.1 \pm 148.5$  bei Patienten, die älter als 18 Jahre waren) (3).

Keiner der Interaktionstests der Subgruppenanalysen lieferte ein signifikantes Ergebnis. Die Ergebnisse der einzelnen Subgruppen sind in den nachfolgenden Tabellen dargestellt.

Tabelle 4-64: Subgruppenanalysen für den Endpunkt „6MWT“ (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel in der ITT-Population (9)

<b>MOR-004</b>	<b>N</b>	<b>Ausgangsge- strecke (m) LSM (± SE)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b>≤200 m Baseline 6MWT</b>					
Placebo	23	137.5 (9.99)	13.5 (-7.5, 34.4)	-	-
BMN 110 qow	24	123.0 (9.76)	18.5 (-1.9, 38.9)	5.0 (-24.0, 34.0)	0.7324
BMN 110 pro Woche	23	125.5 (9.98)	53.8 (33.0, 74.7)	40.4 (11.0, 69.8)	0.0074
<b>&gt;200 m Baseline 6MWT<sup>1</sup></b>					
Placebo	36	260.5 (5.68)	13.5 (-3.1, 30.1)	-	-
BMN 110 qow	35	262.3 (5.75)	11.3 (-5.6, 28.2)	-2.2 (-25.8, 21.4)	0.8550
BMN 110 pro Woche	35	254.4 (5.76)	24.3 (7.4, 41.3)	10.8 (-12.8, 34.4)	0.3657
<b>6MWT &gt;200 m - 6MWT ≤200 m</b>					
BMN 110 qow	-	-	-	-7.22 (-44.60,30.16)	0.7034
BMN 110 pro Woche	-	-	-	-29.54 (-67.18,8.10)	0.1232
<b>Altersgruppe 5-11 Jahre</b>					
Placebo	30	238.4 (5.82)	22.3 (4.0, 40.6)	-	-
BMN 110 qow	31	218.7 (5.72)	10.6 (-7.3, 28.6)	-11.7 (-37.2, 13.8)	0.3678
BMN 110 pro Woche	32	228.9 (5.63)	36.1 (18.4, 53.8)	13.8 (-11.5, 39.1)	0.2844
<b>Altersgruppe 12-18 Jahre</b>					
Placebo	15	196.1 (8.23)	-6.5 (-32.3, 19.2)	-	-
BMN 110 qow	16	199.6 (7.96)	17.2 (-7.7, 42.2)	23.8 (-12.0, 59.6)	0.1914
BMN 110 pro Woche	16	187.3 (7.96)	41.7 (16.8, 66.6)	48.2 (12.4, 84.0)	0.0086
<b>Altersgruppe ≥19 Jahre</b>					
Placebo	14	171.8 (8.51)	16.5 (-10.3, 43.2)	-	-
BMN 110 qow	12	180.1 (9.20)	18.8 (-10.1, 47.7)	2.3 (-36.8, 41.5)	0.9063
BMN 110 pro Woche	10	150.7 (10.07)	26.9 (-4.9, 58.7)	10.4 (-30.9, 51.8)	0.6194
<b>Alter 12-18 Jahre – Alter 5-11 Jahre</b>					
BMN 110 qow	-	-	-	35.45 (-8.50,79.39)	0.1132 0.1224
BMN 110 pro Woche	-	-	-	34.46 (-9.36,78.29)	
<b>Alter ≥19 – Alter 5-11 Jahre</b>					
BMN 110 qow	-	-	-	14.01 (-32.74,60.75)	0.5549

<b>MOR-004</b>	<b>N</b>	<b>Ausgangsgeh- strecke (m) LSM (± SE)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
BMN 110 pro Woche	-	-	-	-3.34 (-51.83,45.15)	0.8921
<b>Weibliche Patienten</b>					
Placebo	32	198.9 (5.92)	17.6 (-0.0, 35.2)	-	-
BMN 110 qow	25	199.2 (6.70)	12.5 (-7.4, 32.5)	-5.1 (-31.7, 21.6)	0.7076
BMN 110 pro Woche	32	196.9 (5.92)	36.2 (18.6, 53.9)	18.6 (-6.3, 43.6)	0.1423
<b>Männliche Patienten</b>					
Placebo	27	227.3 (6.45)	8.8 (-10.4, 28.1)	-	-
BMN 110 qow	34	210.4 (5.74)	15.3 (-1.9, 32.4)	6.4 (-19.3, 32.2)	0.6227
BMN 110 pro Woche	26	212.6 (6.57)	35.8 (16.2, 55.4)	27.0 (-0.4, 54.4)	0.0538
<b>Weiblich – männlich</b>					
BMN 110 qow	-	-	-	-12.22 (-49.62,25.18)	0.5197
BMN 110 pro Woche	-	-	-	-8.89 (- 46.21,28.43)	0.6388
<b>Rasse: Weiß</b>					
Placebo	44	209.4 (5.03)	16.5 (1.8, 31.3)	-	-
BMN 110 qow	35	194.1 (5.64)	10.6 (-6.0, 27.2)	-5.9 (-28.2, 16.3)	0.5981
BMN 110 pro Woche	36	193.2 (5.56)	48.0 (31.6, 64.3)	31.4 (9.4, 53.5)	0.0054
<b>Rasse: Nicht weiß</b>					
Placebo	15	219.0 (8.62)	4.9 (-20.5, 30.3)	-	-
BMN 110 qow	24	222.6 (6.81)	19.2 (-0.8, 39.3)	14.4 (-17.9, 46.6)	0.3810
BMN 110 pro Woche	22	221.6 (7.12)	16.5 (-4.5, 37.5)	11.6 (-21.2, 44.5)	0.4859
<b>Nicht weiß – weiß</b>					
BMN 110 qow	-	-	-	20.05 (-19.38,59.48)	0.3168
BMN 110 pro Woche	-	-	-	-20.22 (-60.06,19.61)	0.3176
<b>Region Nordamerika</b>					
Placebo	16	213.7 (8.31)	10.8 (-14.0, 35.5)	-	-
BMN 110 qow	16	191.0 (8.31)	20.2 (-4.8, 45.1)	9.4 (-25.8, 44.5)	0.5988
BMN 110 pro Woche	15	171.3 (8.59)	54.2 (28.3, 80.0)	43.4 (7.5, 79.3)	0.0180
<b>Region Europa</b>					
Placebo	27	213.3 (6.40)	17.1 (-2.0, 36.1)	-	-
BMN 110 qow	21	205.2 (7.26)	9.6 (-12.1, 31.3)	-7.5 (-36.3, 21.4)	0.6112
BMN 110 pro Woche	25	205.0 (6.65)	40.2 (20.4, 60.0)	23.2 (-4.4, 50.7)	0.0985

<b>MOR-004</b>	<b>N</b>	<b>Ausgangsgeh- strecke (m) LSM (± SE)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (m) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b>Region Andere</b>					
Placebo	16	207.7 (8.31)	10.5 (-14.3, 35.2)	-	-
BMN 110 qow	22	216.8 (7.09)	14.1 (-7.1, 35.2)	3.6 (-29.0, 36.2)	0.8279
BMN 110 pro Woche	18	229.7 (7.84)	15.1 (-8.3, 38.6)	4.7 (-29.4, 38.8)	0.7867
<b>Europa -Nordamerika</b>					
BMN 110 qow	-	-	-	-16.24 (-62.96,30.47)	0.4933
BMN 110 pro Woche	-	-	-	-19.11 (-66.09,27.86)	0.4228
<b>Andere - Nordamerika</b>					
BMN 110 qow	-	-	-	-5.37 (-53.88,43.13)	0.8271
BMN 110 pro Woche	-	-	-	-37.95 (-88.40,12.50)	0.1393

ANCOVA (Änderung von Baseline bis Woche 24) mit den Kovariaten Behandlung, Altersgruppen und Baseline 6MWT-Kategorien, Subgruppe und Interaktion Behandlung nach Subgruppe.  
p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell.

Die Mehrzahl der Subgruppen zeigte hinsichtlich der Änderung der Gehstrecke im 6MWT bis Woche 24 einen numerisch größeren Effekt in der BMN 110 2.0 mg/kg/Woche Gruppe und keinen Effekt der Gabe von BMN 110 alle zwei Wochen verglichen mit Placebo (Tabelle 4-64). Die Ergebnisse entsprechen somit in ihrer Ausrichtung mehrheitlich der Primäranalyse. Die p-Werte im Test auf Interaktionen schwankten zwischen 0.1132 und 0.8271 für BMN 110 2.0 mg/kg/qow vs. Placebo und zwischen 0.1224 und 0.8921 für BMN 110 2.0 mg/kg/Woche vs. Placebo. Der größte Unterschied des Behandlungseffektes zeigte sich in den Subgruppenanalysen der 6MWT-Kategorien ( $\leq 200$  m und  $> 200$  m,  $p=0.1232$ ). In beiden Subgruppen ( $\leq 200$  m,  $> 200$  m) war der Effekt in der BMN 2.0 mg/kg/Woche Gruppe größer als in der Placebo-Gruppe, wobei der Effekt in der Kategorie  $\leq 200$  m am deutlichsten war.

Einige Subgruppen-Analysen zum 6MWT in Woche 24 zeigten signifikante Unterschiede zwischen der Placebo-Gruppe und den Patienten, die BMN 110 einmal pro Woche erhalten hatten. Es handelte sich dabei um Patienten, deren Gehstrecke im 6MWT zu Beginn der Studie 200 m oder weniger war, um Patienten im Alter von 12-18 Jahren, die Subgruppe der weißen Patienten sowie Patienten aus Nordamerika. Die Aussagekraft dieser Ergebnisse wird durch die kleinen Gruppengrößen sowie den fehlenden Nachweis einer signifikanten Interaktion eingeschränkt, allerdings sind die Effekte in ihrer Ausprägung deutlich und weichen zum Teil von dem Gesamtergebnis ab (insbesondere 6MWT-Kategorie). In ihrer Gesamtheit deuten die Subgruppenergebnisse an, dass der Effekt von BMN 110 (einmal

wöchentlich) im 6MWT von der Ausgangsgehstrecke abhängen könnte (mit deutlicheren Effekten bei geringerer Mobilität zu Beginn der Studie), und von der ethnischen Herkunft (mit beträchtlich deutlicheren Effekten bei Kaukasiern).

In Tabelle 4-65 sind die Subgruppenergebnisse für den Endpunkt „3MSCT“ dargestellt.

Tabelle 4-65: Subgruppenanalysen für den Endpunkt „3MSCT“ (mittlere absolute Änderung von Baseline bis Woche 24) aus RCT mit dem zu bewertenden Arzneimittel in der ITT-Population (9)

<b>MOR-004</b>	<b>N</b>	<b>Ausgangswert im 3MSC (Stufen/min) LSM (<math>\pm</math> SE)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (Stufen/min) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (Stufen/min) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b><math>\leq 200</math> m Baseline 6MWT</b>					
Placebo	23	18.5 (2.20)	0.2 (-3.7, 4.1)	-	-
BMN 110 qow	24	15.5 (2.15)	1.3 (-2.6, 5.3)	1.2 (-4.0, 6.4)	0.6553
BMN 110 pro Woche	23	16.8 (2.20)	3.3 (-0.6, 7.3)	3.2 (-2.1, 8.4)	0.2368
<b>&gt;200 m Baseline 6MWT</b>					
Placebo	36	37.7 (2.09)	5.9 (2.8, 9.0)	-	-
BMN 110 qow	35	35.1 (2.11)	4.4 (1.3, 7.5)	-1.5 (-5.8, 2.7)	0.4766
BMN 110 pro Woche	35	37.6 (2.12)	5.8 (2.6, 8.9)	-0.2 (-4.4, 4.1)	0.9408
<b>6MWT &gt;200 m - 6MWT <math>\leq 200</math> m</b>					
BMN 110 qow	-	-	-	-2.70 (-9.38,3.97)	0.4253 0.3313
BMN 110 pro Woche	-	-	-	-3.32 (-10.06,3.41)	
<b>Altersgruppe 5-11 Jahre</b>					
Placebo	30	33.6 (1.19)	5.4 (2.2, 8.7)	-	-
BMN 110 qow	31	30.1 (1.17)	1.6 (-1.6, 4.8)	-3.8 (-8.4, 0.7)	0.0949
BMN 110 pro Woche	32	33.6 (1.15)	4.9 (1.8, 8.0)	-0.6 (-5.0, 3.9)	0.8049
<b>Altersgruppe 12-18 Jahre</b>					
Placebo	15	32.0 (1.69)	1.9 (-2.6, 6.5)	-	-
BMN 110 qow	16	25.7 (1.63)	2.6 (-1.8, 7.0)	0.7 (-5.6, 7.1)	0.8210
BMN 110 pro Woche	16	28.2 (1.63)	5.4 (1.0, 9.8)	3.5 (-2.8, 9.8)	0.2782
<b>Altersgruppe <math>\geq 19</math> Jahre</b>					
Placebo	14	20.1 (1.75)	1.8 (-3.0, 6.5)	-	-
BMN 110 qow	12	21.6 (1.89)	7.9 (2.7, 13.0)	6.1 (-0.8, 13.0)	0.0831
BMN 110 pro Woche	10	18.8 (2.06)	3.4 (-2.3, 9.1)	1.6 (-5.7, 9.0)	0.6577

<b>MOR-004</b>	<b>N</b>	<b>Ausgangswert im 3MSC (Stufen/min) LSM (<math>\pm</math> SE)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (Stufen/min) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (Stufen/min) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t- Test) BMN 110 vs. Placebo</b>
<b>Alter 12-18 Jahre – Alter 5-11 Jahre</b>					
BMN 110 qow	-		-	4.57 (-3.19,12.32)	0.2465
BMN 110 pro Woche	-		-	4.04 (-3.69,11.77)	0.3038
<b>Alter <math>\geq</math>19 Jahre – Alter 5-11 Jahre</b>					
BMN 110 qow	-		-	9.93 (1.68,18.19)	0.0187
BMN 110 pro Woche	-		-	2.21 (-6.37,10.78)	0.6122
<b>Weibliche Patienten</b>					
Placebo	32	26.6 (1.20)	2.6 (-0.4, 5.7)	-	-
BMN 110 qow	35	26.3 (1.36)	0.1 (-3.4, 3.5)	-2.6 (-7.2, 2.1)	0.2746
BMN 110 pro Woche	32	29.5 (1.20)	5.8 (2.7, 8.9)	3.2 (-1.2, 7.5)	0.1521
<b>Männliche Patienten</b>					
Placebo	27	34.1 (1.31)	4.9 (1.5, 8.3)	-	-
BMN 110 qow	34	27.8 (1.17)	5.4 (2.4, 8.4)	0.5 (-4.0, 5.0)	0.8306
BMN 110 pro Woche	26	29.7 (1.33)	3.5 (0.1, 6.9)	-1.4 (-6.2, 3.4)	0.5588
<b>Weiblich – männlich</b>					
BMN 110 qow	-		-	-3.32 (-9.86,3.21)	0.3171
BMN 110 pro Woche	-		-	4.47 (-2.07,11.01)	0.1790
<b>Rasse: Weiß</b>					
Placebo	44	29.4 (1.03)	3.4 (0.8, 6.1)	-	-
BMN 110 qow	35	25.1 (1.16)	5.0 (2.0, 8.0)	1.6 (-2.4, 5.6)	0.4423
BMN 110 pro Woche	36	29.4 (1.14)	5.3 (2.4, 8.3)	1.9 (-2.0, 5.9)	0.3345
<b>Rasse: Nicht weiß</b>					
Placebo	15	31.8 (1.77)	4.4 (-0.1, 9.0)	-	-
BMN 110 qow	24	30.1 (1.40)	0.6 (-3.0, 4.2)	-3.9 (-9.6, 1.9)	0.1885
BMN 110 pro Woche	22	29.8 (1.46)	3.8 (0.1, 7.6)	-0.6 (-6.5, 5.3)	0.8418

<b>MOR-004</b>	<b>N</b>	<b>Ausgangswert im 3MSC (Stufen/min) LSM (± SE)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (Stufen/min) LSM (95% KI)</b>	<b>Absolute Änderung von Baseline bis Woche 24 (Stufen/min) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t- Test) BMN 110 vs. Placebo</b>
<b>Nicht weiß – weiß</b>					
BMN 110 qow	-		-	-5.42 (-12.49,1.65)	0.1317
BMN 110 pro Woche	-		-	-2.53 (-9.68,4.61)	0.4846
<b>Region Nordamerika</b>					
Placebo	16	31.4 (1.69)	1.6 (-2.8, 6.1)	-	-
BMN 110 qow	16	26.4 (1.69)	4.4 (-0.1, 8.9)	2.8 (-3.5, 9.1)	0.3827
BMN 110 pro Woche	15	25.0 (1.74)	6.1 (1.5, 10.8)	4.5 (-1.9, 11.0)	0.1691
<b>Region Europa</b>					
Placebo	27	29.3 (1.30)	4.1 (0.7, 7.5)	-	-
BMN 110 qow	21	23.4 (1.47)	3.2 (-0.8, 7.2)	-0.9 (-6.2, 4.3)	0.7338
BMN 110 pro Woche	25	28.2 (1.35)	3.6 (0.0, 7.1)	-0.5 (-5.5, 4.4)	0.8352
<b>Region Andere</b>					
Placebo	16	29.8 (1.69)	5.0 (0.6, 9.5)	-	-
BMN 110 qow	22	31.2 (1.44)	2.2 (-1.5, 6.0)	-2.8 (-8.6, 3.1)	0.3519
BMN 110 pro Woche	18	35.2 (1.59)	5.3 (1.1, 9.5)	0.3 (-5.8, 6.4)	0.9239
<b>Europa - Nordamerika</b>					
BMN 110 qow	-		-	-3.45 (-11.89,4.98)	0.4200
BMN 110 pro Woche	-		-	-4.79 (-13.23,3.65)	0.2639
<b>Andere - Nordamerika</b>					
BMN 110 qow	-		-	-5.44 (-14.14,3.26)	0.2188
BMN 110 pro Woche	-		-	-4.08 (-13.15,4.99)	0.3759

ANCOVA (Änderung von Baseline bis Woche 24) mit den Kovariaten Behandlung, Altersgruppen, Baseline 6MWT-Kategorien und Baseline-3MSC, Subgruppe und Interaktion Behandlung nach Subgruppe.

p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell.

Die Gehstrecke ( $\leq 200$  m,  $> 200$  m) sowie der 3MSCT zu Beginn der Studie, die Altersklassen, die Rasse und die geographische Region hatten keinen Einfluss auf das Ergebnis des 3MSCT (Tabelle 4-65). Die p-Werte der Interaktion schwankten zwischen 0.0187 (BMN 110 qow) und 0.4253 für den Vergleich von BMN 110 2.0 mg/kg qow vs. Placebo und zwischen 0.1790

und 0.6122 für den Vergleich von BMN 110 2.0 mg/kg/Woche vs. Placebo. Diese Subgruppenanalysen haben keine Konsequenzen für die Bewertung des Ausmaßes des Zusatznutzens.

In Tabelle 4-66 sind die Subgruppenanalysen für den Endpunkt „uKS“ dargestellt.

Tabelle 4-66: Subgruppenanalysen für den Endpunkt „uKS“ aus RCT mit dem zu bewertenden Arzneimittel in der ITT-Population (9)

<b>MOR-004</b>	<b>N</b>	<b>uKS Baseline (µg/mg) LSM (± SE)</b>	<b>Prozentuale Änderung von Baseline bis Woche 24 (%) LSM (95% KI)</b>	<b>Prozentuale Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b>≤200 m Baseline 6MWT</b>					
Placebo	23	24.4 (2.91)	1.6 (-7.8, 10.9)	-	-
BMN 110 qow	24	27.6 (2.83)	-38.0 (-46.9, -29.1)	-39.6 (-52.4, -26.7)	<0.0001
BMN 110 pro Woche	23	23.9 (2.90)	-44.6 (-54.2, -34.9)	-46.1 (-59.1, -33.2)	<0.0001
<b>&gt;200 m Baseline 6MWT<sup>1</sup></b>					
Placebo	36	28.6 (2.18)	-7.4 (-15.0, 0.3)	-	-
BMN 110 qow	35	29.3 (2.21)	-31.5 (-38.9, -24.1)	-24.1 (-34.7, -13.5)	<0.0001
BMN 110 pro Woche	35	27.6 (2.21)	-44.6 (-52.2, -37.0)	-37.2 (-47.8, -26.6)	<0.0001
<b>6MWT &gt;200 m - 6MWT ≤200 m</b>					
BMN 110 qow	-	-	-	15.43 (-1.07,31.92)	0.0665
BMN 110 pro Woche	-	-	-	8.90 (-7.64,25.44)	0.2894
<b>Altersgruppe 5-11 Jahre</b>					
Placebo	30	36.9 (2.44)	-0.7 (-10.2, 8.9)	-	-
BMN 110 qow	31	39.2 (2.40)	-24.2 (-32.6, -15.9)	-23.6 (-35.4, -11.8)	0.0001
BMN 110 pro Woche	32	34.1 (2.36)	-33.2 (-41.1, -25.3)	-32.5 (-44.2, -20.9)	<0.0001
<b>Altersgruppe 12-18 Jahre</b>					
Placebo	15	20.4 (3.45)	-5.7 (-17.5, 6.1)	-	-
BMN 110 qow	16	23.8 (3.34)	-40.1 (-51.0, -29.3)	-34.4 (-50.3, -18.6)	<0.0001
BMN 110 pro Woche	16	22.0 (3.34)	-50.7 (-61.9, -39.5)	-45.0 (-60.7, -29.4)	<0.0001
<b>Altersgruppe ≥19 Jahre</b>					
Placebo	14	9.3 (3.57)	-11.4 (-28.1, 5.4)	-	-
BMN 110 qow	12	7.6 (3.85)	-51.9 (-65.9, -37.8)	-40.5 (-58.6, -22.4)	<0.0001
BMN 110 pro Woche	10	11.7 (4.22)	-67.5 (-84.2, -50.7)	-56.1 (-76.2, -36.1)	<0.0001
<b>Alter 12-18 Jahre – Alter 5-11 Jahre</b>					
BMN 110 qow	-	-	-	-10.86 (-30.51,8.79)	0.2763
BMN 110 pro Woche	-	-	-	-12.52 (-31.81,6.77)	0.2017

<b>MOR-004</b>	<b>N</b>	<b>uKS Baseline (µg/mg) LSM (± SE)</b>	<b>Prozentuale Änderung von Baseline bis Woche 24 (%) LSM (95% KI)</b>	<b>Prozentuale Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b>Alter ≥19 – Alter 5-11 Jahre</b>					
BMN 110 qow	-		-	-16.92 (-39.41,5.56)	0.1377
BMN 110 pro Woche	-		-	-23.61 (-47.73,0.51)	0.0549
<b>Weibliche Patienten</b>					
Placebo	32	23.8 (3.05)	-5.5 (-14.5, 3.4)	-	
BMN 110 qow	25	27.9 (3.45)	-40.7 (-49.9, -31.6)	-35.2 (-47.8, -22.6)	<0.0001
BMN 110 pro Woche	32	25.5 (3.05)	-43.4 (-52.5, -34.3)	-37.9 (-49.8, -25.9)	<0.0001
<b>Männliche Patienten</b>					
Placebo	27	28.9 (3.32)	-2.5 (-11.3, 6.2)	-	
BMN 110 qow	34	29.1 (2.96)	-30.1 (-37.9, -22.3)	-27.5 (-39.2, -15.8)	<0.0001
BMN 110 pro Woche	26	28.7 (3.38)	-44.4 (-53.3, -35.5)	-41.8 (-54.3, -29.4)	<0.0001
<b>Weiblich – männlich</b>					
BMN 110 qow	-		-	-5.97 (-22.53,10.58)	0.4770
BMN 110 pro Woche	-		-	3.52 (-13.10,20.15)	0.6759
<b>Rasse: Weiß</b>					
Placebo	44	25.3 (2.60)	-2.0 (-9.4, 5.3)	-	
BMN 110 qow	35	28.7 (2.91)	-34.6 (-42.3, -26.8)	-32.6 (-43.4, -21.8)	<0.0001
BMN 110 pro Woche	36	24.5 (2.87)	-45.1 (-52.8, -37.5)	-43.1 (-53.7, -32.6)	<0.0001
<b>Rasse: Nicht weiß</b>					
Placebo	15	28.6 (4.45)	-10.4 (-26.6, 5.8)	-	
BMN 110 qow	24	28.5 (3.52)	-34.6 (-44.0, -25.2)	-24.2 (-41.8, -6.6)	0.0080
BMN 110 pro Woche	22	30.9 (3.67)	-41.7 (-53.2, -30.3)	-31.3 (-48.3, -14.3)	0.0005
<b>Nicht weiß – weiß</b>					
BMN 110 qow	-		-	5.03 (-15.13,25.20)	0.6181
BMN 110 pro Woche	-		-	10.37 (-9.19,29.93)	0.2935
<b>Region Nordamerika</b>					
Placebo	16	26.3 (4.31)	5.8 (-7.6, 19.2)	-	
BMN 110 qow	16	23.1 (4.30)	-32.5 (-43.9, -21.1)	-38.2 (-56.4, -20.1)	<0.0001
BMN 110 pro Woche	15	29.9 (4.44)	-40.5 (-53.5, -27.5)	-46.3 (-65.4, -27.1)	<0.0001

<b>MOR-004</b>	<b>N</b>	<b>uKS Baseline (µg/mg) LSM (± SE)</b>	<b>Prozentuale Änderung von Baseline bis Woche 24 (%) LSM (95% KI)</b>	<b>Prozentuale Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b>Region Europa</b>					
Placebo	27	24.6 (3.31)	-4.7 (-14.5, 5.2)	-	
BMN 110 qow	21	27.4 (3.76)	-33.0 (-42.9, -23.1)	-28.4 (-42.0, -14.7)	<0.0001
BMN 110 pro Woche	25	23.9 (3.44)	-50.1 (-59.1, -41.0)	-45.4 (-58.8, -32.0)	<0.0001
<b>Region Andere</b>					
Placebo	16	28.7 (4.30)	-13.2 (-24.9, -1.4)	-	
BMN 110 qow	22	33.7 (3.67)	-37.7 (-47.3, -28.0)	-24.5 (-39.5, -9.4)	0.0016
BMN 110 pro Woche	18	28.6 (4.06)	-38.0 (-48.9, -27.1)	-24.8 (-40.4, -9.1)	0.0021
<b>Europa -Nordamerika</b>					
BMN 110 qow	-	-	-	10.54 (-12.82,33.91)	0.3687
BMN 110 pro Woche	-	-	-	2.01 (-23.02,27.04)	0.8712
<b>Andere - Nordamerika</b>					
BMN 110 qow	-	-	-	16.43 (-6.54,39.39)	0.1584
BMN 110 pro Woche	-	-	-	22.13 (-1.77,46.02)	0.0690

ANCOVA (Änderung vom Ausgangswert) mit den Kovariaten Behandlung, Altersgruppen, 6MWT-Kategorien und normalisiertes uKS (Baseline), Subgruppe und Interaktion Behandlung nach Subgruppe  
p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell.

Die Ergebnisse der Subgruppenanalysen zum uKS waren dem Ergebnis in der Gesamtpopulation vergleichbar. Die p-Werte der Interaktionstests schwankten zwischen 0.0665 bis 0.6181 für BMN 110 2.0 mg/kg/qow vs. Placebo und 0.0549 bis 0.8712 für BMN 110 2.0 mg/kg einmal wöchentlich. Die größten Unterschiede hinsichtlich der Interaktion zwischen Behandlung und Subgruppe zeigten sich zwischen den Altersgruppen  $\geq 19$  Jahre vs. 5-11 Jahre ( $p=0.0549$ ). Wie erwartet unterschieden sich die uKS-Werte zu Beginn der Studie zwischen den Altersgruppen und waren bei den jüngeren Patienten höher als bei älteren Patienten. Bei niedrigeren uKS Ausgangswerten war bereits eine geringe absolute Änderung bei den älteren Patienten ( $\geq 19$  Jahre) mit einer relativ hohen prozentualen Änderung verbunden. Ausgehend von höheren uKS-Werten führte bei den jüngeren Patienten eine relativ deutliche absolute Änderung zu einer niedrigeren prozentualen Änderung verglichen mit den älteren Patienten. Aus den Subgruppenanalysen zum uKS ergeben sich keine Konsequenzen für die Bewertung des Ausmaßes des Zusatznutzens von BMN 110.

In nachfolgender Tabelle 4-67 sind die Subgruppenanalysen für den Endpunkt „MVV“ dargestellt.

Tabelle 4-67: Subgruppenanalysen für den Endpunkt „MVV“ aus RCT mit dem zu bewertenden Arzneimittel in der ITT-Population (9)

<b>MOR-004</b>	<b>N</b>	<b>MVV Baseline (l/min) LSM (± SE)</b>	<b>Änderung von Baseline bis Woche 24 (l/min) LSM (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b>≤200 m Baseline 6MWT</b>					
Placebo	23	31.8 (4.21)	4.8 (-8.7, 18.4)	-	-
BMN 110 qow	24	28.5 (3.93)	6.1 (-8.4, 20.6)	1.2 (-18.7, 21.2)	0.9024
BMN 110 pro Woche	23	27.1 (4.03)	14.5 (0.5, 28.5)	9.7 (-9.6, 29.0)	0.3223
<b>&gt;200 m Baseline 6 MWT</b>					
Placebo	36	34.4 (3.46)	3.6 (-8.8, 16.0)	-	-
BMN 110 qow	35	34.6 (3.44)	8.4 (-3.4, 20.2)	4.8 (-13.9, 23.4)	0.6032
BMN 110 pro Woche	35	28.7 (3.45)	14.3 (-0.1, 28.7)	10.7 (-5.7, 27.1)	0.1965
<b>6MWT &gt;200 m - 6MWT ≤200 m</b>					
BMN 110 qow	-	-	-	3.56 (-24.59,31.71)	0.7995
BMN 110 pro Woche	-	-	-	1.03 (-24.88,26.94)	0.9372
<b>Altersgruppe 5-11 Jahre</b>					
Placebo	30	24.5 (3.90)	-2.4 (-17.0, 12.2)	-	-
BMN 110 qow	31	27.0 (3.62)	12.3 (0.1, 24.5)	14.7 (-6.0, 35.5)	0.1560
BMN 110 pro Woche	32	24.8 (3.60)	23.4 (6.4, 40.4)	25.8 (7.4, 44.3)	0.0073
<b>Altersgruppe 12-18 Jahre</b>					
Placebo	15	37.9 (5.16)	12.9 (-3.1, 28.9)	-	-
BMN 110 qow	16	32.7 (5.00)	-0.2 (-15.7, 15.2)	-13.1 (-35.2, 9.0)	0.2426
BMN 110 pro Woche	16	29.8 (4.99)	6.7 (-8.8, 22.2)	-6.2 (-28.5, 16.1)	0.5828
<b>Altersgruppe ≥19 Jahre</b>					
Placebo	14	49.9 (5.48)	8.3 (-9.3, 26.0)	-	-
BMN 110 qow	12	44.5 (5.77)	5.2 (-13.0, 23.5)	-3.1 (-27.5, 21.3)	0.8036
BMN 110 pro Woche	10	32.2 (6.35)	-2.3 (-23.8, 19.3)	-10.6 (-38.7, 17.5)	0.4553
<b>Alter 12-18 Jahre – Alter 5-11 Jahre</b>					
BMN 110 qow	-	-	-	-27.84 (-57.92,2.24)	0.0692
BMN 110 pro Woche	-	-	-	-32.03 (-61.02,-3.03)	0.0308

<b>MOR-004</b>	<b>N</b>	<b>MVV Baseline (l/min) LSM (± SE)</b>	<b>Änderung von Baseline bis Woche 24 (l/min) LSM (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b>Alter ≥19 – Alter 5-11 Jahre</b>					
BMN 110 qow	-		-	-17.79 (-49.72,14.15)	0.2704
BMN 110 pro Woche	-		-	-36.43 (-73.13,0.28)	0.0517
<b>Weibliche Patienten</b>					
Placebo	32	30.6 (3.80)	4.0 (-10.2, 18.3)	-	
BMN 110 qow	25	32.3 (4.21)	7.3 (-6.0, 20.7)	3.3 (-15.8, 22.4)	0.7321
BMN 110 pro Woche	32	29.9 (3.74)	13.3 (-0.3, 27.0)	9.3 (-7.9, 26.5)	0.2846
<b>Männliche Patienten</b>					
Placebo	27	37.9 (4.12)	4.0 (-9.1, 17.1)	-	
BMN 110 qow	34	32.0 (3.63)	7.6 (-3.6, 18.8)	3.6 (-13.5, 20.7)	0.6765
BMN 110 pro Woche	26	24.4 (4.15)	15.7 (3.0, 28.3)	11.7 (-6.8, 30.1)	0.2128
<b>Weiblich – männlich</b>					
BMN 110 qow	-		-	-0.76 (-25.80,24.28)	0.9521
BMN 110 pro Woche	-		-	-3.53 (-29.58,22.52)	0.7881
<b>Rasse: Weiß</b>					
Placebo	44	32.9 (3.22)	3.2 (-8.4, 14.8)	-	
BMN 110 qow	35	33.6 (3.61)	6.9 (-4.3, 18.0)	3.6 (-13.2, 20.5)	0.6652
BMN 110 pro Woche	36	28.5 (3.55)	17.0 (3.1, 30.8)	13.8 (-3.2, 30.7)	0.1087
<b>Rasse: Nicht weiß</b>					
Placebo	15	36.9 (5.84)	6.5 (-11.6, 24.6)	-	
BMN 110 qow	24	30.0 (4.32)	8.4 (-4.6, 21.4)	2.0 (-20.3, 24.2)	0.8609
BMN 110 pro Woche	22	25.8 (4.53)	10.1 (-5.2, 25.3)	3.6 (-19.4, 26.6)	0.7551
<b>Nicht weiß – weiß</b>					
BMN 110 qow	-		-	-1.99 (-30.20,26.22)	0.8884
BMN 110 pro Woche	-		-	-10.70 (-42.04,20.64)	0.4916
<b>Region Nordamerika</b>					
Placebo	16	32.1 (5.34)	3.4 (-12.5, 19.2)	-	
BMN 110 qow	16	34.8 (5.32)	11.6 (-4.0, 27.3)	8.3 (-14.1, 30.6)	0.4659
BMN 110 pro Woche	15	29.5 (5.56)	29.4 (-7.7, 66.4)	26.0 (-15.2, 67.2)	0.1872

<b>MOR-004</b>	<b>N</b>	<b>MVV Baseline (l/min) LSM (± SE)</b>	<b>Änderung von Baseline bis Woche 24 (l/min) LSM (95% KI)</b>	<b>Änderung von Baseline bis Woche 24 (%) BMN 110 vs. Placebo LSM (95% KI)</b>	<b>p-Wert (t-Test) BMN 110 vs. Placebo</b>
<b>Region Europa</b>					
Placebo	27	36.2 (4.14)	1.3 (-15.0, 17.6)	-	
BMN 110 qow	21	29.6 (4.67)	3.0 (-12.8, 18.7)	1.6 (-24.0, 27.3)	0.8941
BMN 110 pro Woche	25	26.0 (4.27)	4.7 (-8.1, 17.6)	3.4 (-16.3, 23.2)	0.7295
<b>Region Andere</b>					
Placebo	16	32.0 (5.86)	9.4 (-6.5, 25.3)	-	
BMN 110 qow	22	32.5 (4.56)	8.9 (-5.3, 23.1)	0.5 (-21.9, 20.9)	0.9621
BMN 110 pro Woche	18	27.9 (5.06)	15.1 (-4.6, 34.8)	5.7 (-17.8, 29.2)	0.6298
<b>Europa -Nordamerika</b>					
BMN 110 qow	-	-	-	-5.97 (-42.40,30.46)	0.7399
BMN 110 pro Woche	-	-	-	-21.01 (-72.19,30.17)	0.3820
<b>Andere - Nordamerika</b>					
BMN 110 qow	-	-	-	-8.32 (-39.21,22.58)	0.5955
BMN 110 pro Woche	-	-	-	-19.30 (-73.90,35.30)	0.4495

ANCOVA der prozentualen Änderung vom Ausgangswert unter Berücksichtigung der Kovariaten Behandlung, Altersgruppen, 6MWT-Kategorie, MVV (Baseline), Subgruppe und Interaktion Behandlung nach Subgruppe. p-Wert mittels t-Test vs. Placebo im ANCOVA-Modell.

Der Behandlungseffekt war in den Subgruppen, mit Ausnahme der Altersklassen, vergleichbar mit der Gesamtanalyse, unabhängig von der 6MWT-Kategorie zu Beginn der Studie, des Geschlechts, der Rasse und der geographischen Region. Die p-Werte der Interaktionstests vs. Placebo schwankten zwischen 0.0692 bis 0.9521 für BMN 110 2.0 mg/kg/qow und 0.0308 bis 0.9372 für BMN 110 2.0 mg/kg einmal pro Woche. Die größten Unterschiede hinsichtlich einer Interaktion zwischen Behandlung und Subgruppe zeigte sich beim Vergleich der Altersgruppen 12-18 Jahre vs. 5-11 Jahre (p=0.0308) und ≥ 19 Jahre vs. 5-11 Jahre (p=0.0517). Patienten im Alter zwischen fünf und 11 Jahren sprachen ausgehend von niedrigeren Werten zu Beginn der Studie (vs. älteren Patienten) offensichtlich am besten auf die Behandlung mit BMN 110 einmal pro Woche an, und zeigten sogar für diesen Endpunkt schon nach 24 Wochen eine signifikante Verbesserung. Für die Gesamtpopulation zeigte sich innerhalb der 24-wöchigen Studiendauer erwartungsgemäß noch keine signifikante Veränderung von respiratorischen Parametern.

#### 4.3.1.3.11 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.*

Die nachfolgenden Tabellen geben eine Übersicht der Ergebnisse der randomisierten, Placebo-kontrollierten Doppelblindstudie mit dem gegenständlichen Arzneimittel.

In einer klassischen Responderanalyse unter - Anwendung von unterschiedlichen Responsekriterien - zeigte sich für Gehstreckenänderungen von über 60 m gegenüber Baseline eine signifikante Überlegenheit von BMN 110 (einmal wöchentlich) gegenüber Placebo (s. Abbildung 4). Gehstreckenänderungen von mehr als 60 m sind für die untersuchte Patientenpopulation mit Sicherheit als Gehstrecken einzustufen, die über klinisch relevanten Schwellen liegen.

In einer ergänzenden quantitativen Auswertung der kumulativen Verteilungsfunktionen der Gehstreckenveränderungen konnte eine signifikante Überlegenheit von BMN 110 (einmal wöchentlich) gegenüber Placebo für die gesamte Population und für alle in der Studie erfassten Gehstreckenänderungen demonstriert werden (Wilcoxon-Mann-Whitney Effektgröße, Tabelle 4-25). Die signifikante Überlegenheit von BMN 110 (einmal wöchentlich) war auch für positive Gehstreckenveränderungen mit Responsekriterien von  $\geq 14$  m bis mehr als 95 m unter Verwendung von approximierten relativen Risiken nachweisbar. Diese Responsekriterien beinhalten mit Sicherheit klinisch relevante Veränderungen. Für die Subgruppe der weißen Patienten waren die Effektunterschiede auf Basis der relativen Risiken noch größer. Die Konfidenzintervalle der relativen Risiken lagen in dem Vergleich von Placebo und BMN 110 (einmal wöchentlich) für alle Responsekriterien unter einem Wert von 0.9 (Tabelle 4-68). Auf Basis der therapierelevanten Abschwächung schwerwiegender Krankheitssymptome gegenüber der Vergleichsintervention ist für Elosulfase alfa ein beträchtlicher Zusatznutzen zu sehen.

Tabelle 4-68: Approximierte relative Risiken für den Vergleich der individuellen Gehstreckenunterschiede in den Behandlungsgruppen BMN 110 (einmal wöchentlich) versus Placebo

		<b>RR derived from OR</b> OR (= $\beta$ ): 2.0368 (95% CI: 1.0766 - 4.1090)			<b>Invertiertes Relatives Risiko</b> OR (= $\beta$ ): 0.4910 (95% CI: 0.2434 - 0.9289)		
CER	Interpolierte Gehstreckenveränderung (Meter)	RR (CER)	RR (CER) unteres Limit	RR (CER) oberes Limit	RR	RR unteres Limit	RR oberes Limit
<b>Gesamtpopulation</b>							
0.0500	≥ 95.1	1.9364	1.0725	3.5562	0.5164	0.2812	0.9324
0.1000	≥ 73	1.8455	1.0684	3.1345	0.5419	0.3190	0.9360
0.2000	≥ 57.5	1.6870	1.0604	2.5336	0.5928	0.3947	0.9431
0.3000	≥ 33	1.5536	1.0524	2.1260	0.6437	0.4704	0.9502
0.4000	≥ 14.1	1.4397	1.0446	1.8314	0.6946	0.5460	0.9573
<b>Weißer Population</b>							
0.0500	≥ 95.1	2.4445	1.2047	5.1621	0.4091	0.1937	0.8301
0.1000	≥ 75.6	2.2718	1.1918	4.2345	0.4402	0.2362	0.8390
0.2000	≥ 57.7	1.9905	1.1670	3.1150	0.5024	0.3210	0.8569
0.3000	≥ 33	1.7712	1.1431	2.4637	0.5646	0.4059	0.8748
0.4000	≥ 19	1.5954	1.1202	2.0376	0.6268	0.4908	0.8927

Für den primären Endpunkt (mittlere absolute Änderung von Baseline bis Woche 24, Abbildung 2) der Studie MOR-004 zeigte sich eine signifikante Überlegenheit von BMN 110 2 mg/kg/KG einmal wöchentlich gegenüber der Vergleichstherapie. Dieser Behandlungseffekt wurde in einer Zwischenauswertung einer Langzeit-Verlängerungsstudie (MOR-005; Abschnitt 4.3.2.3.3) bestätigt.

Auch für nahezu alle anderen sekundären und tertiären Endpunkte zeigten sich numerische Verbesserungen zugunsten von BMN 110 2 mg/kg/KG (einmal wöchentlich) (Tabelle 4-69, Abbildung 18). Die Ergebnisse einer Zwischenanalyse einer Verlängerungsstudie sind in Abschnitt 4.3.2.3.3 zusammengefasst.

Tabelle 4-69: Zusammenfassung der Ergebnisse der Wirksamkeitsanalysen aus RCT mit dem zu bewertenden Arzneimittel (9)

Parameter	Gruppe	Änderung bis Woche 24 vs. Placebo		p-Wert
		LSM (95% KI)		
6MWT (m)	BMN 110 qow (n=59)	0.5	(-17.8, 18.9)	0.9542
	BMN 110 pro Woche (n=58)	22.5	(4.0, 40.9)	0.0174
3MSCT (Stufen/min)	BMN 110 qow (n=59)	-0.5	(-3.7, 2.8)	0.7783
	BMN 110 pro Woche (n=58)	1.1	(-2.1, 4.4)	0.4935
uKS (%)	BMN 110 qow (n=59)	-30.2	(-38.5, -22.0)	<0.0001
	BMN 110 pro Woche (n=58)	-40.7	(-49.0, -32.4)	<0.0001
MVV (%)	BMN 110 qow (n=59)	3.4	(-9.9, 16.6)	0.6111
	BMN 110 pro Woche (n=58)	10.3	(-1.8, 22.4)	0.0943
FVC (%)	BMN 110 qow (n=59)	3.0	(-2.4, 8.3)	0.2751
	BMN 110 pro Woche (n=58)	3.3	(-3.1, 9.6)	0.3041
FEV1 <sup>1</sup> (%)	BMN 110 qow (n=59)	0.2	(-7.4, 7.9)	0.9546
	BMN 110 pro Woche (n=58)	1.8	(-5.5, 9.2)	0.6129
FIVC <sup>1</sup> (%)	BMN 110 qow (n=59)	-20.9	(-63.6, 21.9)	0.3344
	BMN 110 pro Woche (n=58)	-2.6	(-65.7, 60.5)	0.9303
FET <sup>1</sup> (%)	BMN 110 qow (n=59)	-3.7	(-59.8, 52.4)	0.8962
	BMN 110 pro Woche (n=58)	42.5	(-14.9, 99.9)	0.1457
MPS HAQ Selbstversorgung <sup>1</sup>	BMN 110 qow (n=59)	-0.1	(-0.5, 0.3)	0.7242
	BMN 110 pro Woche (n=58)	0.1	(-0.3, 0.5)	0.7367
MPS HAQ Mobilität <sup>1</sup>	BMN 110 qow (n=59)	-0.3	(-0.8, 0.2)	0.2611
	BMN 110 pro Woche (n=58)	-0.3	(-0.8, 0.3)	0.3355
MPS HAQ Hilfe durch Betreuer <sup>1</sup>	BMN 110 qow (n=59)	-0.4	(-2.4, 1.6)	0.6794
	BMN 110 pro Woche (n=58)	-0.9	(-2.8, 1.1)	0.3990
Körpergröße im Stehen (z-Score) <sup>1,2</sup>	BMN 110 qow (n=45)	0.1	(-0.1, 0.3)	0.2218
	BMN 110 pro Woche (n=44)	0.1	(-0.0, 0.3)	0.1149
Wachstumsrate (z-Score) <sup>1,2</sup>	BMN 110 qow (n=59)	0.4	(-0.1, 0.9)	0.1384
	BMN 110 pro Woche (n=58)	0.4	(-0.1, 0.9)	0.1032

ANCOVA mit den Kovariaten Altersgruppen und 6MWT-Kategorie sowie des jeweiligen Endpunkts zu Beginn der Studie.

<sup>1</sup> Behandlung als zusätzliche Kovariate.

<sup>2</sup> Männliche Patienten im Alter von ≤18 Jahren und weibliche Patienten im Alter von ≤15 Jahren.

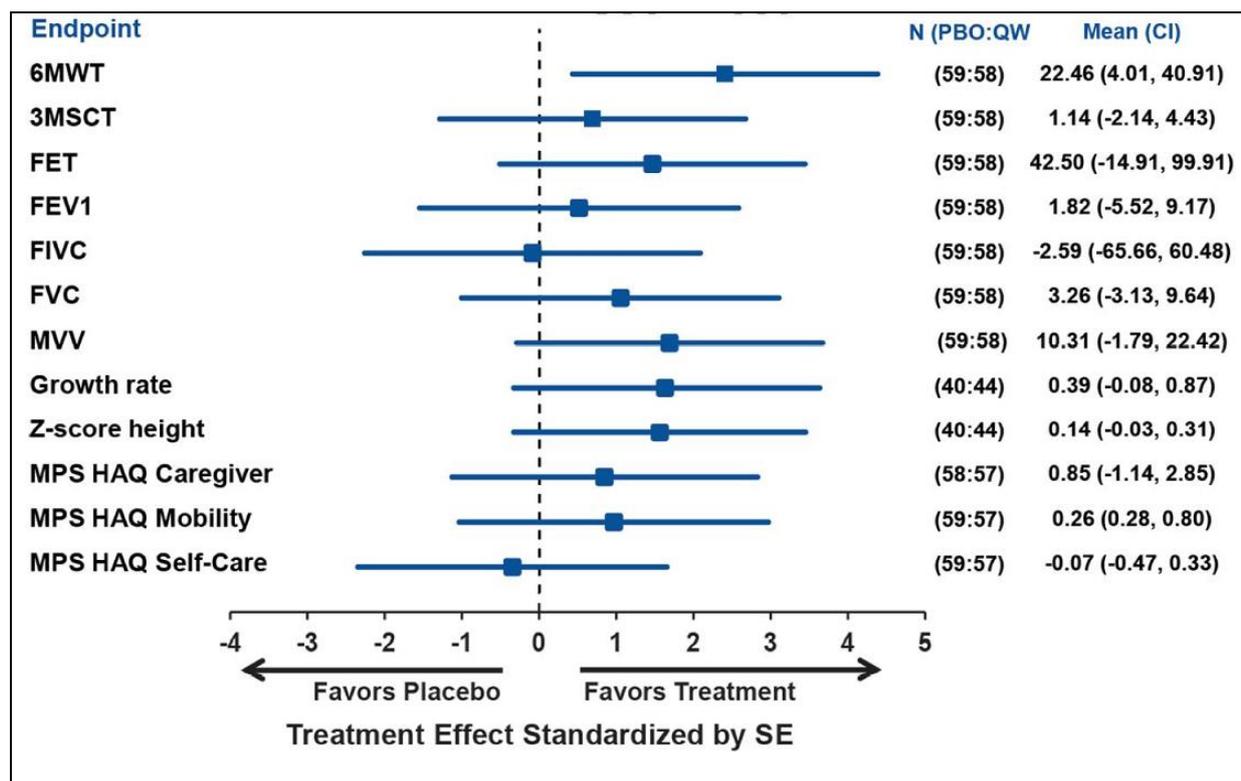


Abbildung 18: Zusammenfassung der Behandlungseffekte von BMN 110 (einmal wöchentlich) gegenüber Placebo für die Wirksamkeitsendpunkte. Analysepopulation: ITT

Die Beurteilung aller in der Studie erhobenen Sicherheitsparameter demonstrierte, dass die intravenöse Infusion von BMN 110 2.0 mg/kg/KG (einmal wöchentlich) im Vergleich zu einer Placebointervention gut toleriert wurde (Tabelle 4-70). Keiner der Patienten brach die Studie auf Grund von unerwünschten Ereignissen ab. Die häufigsten Therapie-assoziierten UEs waren Erbrechen, Übelkeit, Fieber, und Kopfschmerz, die auch bei Infusion von anderen Enzyersatztherapien beobachtet wurden. Schwere unerwünschte Ereignisse waren entweder infusionsbedingt oder krankheitsbedingt, und konnten symptomatisch behandelt werden. Hinsichtlich der gesamten mit der Behandlung in Zusammenhang stehenden Überempfindlichkeitsreaktionen einschließlich anaphylaktische Reaktionen und Angioödeme ergab sich zwischen Placebo und BMN 110 einmal pro Woche kein signifikanter Unterschied. Bei der Mehrzahl der Patienten trat unabhängig von der Gruppenzugehörigkeit mindestens eine infusionsbedingte Reaktion (IAR) auf. Nur jeweils drei der gesamten BMN 110-Infusionen (qow bzw. einmal wöchentlich) wurden abgebrochen und keiner der Patienten brach die Studie aufgrund von UEs einschließlich IARs ab. Alle Patienten, die eine IAR entwickelt hatten, erhielten und tolerierten spätere Infusionen.

Das Sicherheitsprofil von BMN 110 entsprach demjenigen anderer ERTs. Mit Ausnahme von IARs ergab sich kein kausaler Zusammenhang zwischen der Behandlung und klinisch bedeutungsvollen Änderungen der Laborparameter und Vitalzeichen.

Tabelle 4-70: Zusammenfassung der Ergebnisse der Sicherheitsanalysen aus RCT mit dem zu bewertenden Arzneimittel (9)

UE	Placebo (N=59) n (%)	BMN 110 2.0 mg/kg (N=59) n (%)	BMN 110 2.0 mg/kg/Woche (N=58) n (%)
Gesamte UEs	57 (96.6%)	59 (100.0%)	56 (96.6%)
Therapie-assoziierte gesamte UEs	36 (61.0%)	42 (71.2%)	42 (72.4%)
SUEs	2 (3.4%)	4 (6.8%)	9 (15.5%)
Therapie-assoziierte SUEs	0 (0.0%)	1 (1.7%)	2 (3.4%)
Gesamte Überempfindlichkeitsreaktionen <sup>1</sup>	7 (11.9%)	16 (27.1%)	12 (20.7%)
Gesamte anaphylaktische Reaktionen SMQ <sup>1</sup>	1 (1.7%)	2 (3.4%)	3 (5.2%)
Angioödem SMQ <sup>1</sup> (gesamt)	7 (11.9%)	14 (23.7%)	10 (17.2%)
Mit den Prüfpräparaten assoziierte Überempfindlichkeitsreaktionen <sup>1</sup>	2 (3.4%)	9 (15.3%)	7 (12.1%)
Mit der Behandlung in Zusammenhang stehende anaphylaktische Reaktionen SMQ <sup>1</sup>	0 (0.0%)	2 (3.4%)	3 (5.2%)
Mit der Behandlung in Zusammenhang stehende Angioödeme SMQ <sup>1</sup>	2 (3.4%)	7 (11.9%)	5 (8.6%)
Gesamte IARs <sup>2</sup>	54 (91.5%)	56 (94.9%)	52 (89.7%)
Gesamte IARs <sup>2</sup> während der Infusion	48 (81.4%)	53 (89.8%)	52 (89.7%)
Gesamte schwerwiegende IARs <sup>1</sup> während der Infusion	0 (0.0%)	2 (3.4%)	2 (3.4%)

MedDRA Version 15.0

Inzidenz: Patienten, die mehr als ein UE innerhalb eines MedDRA PT entwickelten, wurden nur einmal gezählt.

<sup>1</sup> Überempfindlichkeitsreaktionen wurden mittels einer standardisierten MedDRA-Abfrage (SMQ) unter Einsatz der Algorithmen „Anaphylaktische Reaktionen“ und „Angioödem“ identifiziert.

<sup>2</sup> IARs (MedDRA PT) waren als mit der Verabreichung der Prüfpräparate assoziiert eingestuft wenn sie nach Beginn der Infusion oder innerhalb eines Tages nach Ende der Infusion auftraten (unabhängig davon, ob ein Zusammenhang mit den Prüfpräparaten bestand oder nicht).

## 4.3.2 Weitere Unterlagen

### 4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

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

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

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

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

Nicht zutreffend

**4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche**

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

Nicht zutreffend

**4.3.2.1.3 Ergebnisse aus indirekten Vergleichen**

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

Tabelle 4-71: 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 4-72: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

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

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend

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

Tabelle 4-73: 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 4-74: 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.

Nicht zutreffend

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

Tabelle 4-75: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

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

Nicht zutreffend

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

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

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

Nicht zutreffend

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

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

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.10.*

Nicht zutreffend

#### **4.3.2.2 Nicht randomisierte vergleichende Studien**

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

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

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

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

Nicht zutreffend

##### **4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien**

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

*Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit*

den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-76: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

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

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend

#### 4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

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

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

Tabelle 4-77: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
MOR-005	

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 4-78: Verzerrungsaspekte für &lt;Endpunkt xxx&gt; – nicht randomisierte vergleichende Studien

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

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

Nicht zutreffend

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

Nicht zutreffend

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

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

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.10.***

Nicht zutreffend

#### **4.3.2.3 Weitere Untersuchungen**

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 0 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 0 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

Tabelle 4-79: Liste der Studien des pharmazeutischen Unternehmers – nicht RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer	Therapiearme
MOR-005	ja	ja	laufend	Geplante Studiendauer von bis zu 240 Wochen	Part 1: BMN 110 2.0 mg/kg pro Woche 2.0 mg/kg jede zweite Woche Part 2: BMN 110 2.0 mg/kg pro Woche

Tabelle 4-80: Studienpool – nicht RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
<b>ggf. Zwischenüberschrift zur Strukturierung des Studienpools</b>						
<b>Nicht RCT</b>						
MOR-005	ja	ja	nein	ja (52)	ja (53)	nein
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.						

Studie	Studienkategorie			verfügbare Quellen <sup>a</sup>		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie <sup>b</sup> (ja/nein)	Studie Dritter (ja/nein)	Studienbericht (ja/nein [Zitat])	Registereintrag <sup>c</sup> (ja/nein [Zitat])	Publikation (ja/nein [Zitat])
b: Studie, für die der Unternehmer Sponsor war.						
c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.						

#### 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.***

Bei der Zulassungsstudie MOR-005 (52) handelt es sich um eine Verlängerungsstudie mit Patienten, die die pivotale Studie MOR-004 abgeschlossen hatten. Ziel der Studie war die Beurteilung der Langzeit-Wirksamkeit und Sicherheit von BMN 110 in Patienten mit MPS IVA. Die Studie basierte auf zwei aufeinander folgenden Abschnitten: In dem doppelblinden Part 1 verblieben diejenigen Patienten, die in MOR-004 mit BMN 110 2 mg/kg/KG jede Woche oder jede zweite Woche behandelt wurden auf ihrem jeweiligen Therapieregime (QW-QW, QOW-QOW). Patienten, die in MOR-004 in den Placeboarm randomisiert waren, wurden in eine der beiden BMN 110 Kohorten randomisiert (BMN 110 2 mg/kg/KG jede Woche oder jede zweite Woche) (PBO-QOW, PBO-QW). In dem offenen Part 2 wurden nach Auswertung der MOR-004 Studiendaten, und auf Empfehlung des Data Monitoring Committee, alle Patienten unter Behandlung mit dem optimalen Dosierungsregime BMN 110 2 mg/kg/KG jede Woche gestellt (Abbildung 19).

An dieser Stelle sollen nur supportive Daten zur Langzeitwirksamkeit einer Zwischenanalyse vom 03 September 2013 zusammengefasst werden. In Bezug auf Details der Methodologie und der Studienpopulation wird daher auf den Studienbericht der Studie MOR-005 (52) und auf die Advisory Committee Briefing Materials (15) verwiesen.

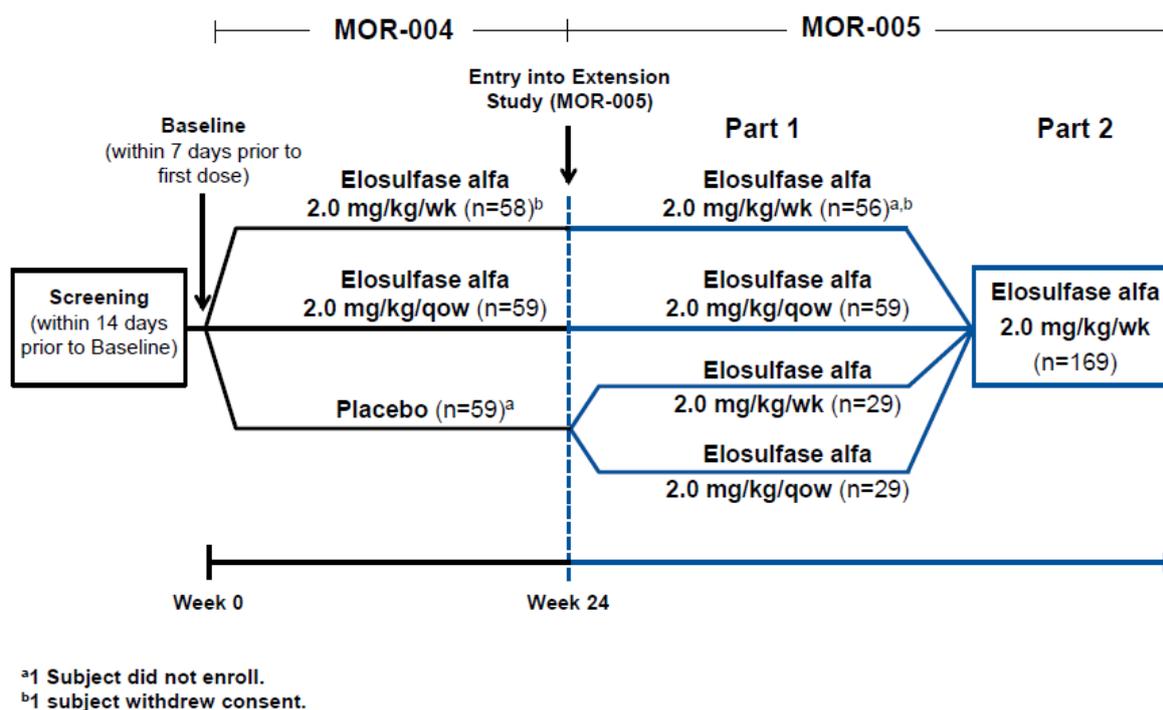


Abbildung 19: Studiendesign der Studien MOR-004 und MOR-005 (52)

#### 4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

##### 4.3.2.3.3.1 Zwischenanalyse Langzeitwirksamkeit – 6MWT– 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 4-81: Operationalisierung von Langzeitwirksamkeit – 6MWT - weitere Untersuchungen

Studie	Operationalisierung
MOR-005	Primärer Endpunktparameter der Studie MOR-004 war die Änderung der Gehstrecke im 6MWT vom Ausgangspunkt in Woche 12 und Woche 24. In der Studie MOR-005 wurde dieser Endpunkt zu Woche 0 (Woche 24 der Studie MOR-004) und zu Woche 12, 24, und 48 erhoben. In die statistische Analyse wurde jeweils der Mittelwert beider Tests pro Erhebungszeitpunkt einbezogen oder eines Einzeltests, wenn nur ein Testergebnis vorlag. Die Änderung der Gehstrecke (in Metern) im 6MWT bis Woche 48 der Studie MOR-005 verglichen mit Baseline wurde mittels ANCOVA (mit Messwiederholung) mit den Faktoren Behandlung, Altersklasse (5-11, 12-18, $\geq 19$ Jahre) und 6MWT-Kategorien bei der Baseline-Visite ( $\leq 200$ m und $> 200$ m) analysiert (primäre Analyseverfahren)

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht zutreffend

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

Descriptive Statistik der Gehstrecken für alle Behandlungskohorten bis zur Woche 72 nach Baselineerhebung (Woche 48 der Studie MOR-005) ist in Tabelle 82 dargestellt. Bei der Interpretation der Daten sollte darauf geachtet werden, dass auf Grund der unterschiedlichen Auswertungspläne in Part 1 und Part 2 nicht für alle Patienten die Werte in Woche 48 erhoben werden konnten. Weiters erhielten mit Beginn des Part 2 alle Patienten der Kohorten PBO-QOW und QOW-QOW die effektive Dosierung von BMN 110 2 mg/kg/KG jede Woche. Der Zeitpunkt des Übergangs auf die effektive (zugelassene) Dosierung war für jeden Patienten unterschiedlich, die meisten Patienten wurden zwischen der Woche 36 und der Woche 72 auf die effektive Dosierung umgestellt.

Tabelle 82: Gehstrecken im 6MWT von Baseline bis Woche 72 (ITT-Population)

Gehstrecke (Meter)	Placebo (n=59)	PBO-QOW (n=29)	PBO-QW (n=29)	QOW-QOW (n=59)	QW-QW (n=58)
Baseline					
n	59	29	29	59	58
Mean (SD)	211.9 (69.88)	219.7 (74.22)	207.2 (64.87)	205.7 (81.19)	203.9 (76.32)
Median	228.9	239.5	217.2	218.0	216.5
25th , 75th Percentile	151.0 , 275.0	167.6 , 285.6	151.0 , 238.7	145.1 , 266.9	149.4 , 259.8
Min, Max	36.2 , 312.2	36.2 , 309.9	93.0 , 312.2	47.1 , 319.6	42.4 , 321.5
Woche 12					
n	59	NA	NA	59	58
Mean (SD)	224.6 (78.50)			219.1 (78.44)	227.6 (76.38)
Median	231.3			232.1	237.1
25th , 75th Percentile	176.5 , 288.4			174.4 , 269.6	192.4 , 292.4
Min , Max	51.5 , 431.5			54.7 , 377.3	48.6 , 350.7
Woche 24 (MOR-005 Woche 0)					
n	59	29	29	58	57
Mean (SD)	225.4 (83.22)	243.5 (93.18)	212.2 (66.30)	220.5 (88.24)	243.3 (83.53)
Median	229.4	252.1	225.4	238.1	251.0
25th , 75th Percentile	159.7 , 286.9	177.0 , 304.5	159.7 , 260.9	148.6 , 285.0	206.4 , 287.5
Min , Max	50.6 , 501.0	74.3 , 501.0	50.6 , 335.0	44.1 , 370.4	52.0 , 399.9
Woche 36 (MOR-005 Woche 12)					

Gehstrecke (Meter)	Placebo (n=59)	PBO-QOW (n=29)	PBO-QW (n=29)	QOW-QOW (n=59)	QW-QW (n=58)
n	NA	28	28	58	54
Mean (SD)		254.7 (97.15)	210.4 (83.43)	227.4 (88.77)	253.0 (80.76)
Median		278.4	225.4	244.9	253.8
25th , 75th Percentile		192.4 , 316.3	162.6 , 262.4	168.0 , 284.2	196.9 , 313.5
Min , Max		58.2 , 462.0	0.0 , 333.8	28.8 , 391.0	43.0 , 405.0
Woche 48 (MOR-005 Woche 24)					
n	NA	14	13	26	27
Mean (SD)		245.2 (141.52)	197.1 (99.67)	203.7 (105.08)	242.0 (95.00)
Median		247.8	212.7	222.4	253.7
25th , 75th Percentile		148.2 , 317.1	181.9 , 244.0	105.0 , 277.1	192.0 , 311.0
Min , Max		0.0 , 522.0	0.0 , 348.0	4.3 , 424.0	0.0 , 386.8
Woche 72 (MOR-005 Woche 48)					
n	NA	28	27	56	54
Mean (SD)		262.0 (121.98)	202.7 (106.50)	234.7 (95.41)	240.4 (103.50)
Median		267.4	225.9	249.7	255.0
25th , 75th Percentile		164.8 , 330.2	159.6 , 266.0	171.9 , 294.8	192.5 , 315.9
Min , Max		63.4 , 616.5	0.0 , 379.0	3.8 , 443.7	0.0 , 422.5

PBO-QOW, Placebo-BMN 110 2.0 mg/kg/KG jede zweite Woche;

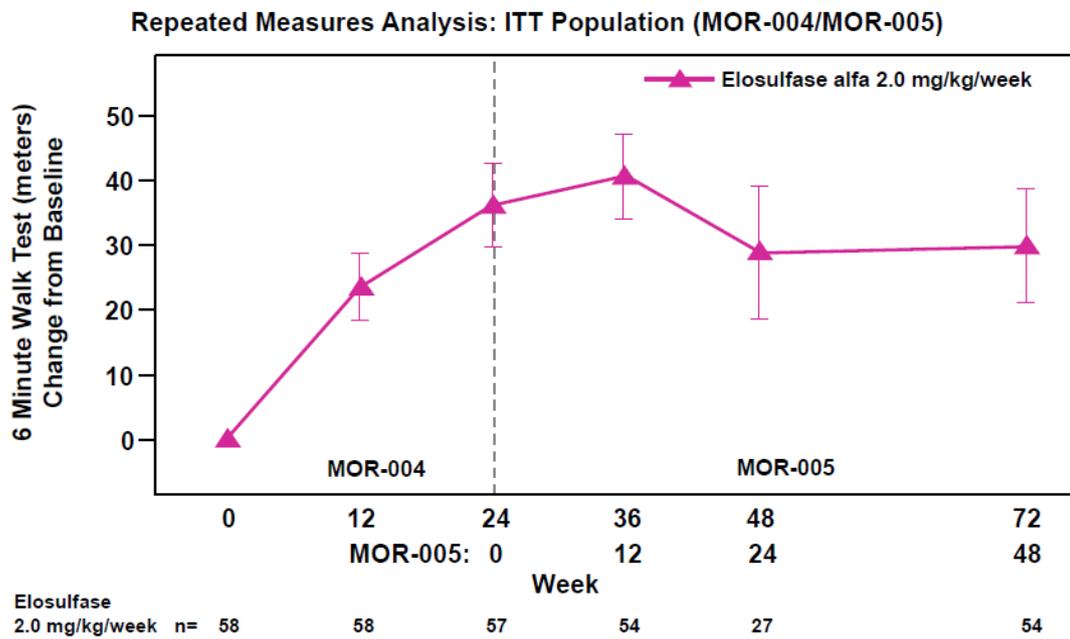
PBO-QW, Placebo-BMN 110 2.0 mg/kg/KG jede Woche;

QOW-QOW, BMN 110-BMN 110 2.0 mg/kg/KG jede zweite Woche;

QW-QW, BMN 110-BMN 110 2.0 mg/kg/KG jede Woche

In Abbildung 20 sind die Gehstreckenänderungen nach der ANCOVA Analyse (mit Messwiederholung) für diejenigen Patienten dargestellt, die seit Beginn der Studie MOR-004 bis zur Woche 72 (Woche 48 der Studie MOR-005) mit BMN 110 2.0 mg/kg/KG jede Woche behandelt wurden (QW-QW). Als primäre Analysepopulation wurde die ITT Population festgelegt. Eine Sensitivitätsanalyse wurde mit der PP Population durchgeführt. Die Ergebnisse zeigten, dass die sich der Behandlungseffekt der Intervention mit BMN 110 2.0 mg/kg/KG jede Woche über einen Zeitraum von 72 Wochen verbesserte (der Behandlungseffekt in der ITT Population war geringer als in der PP Population, da die ITT Population Patienten enthielt, die sich nach 24 Wochen einem chirurgischen Eingriff unterziehen mussten). Vor dem Hintergrund einer sich chronisch verschlechternden Erkrankungssymptomatik ist dieses Ergebnis als klinisch bedeutsam zu werten, und unterstützt die Schlussfolgerungen der Datenauswertung der pivotalen Studie MOR-004.

A.



B.

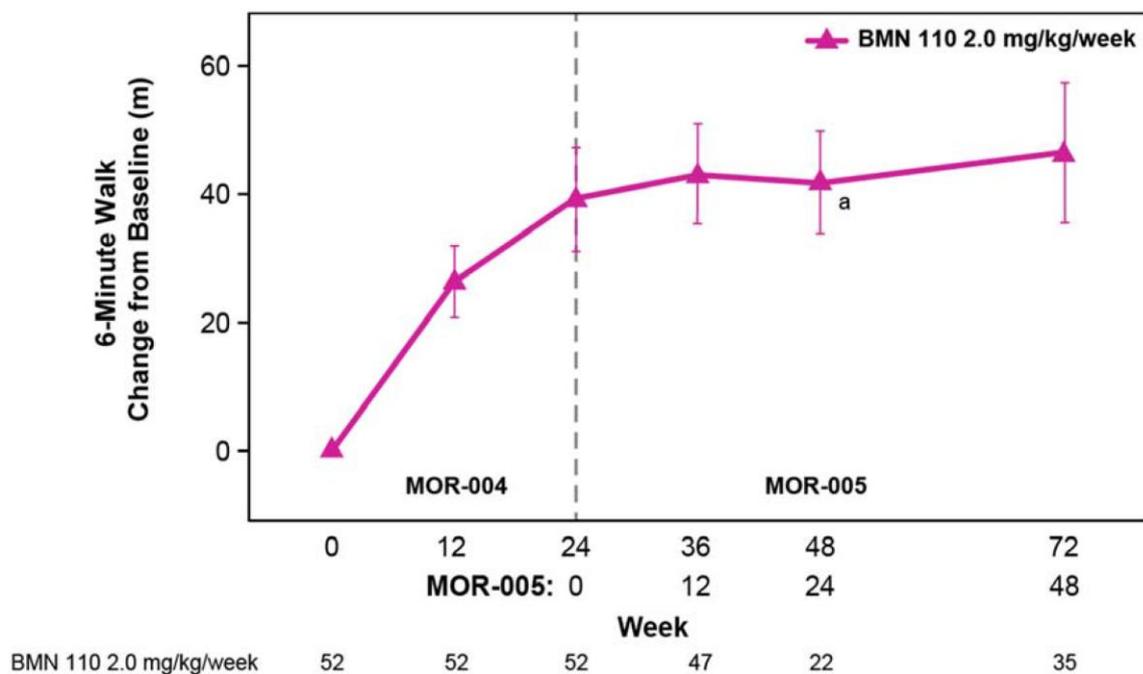


Abbildung 20: Änderung der Gehstrecke im 6MWT über 72 Wochen (Repeated Measures ANCOVA Mittelwert ± SE). Analyse Population: ITT (A); Per Protocol (B)

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

#### 4.3.2.3.3.2 Zwischenanalyse Langzeitwirksamkeit – FVC– 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 4-83: Operationalisierung von Langzeitwirksamkeit – FVC - weitere Untersuchungen

Studie	Operationalisierung
MOR-005	Der Endpunktparameter FVC wurde in der Studie MOR-004 während der Eingangsuntersuchung und in Woche 24 erhoben. In der Studie MOR-005 wurde dieser Endpunkt zu Woche 0 (Woche 24 der Studie MOR-004) und zur Woche 24 und 48 erhoben. Primäre Analyseverfahren war die prozentuelle Veränderung von Baseline mittels ANCOVA (mit Messwiederholung) mit den Faktoren Behandlung, Altersklasse (5-11, 12-18, $\geq 19$ Jahre) und 6MWT-Kategorien bei der Baseline-Visite ( $\leq 200$ m und $> 200$ m) für die ITT Population. Deskriptiv dargestellt wurden Kleinstquadrat-Mittelwerte und Konfidenzintervalle zu jeder Visite.

*Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nichtrandomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.*

Nicht zutreffend

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

Patienten mit MPS IVA leiden häufig an restriktiven und obstruktiven Lungenerkrankungen. Hauptursachen sind laryngeale Verengungen, sowie tracheale und bronchiale Abnormalitäten. Verbesserungen in Lungenfunktionsparametern vor dem Hintergrund einer sich progredient verschlechternden Erkrankung sind daher als höchst patientenrelevant für die hier zu bewertende Patientenpopulation zu betrachten.

Die Ergebnisse für den Endpunkt ‚forcierte Vitalkapazität‘ (FVC) aus einer Zwischenanalyse der Verlängerungsstudie MOR-005 (52) sind bis zur Woche 48 (Woche 72 der pivotalen Studie MOR-004) in Abbildung 21, Abbildung 22, und Abbildung 23 zusammengefasst. Für die Patientengruppe, die seit Beginn der Studie MOR-004 mit BMN 110 2.0 mg/kg/KG jede Woche behandelt wurde (QW-QW), war die mittlere Verbesserungen der forcierten

Vitalkapazität bis zur Woche 72 (Woche 48 der Studie MOR-005) unabhängig von der Auswertungsmethode statistisch signifikant. Diese Ergebnisse unterstützen die Schlussfolgerungen der Datenauswertung der pivotalen Studie MOR-004.

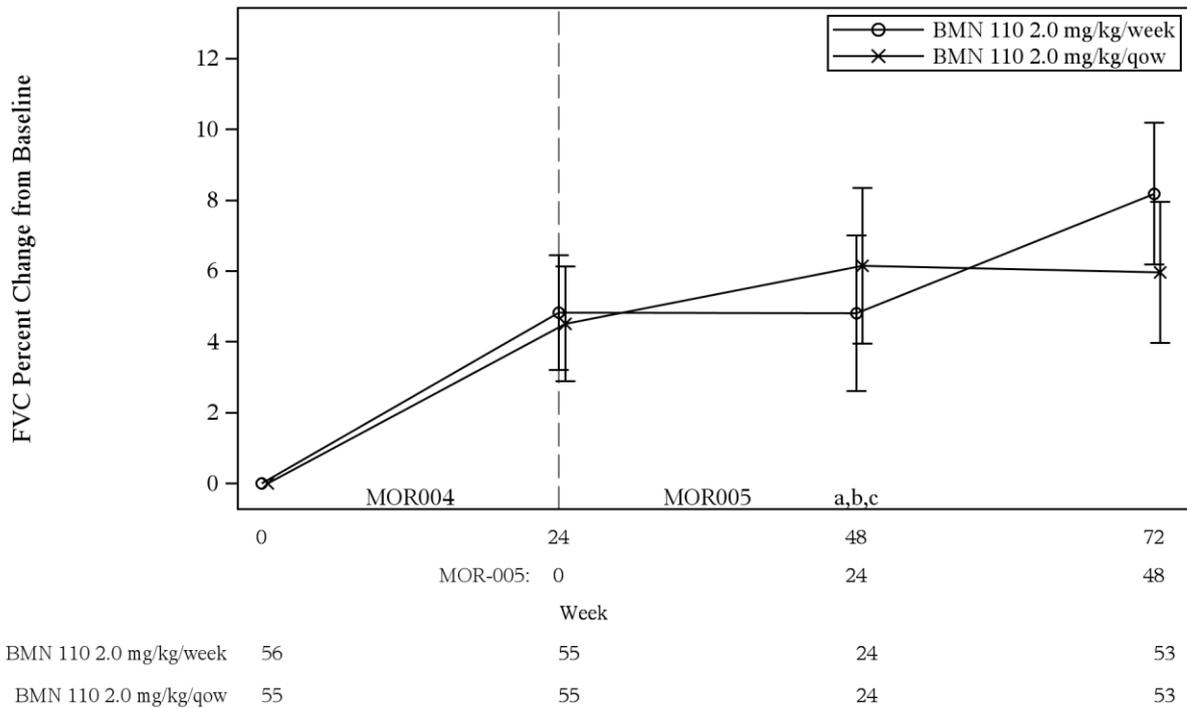


Abbildung 21: ANCOVA mit Messwiederholung – Prozentuale Veränderung des Mittelwertes (least squares mean  $\pm$  SE) der forcierten Vitalkapazität. Analysepopulation ITT

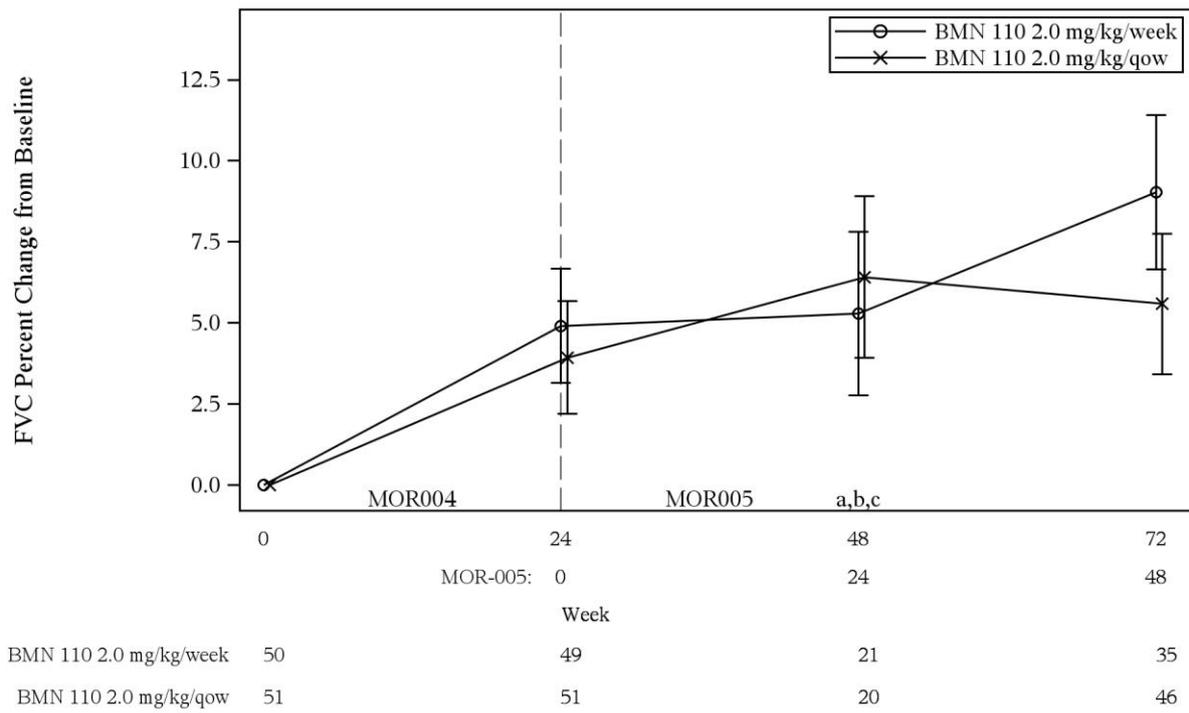
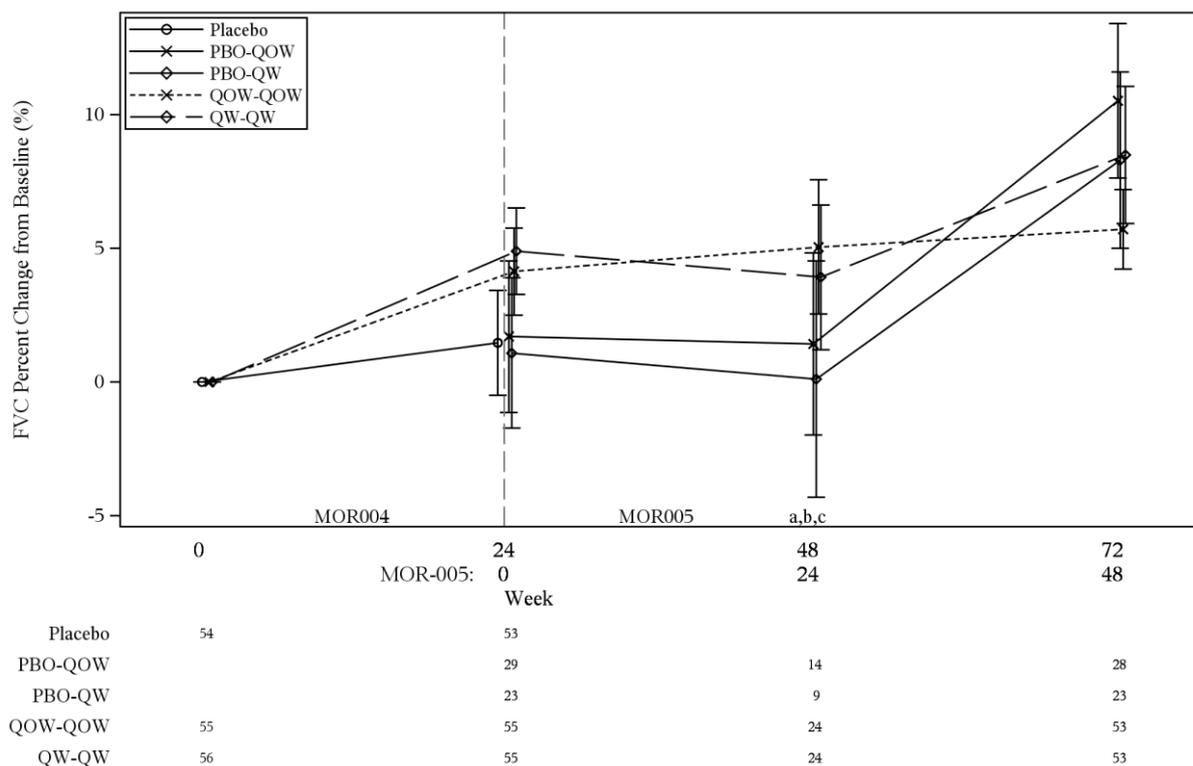


Abbildung 22: ANCOVA mit Messwiederholung – Prozentuale Veränderung des Mittelwertes (least squares mean ± SE) der forcierten Vitalkapaizität. Analysepopulation PP



PBO-QOW, Placebo-BMN 110 2.0 mg/kg/KG jede zweite Woche;  
 PBO-QW, Placebo-BMN 110 2.0 mg/kg/KG jede Woche;  
 QOW-QOW, BMN 110-BMN 110 2.0 mg/kg/KG jede zweite Woche;  
 QW-QW, BMN 110-BMN 110 2.0 mg/kg/KG jede Woche

Abbildung 23: Prozentuale Veränderung des Mittelwertes ( $\pm$  SE) der forcierten Vitalkapazität. Analysepopulation ITT

#### 4.3.2.3.3 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.10.**

Nicht zutreffend

#### 4.3.2.4 Zusammenfassung der Ergebnisse aus weiteren Unterlagen

Der vorliegende Abschnitt soll einen Überblick über die Ergebnisse aus weiteren Unterlagen (Abschnitte 4.3.2.1, 0 und 0) geben. Die Zusammenfassung soll Aussagen zu allen in diesen Abschnitten präsentierten Endpunkten und Subgruppenanalysen enthalten. Dabei sollen, soweit verfügbar, numerische Ergebnisse aus Meta-Analysen einschließlich Konfidenzintervallen dargestellt werden.

Fassen Sie die Ergebnisse aus weiteren Unterlagen zusammen.

Nicht zutreffend

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

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

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

Bei der für die Nutzenbewertung herangezogenen Studie handelt es sich um eine randomisierte, Placebo-kontrollierte Doppelblindstudie der Evidenzstufe 1b. Da es sich bei Morquio IVA um ein sehr seltenes Leiden handelt, ist die Anzahl der in die Studie eingeschlossenen Patienten als hoch einzustufen. Die *a priori* festgelegten Endpunkte wurden adäquat operationalisiert und das Verzerrungspotential wurde auf Studienebene und für die bewertungsrelevanten Endpunkte als niedrig eingestuft, so dass insgesamt von einer hohen Aussagekraft der Nachweise ausgegangen werden kann.

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

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

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

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

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

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

Mit der Zulassung von Vimizim (BMN 110) als Arzneimittel zur Behandlung eines seltenen Leidens („Orphan Drug“) (Verordnung (EG) Nr. 141/2000 des Europäischen Parlaments und des Rates vom 16. Dezember 1999) gilt der Zusatznutzen nach § 35a Abs. 1 Satz 10 SGB V mit der Zulassung als belegt. Das Ausmaß des Zusatznutzens von Vimizin wird anhand der im 5. Kapitel § 5 Abs. 7 Nr. 1 bis 4 der VerfO festgelegten Kriterien qualitativ bewertet (6). Demnach liegt ein beträchtlicher Zusatznutzen vor, wenn eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte deutliche Verbesserung des therapie-relevanten Nutzens im Sinne von § 3 Absatz 1 erreicht wird, insbesondere eine Abschwächung schwerwiegender Symptome, eine moderate Verlängerung der Lebensdauer, eine für die Patientinnen und Patienten spürbare Linderung der Erkrankung, eine relevante Vermeidung schwerwiegender Nebenwirkungen oder eine bedeutsame Vermeidung anderer Nebenwirkungen (6).

Bei der Studie MOR-004 handelt es sich um eine Placebo-kontrollierte Doppelblindstudie der Evidenzstufe 1b, die eine quantitative vergleichende Beurteilung des Ausmaßes des Effektes und eine Quantifizierung des Zusatznutzens nach heutigen wissenschaftlichen Standards erlaubt. Bei der Quantifizierung des Ausmaßes des Zusatznutzens von BMN 110 (Vimizim) ist zu berücksichtigen, dass es sich um eine sehr schwerwiegende, progrediente Erkrankung handelt, für die bislang keine Referenztherapie zugelassen ist und deren Behandlung sich hauptsächlich auf palliative Maßnahmen beschränkt (1,8).

### Wirksamkeit

Bei einer chronischen Erkrankung, die durch fortschreitende Behinderung, häufig bis zur Immobilität und Verlust funktionaler Fähigkeiten charakterisiert ist, ist eine Verbesserung der Gehstrecke patientenrelevant (3). Das CHMP klassifiziert die Verbesserung der Gehstrecke unter Behandlung mit BMN 110 (einmal wöchentlich) in der Studie MOR-004 als statistisch signifikanten und klinisch relevanten therapeutischen Effekt (54).

Bei Patienten mit MPS IVA führt die eingeschränkte Mobilität und Ausdauer zu einer verminderten Lebensqualität und Motivation (2). Bei Patienten mit schwer eingeschränkten Organfunktionen wie MPS sind bereits kleine Verbesserungen der Gehstrecke klinisch relevant und spiegeln die Fähigkeit des Patienten wider, Alltagsaktivitäten zu verrichten (10).

Der geschätzte Behandlungseffekt vom BMN 110 einmal wöchentlich auf die mittlere Gehstreckenverbesserung im 6MWT betrug in der 24. Woche 22.5 m (95% KI 4.0, 40.9;  $p=0.0174$ ) gegenüber Placebo. Der Unterschied der prozentualen Änderung des 6MWT in Woche 24 zwischen der BMN 110 (einmal wöchentlich) und der Placebo-Gruppe betrug 14.9% (95% KI, 2.7, 27.2;  $p=0.017$ ). Die Zwischenauswertung einer Langzeit-Verlängerungsstudie deutete darauf hin, dass der Behandlungseffekt bis Woche 72 zumindest bestehen bleibt.

Die quantitative Auswertung der Gehstreckenunterschiede mittels einer kontinuierlichen Responderanalyse zeigte eine signifikante Überlegenheit für alle mit BMN 110 (einmal wöchentlich) behandelten Patienten gegenüber der gesamten Placebo-Gruppe. Approximierte

relative Risiken der Gehstreckenverbesserungen von 14 m bis  $\geq 95$  m zeigten alle eine signifikante Überlegenheit von BMN 110 (einmal wöchentlich) gegenüber der Vergleichstherapie. Die therapeutisch relevante Abschwächung schwerwiegender Symptome stellt für diesen Endpunkt somit einen beträchtlichen Zusatznutzen dar.

Da in die Studie MOR-004 ein beträchtlicher Anteil von asiatischen Patienten eingeschlossen war, wurde unter Berücksichtigung der Kriterien internationaler Richtlinien (50,51) die kontinuierliche Responderanalyse in weissen Patienten durchgeführt, um die Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext zu demonstrieren. Auch hier zeigte sich, dass alle mit Elosulfase alfa behandelten Patienten gegenüber den Placebo-behandelten Patienten eine therapeutisch relevante Abschwächung schwerwiegender Krankheitssymptome erfuhren. Die Verbesserung der Gehstrecke war für alle mit Elosulfase alfa behandelten Patienten unabhängig von den Baselinewerten so deutlich, dass die Konfidenzintervalle der approximierten relativen Risiken im therapeutisch relevanten Bereich alle unter dem Schwellenwert von 0.9 lagen.

Bei einer derartig schwerwiegenden Erkrankung wie MPS IVA können bereits minimale Änderungen klinisch bedeutsam sein. Ein signifikanter Unterschied zwischen BMN 110 (einmal wöchentlich) und der Placebo-Gruppe kann bei der gegebenen schweren Krankheitssymptomatik *per se* als bedeutsamer Zusatznutzen gewertet werden. Eine durchschnittliche Zunahme der Gehstrecke um 22.5 m unter Behandlung mit BMN 110 (einmal wöchentlich) kann für MPS IVA-Patienten eine durchaus im Alltag relevant Veränderung bedeuten, beispielsweise bei alltagsbezogenen Verrichtungen im Bereich der Körperpflege und Hygiene, oder beim Überqueren einer Straße. Im MPS HAQ gaben mehr Patienten der BMN 110-Gruppen als der Placebo-Gruppe an, einen Rollstuhl zu benutzen (Baselinewerte: 57.6% und 51.7% vs. 37.3%). Dies deutet einen höheren Behinderungsgrad in den BMN 110-Gruppen verglichen mit der Placebo-Gruppe an, welcher deutlichere Verbesserungen im 6MWT unter Behandlung mit BMN 110 verhindert haben könnte. Ansonsten ergaben sich zu Behandlungsbeginn keine relevanten Imbalancen zwischen den Gruppen.

Die publizierten MCID-Werte im 6MWT beziehen sich auf andere Krankheitsbilder als MPS IVA (11) und mehrheitlich auf Patienten, deren Ausgangsgehstrecke im 6MWT deutlich besser war (meist  $>350$  m), als in der pivotalen MOR-004-Studie oder in anderen Studien bei MPS IVA-Patienten. Dennoch ist festzustellen, dass bei diesen Patienten mit gegenüber MPS IVA-Patienten besserer Gehfähigkeit, bereits moderate Änderungen der Gehstrecke zwischen 24 m und 54 m eine für den Patienten spürbare Verbesserung mit sich bringen (11). In einigen Studien waren bereits kleine Änderungen von 5% bis 10% für die Patienten wahrnehmbar (11). Nach einer Studie bei Patienten mit Duchenne-Syndrom schwankt der MCID abhängig von der Ausgangsgehstrecke (32), wie dies auch von den Experten des Delphi-Panels bei MPS IVA angenommen wurde. Bei geringeren Gehstrecken war bereits eine Änderung von nur 5.6 m für die Patienten relevant, während bei denjenigen mit längerer Gehstrecke erst eine Änderung von 45.9 m von den Patienten wahrgenommen wurde (32). Allerdings ist zu beachten, dass in der Studie von Henricson et al. (2013) (32) Jungen im Alter von 4-12 Jahren teilgenommen haben. Bei Duchenne-Patienten im Alter zwischen 5 und 20 Jahren (358 m

Gehstrecke) betrug die MCID des 6MWT 28.5 und 31.7 m (8.0% bzw. 8.9% des mittleren Ausgangswerts) basierend auf zwei statistischen Methoden (33). Es wird eingeräumt, dass bei Duchenne-Patienten im Gegensatz zu MPS IVA-Patienten skeletale Dysplasien kaum eine Rolle spielen. Leider konnte kein MCID bei anderen Erkrankungen, bei denen skeletale Dysplasien vorherrschen, identifiziert werden.

Unter Berücksichtigung dieser Einschränkung legen die Ergebnisse bei Duchenne-Patienten nahe, dass aus Sicht der Patienten die Ausgangsgehstrecke unabhängig von der zugrunde liegenden Erkrankung oder den Beschwerden für die Bewertung des Therapieerfolgs von entscheidender Bedeutung sein könnte. Bei Patienten mit MPS I verlängerte sich die Gehstrecke ausgehend von  $319.1 \pm 131.4$  m ( $366.7 \pm 113.7$  m Placebo) unter Therapie mit Laronidase innerhalb von 26 Wochen um 38 m ( $19.7 \pm 68.6$  m Laronidase vs.  $-18.4 \pm 67.5$  m Placebo) (20). Diese Änderung lag bezogen auf die Ausgangsgehstrecke im 6MWT größenordnungsmäßig innerhalb der von Henricson et al. (2013) bei Duchenne-Patienten festgelegten MCID ( $276.7-304.9$  m: 28.2 m,  $344.3-380.8$  m: 36.5 m) (32). In der pivotalen MOR-004-Studie betrug die mittlere initiale Gehstrecke ( $\pm$  SD) 211.9 (69.88), 205.7 (81.19), bzw. 203.9 (76.32) m in der Placebo- und den BMN 110-Gruppen (jede zweite Woche bzw. jede Woche). Unter Berücksichtigung der von Henricson et al. (2013) (32) bei Duchenne-Patienten ermittelten MCIDs wird angenommen, dass für die Teilnehmer der MOR-004-Studie ausgehend von einer Gehstrecke von etwa 205 m eine Änderung zwischen 12.7 und 20.9 m klinisch relevant sein dürfte (MCID bei Duchenne-Patienten:  $189.5-202.2$  m: 12.7 m;  $225.9-246.8$  m: 20.9 m), wobei bei älteren Patienten mit einer kürzeren Gehstrecke eher eine MCID im unteren Bereich angenommen werden kann.

Die Änderung der Gehstrecke in der Studie MOR-004 unter Behandlung mit BMN 110 einmal wöchentlich lag mit 22.5 m oberhalb des angenommenen MCIDs bezogen auf die Ausgangsgehstrecke. Unter Berücksichtigung, dass bei weniger in ihrer Bewegung eingeschränkten Patienten bereits Änderungen der Gehstrecke im Bereich von 5% bis 10% von den Patienten wahrgenommen werden (11), ist eine mittlere Änderung der Gehstrecke von nahezu 15% unter Behandlung mit BMN 110 einmal wöchentlich bei MPS IVA-Patienten als klinisch bedeutsam einzustufen. Bei der Ausdauerfähigkeit handelt es sich um eine vitale menschliche Funktion, die dementsprechend auch für MPS IVA-Patienten von entscheidender Bedeutung ist (9) und wahrscheinlich einen erheblichen Einfluss auf deren Lebensqualität hat (2).

Die Auswertung des MPS HAQ zeigte positive Trends für die Überlegenheit von BMN 110 hinsichtlich der Mobilität und der Hilfe durch Betreuer bei Alltagsaktivitäten. Möglicherweise war die Studiendauer jedoch zu kurz, um bei diesen Endpunkten signifikante Änderungen erwarten zu können. Die Zahl der Patienten, die einen Rollstuhl benötigten stieg in der Placebo-Gruppe im Verlauf der Studie um fünf Patienten (8.8%) an, während in den BMN 110-Gruppen keiner der Patienten zusätzlich einen Rollstuhl benötigte. Dieser Befund ist potentiell von hoher klinischer Relevanz, die Größe des Effektes kann jedoch zu diesem Zeitpunkt nicht eingeschätzt werden.

Hinsichtlich des Parameters MVV zeigte sich ein Trend zugunsten von BMN 110 einmal wöchentlich verglichen mit Placebo. Andere weniger sensitive respiratorische Parameter unterschieden sich zu Woche 24 im Gruppenvergleich nicht signifikant. Diese Ergebnisse haben keinen Einfluss auf die Bewertung des Ausmaßes des Zusatznutzens von BMN 110. Die Ergebnisse einer Zwischenauswertung der Verlängerungsstudie MOR-005 zeigten jedoch schon signifikante Veränderungen des Endpunktes ‚forcierte Vitalkapazität‘ (MVV) in BMN 110 behandelten Patienten gegenüber dem Ausgangszustand.

Bei noch im Wachstum befindlichen Patienten, zeigten beide BMN 110-Dosen einen positiven Effekt auf die Wachstumsrate und die Körpergröße, der Unterschied verglichen mit Placebo war allerdings nicht signifikant, was auf die geringe Fallzahl zurückzuführen ist. Auch dieser Befund ist potentiell von hoher klinischer Relevanz. Das Ergebnis der Woche 24 hat jedoch keinen Einfluss auf die Bewertung des Ausmaßes des Zusatznutzens von BMN 110.

### **Unerwünschte Ereignisse**

Hinsichtlich der Bewertung der UEs ist zu berücksichtigen, dass zur Behandlung der MPS IVA bislang keine zweckmäßige Vergleichstherapie zugelassen ist. Ein Zusatznutzen im Sinne einer Vermeidung schwerwiegender Nebenwirkungen oder anderer Nebenwirkungen ist dementsprechend für BMN 110 nicht belegbar. Zudem ist zu beachten, dass BMN 110 nur unter streng kontrollierten klinischen Bedingungen und unter enger Beobachtung des Patienten verabreicht wird.

SUEs traten unter Behandlung mit BMN 110 einmal wöchentlich häufiger auf als unter Behandlung mit Placebo. Allerdings ist festzustellen, dass die Mehrzahl der SUEs entweder (temporäre) Infusionsreaktionen überwiegend milder oder mittelschwerer Art oder Begleiterkrankungen des MPS IVA darstellten. Nur ein Patient der BMN 110 qow Gruppe und zwei Patienten, denen BMN 110 einmal wöchentlich verabreicht wurde, entwickelten ein mit der Behandlung in Zusammenhang stehendes (therapieassoziiertes, „drug-dependent“) SUE. Diese SUEs traten während der Infusion auf und waren durch entsprechende Maßnahmen kontrollierbar. Es ergibt sich daher kein Nachteil für BMN 110.

Wie von anderen ERTs bekannt, war die Mehrzahl der UEs unter Behandlung mit BMN 110 Infusionsreaktionen, die mehrheitlich keine Unterbrechung oder einen Abbruch der Infusion und eine medizinische Intervention erforderten (1.26% der wöchentlichen BMN 110-Infusionen). Überempfindlichkeitsreaktionen traten unter Behandlung mit BMN 110 etwas häufiger auf als unter Behandlung mit Placebo. Diese Patienten tolerierten nachfolgende Infusion ohne Wiederauftreten der Symptome. Die UEs führten in keinem Fall zum Abbruch der Therapie und alle Patienten, die eine Infusionsreaktion entwickelt hatten, setzten die Behandlung fort. Die Gesamtzahl der IARs unterschied sich im Gruppenvergleich nicht und nahezu alle Patienten entwickelten mindestens ein IAR. Mehr Patienten der BMN 110-Gruppe (einmal wöchentlich) als der Placebo-Gruppe erbrachen sich oder entwickelten Fieber während oder innerhalb eines Tages nach Ende der Infusion (Tabelle 4-63). Nur einzelne Patienten entwickelten schwerwiegende IARs und deren Inzidenz unterschied sich im

Gruppenvergleich nicht signifikant. Mit Ausnahme von IARs ergab sich kein kausaler Zusammenhang zwischen der Behandlung und klinisch bedeutungsvollen Änderungen der Laborparameter und Vitalzeichen. Daher ist festzustellen, dass es sich bei den Infusionsreaktionen um temporäre UEs handelt, die kontrollierbar und in der überwiegenden Mehrzahl nicht schwerwiegend sind. Unter Berücksichtigung, dass es sich um vorübergehende, meist nicht schwerwiegende UEs handelt, ergibt sich hinsichtlich der UEs insbesondere der Infusionsreaktionen allenfalls ein geringer Nachteil für BMN 110 gegenüber einer Nichtbehandlung, dem ein beträchtlicher Zusatznutzen für den Wirksamkeitsendpunkt gegenüber steht.

Aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten ergibt sich somit insgesamt ein beträchtlicher Zusatznutzen für das zu bewertende Arzneimittel im Vergleich zur zweckmäßigen Vergleichstherapie

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

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

Tabelle 4-84: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

<b>Bezeichnung der Patientengruppen</b>	<b>Ausmaß des Zusatznutzens</b>
Patienten mit MPS IV A	Beträchtlich

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

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

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

Nicht zutreffend.

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

*Sofern mit dem Dossier nichtrandomisierte vergleichende Studien (Abschnitt 0) oder weitere Untersuchungen (Abschnitt 0) eingereicht wurden, nennen Sie die Gründe, nach denen es*

*unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.*

Nicht zutreffend.

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

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

Nicht zutreffend.

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

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

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

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006<sup>15</sup>) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%-Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006<sup>16</sup>) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

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

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

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.*

Entsprechend des IQWiG (14) werden „zur Validierung eines Surrogats umfangreiche Daten benötigt, vorzugsweise eine Meta-Analyse von mehreren randomisierten Studien mit ausreichender Ergebnissicherheit. In diesen Studien müssen sowohl das Surrogat als auch der klinische Endpunkt erhoben worden sein.“ Für die in der vorliegenden Bewertung evaluierten Surrogatendpunkte liegen keine den Kriterien des IQWiG entsprechenden Validierungsstudien vor.

Eine noch andauernde multinationale (einschließlich Deutschland), prospektive Querschnittsstudie (MorCAP) hat zum Ziel, den natürlichen Verlauf unbehandelter MPS IVA-Patienten zu charakterisieren (3). Bislang liegen Ausgangsdaten von 325 Patienten mit bestätigter MPS IVA-Diagnose (molekulargenetische Testung oder reduzierte GALNS-Aktivität) aller Altersgruppen und Schweregrade vor (n=45 0-4 Jahre, n=127 5-11 Jahre, n=84 12-18 Jahre, n=69 >18 Jahre). Ein Follow-up von bis zu 10 Jahren ist geplant. Zu jeder Visite wurde die Krankheitsgeschichte abgefragt und es fanden klinische Tests und eine ärztliche Untersuchung statt. Der 6MWT, 3MSCT und die respiratorischen Funktionstests wurden entsprechend relevanter Richtlinien ausgeführt. Zusätzlich wurde der MPS HAQ von den Patienten oder deren Betreuer ausgefüllt. Das Geschlechterverhältnis war ausgeglichen (52.9% weiblich) und die Mehrzahl der Patienten (73.8%) waren weiß. Die uKS- und Kreatinin-Spiegel wurden im ersten Morgenurin mittels einer validierten Methode bestimmt. Die normalisierten uKS-Spiegel schwankten bei MPS IVA-Patienten in Abhängigkeit vom Alter und waren bei älteren Patienten niedriger als bei jüngeren Patienten (0-4 Jahre:  $56.9 \pm 25.1$   $\mu\text{g uKS /mg Kreatinin}$ , 5-11 Jahre:  $46.8 \pm 30.6$   $\mu\text{g uKS /mg Kreatinin}$ , 12-18 Jahre:  $30.5 \pm 20.4$   $\mu\text{g uKS /mg Kreatinin}$ , >18 Jahre:  $11.9 \pm 8.1$   $\mu\text{g uKS /mg Kreatinin}$ ,  $p \leq 0.0230$ ). Basierend auf der Korrelation zwischen dem uKS und der Gehstrecke im 6MWT wurden Schwellenwerte mittels LOWESS („locally weighted scatterplot smoothing“) identifiziert, um der altersbedingten Abnahme gerecht zu werden. Ein niedriger uKS war definiert als  $\leq 20$   $\mu\text{g uKS /mg Kreatinin}$  bei Patienten im Alter von 18 Jahren und jünger und als  $\leq 10$   $\mu\text{g uKS /mg Kreatinin}$  bei Patienten,

die älter als 18 Jahre waren. Im Mittel zeigten Patienten mit einem hohen uKS-Spiegel größere klinische Einschränkungen (Körpergröße, z-Scores) sowie Ausdauer- und respiratorische Funktionstests). Bei den älteren Patienten war die Gehstrecke im 6MWT (jeweils Mittelwert  $\pm$  SD, 265.7 $\pm$ 140.9 m vs. 136.6 $\pm$ 125.9 m,  $p=0.0002$ ), die Zahl der erklommenen Stufen im 3MSCT (34.0 $\pm$ 20.6 vs. 16.4 $\pm$ 19.3 Stufen/min,  $p=0.0016$ ), die FVC (2.1 $\pm$ 1.3 l vs. 0.8 $\pm$ 0.4 l,  $p<0.0001$ ) und das MVV (59.2 $\pm$ 37.5 l/min vs. 24.2 $\pm$ 10.9 l/min,  $p<0.0001$ ) bei den Patienten mit niedrigem uKS-Spiegel ( $n=32$ ) signifikant höher als bei Patienten mit hohem uKS ( $n=33$ ,  $n=4$  fehlend; ANCOVA: niedriges vs. hohes uKS kontrolliert für die Altersgruppen). Bei jüngeren Patienten ( $\leq 18$  Jahre) zeigten sich signifikante Unterschiede zugunsten der Patienten mit niedrigem uKS ( $n=44$ ) vs. hohem uKS ( $n=201$ ,  $n=1$  fehlend) bei den pulmonalen Parametern FVC (1.9 $\pm$ 1.0 vs. 0.9 $\pm$ 0.4 l,  $p<0.0001$ ) und MVV (56.1 $\pm$ 29.6 l/min vs. 25.7 $\pm$ 13.1 l/min,  $p<0.0001$ ). Die p-Werte des Testes auf Gleichheit („equality“) der niedrigen vs. hohen uKS-Gruppe waren für alle Parameter (6MWT, 3MSCT, FVC, MVV) signifikant. Bei den pädiatrischen Patienten ( $\leq 18$  Jahre) zeigte sich eine negative Korrelation zwischen den uKS-Spiegeln und der Körpergröße. Entsprechend der Autoren könnte ein hoher uKS-Spiegel eine rascher voranschreitende Erkrankung anzeigen. Die Autoren merken allerdings an, dass weitere longitudinale Daten erforderlich sind, um den prognostischen Wert von uKS und die Assoziation zwischen diesem Biomarker und dem klinischen Verlauf definitiv zu bestimmen (3). In den Analysen zum Einfluss von uKS auf klinische Endpunkte wurden die Daten bei der Gruppe der Patienten im Alter von 18 Jahren und jünger zusammen gefasst und nicht nach Altersgruppen stratifiziert ausgewertet.

Bei 47 Patienten mit schwerer und 12 Patienten mit einer leichteren Form der MPS IVA wurden 70 Urinproben zur Bestimmung des uKS mittels eines neu entwickelten und validierten ELISAs entnommen. Patienten mit einer schweren Form der MPS IVA hatten signifikant höhere uKS-Spiegel als Patienten mit einer milden Form (Mittelwert  $\pm$  SD: 11.21 $\pm$ 9.72 vs. 1.67 $\pm$ 1.43 ng uKS /g Kreatinin,  $p<0.0001$ ) (20).

Da sich hinsichtlich anderer Surrogatendpunkte wie MVV oder FEV1 in der Studie MOR-004 keine signifikanten Unterschiede zwischen BMN 110 und Placebo ergaben, wird auf eine Diskussion dieser Endpunkte verzichtet.

#### 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).*

- [9] Dummer W, Slasor P, Boyd H, Lounsbury D, Glasscock B, Gagnon B. A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Multinational Clinical Study to Evaluate the Efficacy and Safety of 2.0 mg/kg/week and 2.0 mg/kg/every other week BMN 110 in Patients with Mucopolysaccharidosis IVA (Morquio A Syndrome). Clinical Study Report MOR-004, 11 March 2013.
- [39] BioMarin Pharmaceutical. A Double-Blind Study to Evaluate the Efficacy and Safety of BMN 110 in Patients With Mucopolysaccharidosis IVA (Morquio A Syndrome). NCT01275066 [online]. 2013 Aug 2 [cited 2013 Aug 22];[3 screens]. Available from: <http://clinicaltrials.gov/ct2/show/study/NCT01275066?term=BMN+110&rank=2> URL:
- [40] BioMarin Pharmaceutical. A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Multinational Clinical Study to Evaluate the Efficacy and Safety of 2.0 mg/kg/week and 2.0 mg/kg/every other week BMN 110 in Patients with Mucopolysaccharidosis IVA (Morquio A Syndrome) [online]. 2012 Nov 15 [cited 2013 Aug 22];[2 screens]: Available from: [URL: http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-020198-18-GB](http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-020198-18-GB)
- [52] Dummer W, Slasor P, Boyd H, Lounsbury D, Glasscock B, Gagnon B. A. A Multicenter, Multinational, Extension Study to Evaluate the Long-Term Efficacy and Safety of BMN 110 in Patients with Mucopolysaccharidosis IVA (Morquio A Syndrome). Clinical Study Report MOR-005, 11 March 2013.
- [53] BioMarin Pharmaceutical. Long-Term Efficacy and Safety Extension Study of BMN 110 in Patients With Mucopolysaccharidosis IVA (Morquio A Syndrome) [online]. September 2012 [cited 2014 Feb 19]: Available from: [URL: http://clinicaltrials.gov/ct2/results/displayOpt?flds=a&flds=b&flds=q&submit\\_fld\\_opt=on&term=mor-005&show\\_flds=Y](http://clinicaltrials.gov/ct2/results/displayOpt?flds=a&flds=b&flds=q&submit_fld_opt=on&term=mor-005&show_flds=Y)

#### 4.7 Referenzliste

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

1. Tomatsu S, Montañó AM, Oikawa H, Smith M, Barrera L, Chinen Y, u. a. Mucopolysaccharidosis type IVA (Morquio A disease): clinical review and current treatment. *Curr Pharm Biotechnol.* Juni 2011;12(6):931–45.
2. Hendriksz CJ, Al-Jawad M, Berger KI, Hawley SM, Lawrence R, Mc Ardle C, u. a. Clinical overview and treatment options for non-skeletal manifestations of mucopolysaccharidosis type IVA. *J Inher Metab Dis.* März 2013;36(2):309–22.

3. Harmatz P, Mengel KE, Giugliani R, Valayannopoulos V, Lin S-P, Parini R, u. a. The Morquio A Clinical Assessment Program: baseline results illustrating progressive, multisystemic clinical impairments in Morquio A subjects. *Mol Genet Metab.* Mai 2013;109(1):54–61.
4. Montaña AM, Tomatsu S, Gottesman GS, Smith M, Orii T. International Morquio A Registry: clinical manifestation and natural course of Morquio A disease. *J Inher Metab Dis.* April 2007;30(2):165–74.
5. European Medicines Agency (EMA). Committee for Orphan Medicinal Products. Public summary of positive opinion for orphan designation of recombinant human N-acetylgalactosamine-6-sulfatase for the treatment of mucopolysaccharidosis, type IVA (Morquio A syndrome) [Internet]. 2009. Verfügbar unter: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Orphan\\_designation/2009/10/WC500005877.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Orphan_designation/2009/10/WC500005877.pdf)
6. Gemeinsamer Bundesausschuss. Verfahrensordnung des Gemeinsamen Bundesausschusses [Internet]. 2014. Verfügbar unter: [https://www.g-ba.de/downloads/62-492-861/VerfO\\_2014-01-23.pdf](https://www.g-ba.de/downloads/62-492-861/VerfO_2014-01-23.pdf)
7. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden Version 4.1 [Internet]. 2013. Verfügbar unter: [https://www.iqwig.de/download/IQWiG\\_Methoden\\_Version\\_4-1.pdf](https://www.iqwig.de/download/IQWiG_Methoden_Version_4-1.pdf)
8. Tomatsu S, Mackenzie W, Theroux M, Mason R, Thacker M, Shaffer T, u. a. Current and emerging treatments and surgical interventions for Morquio A syndrome: a review. *Res Rep Endocr Disord.* 2012;2:65–77.
9. Dummer W, Slasor P, Boyd H, Lounsbury D, Glasscock B, Gagnon B. A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Multinational Clinical Study to Evaluate the Efficacy and Safety of 2.0 mg/kg/week and 2.0 mg/kg/every other week BMN 110 in Patients with Mucopolysaccharidosis IVA (Morquio A Syndrome). 2013.
10. McDonald A, Steiner R, Kuehl K, Turbeville S. Clinical utility of endurance measures for evaluation of treatment in patients with mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). *J Pediatr Rehabil Med.* 2010;3(2):119–27.
11. Lachmann R, Schoser B. The clinical relevance of outcomes used in late-onset Pompe disease: can we do better? *Orphanet J Rare Dis.* 2013;8:160.
12. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Ticagrelor-Nutzenbewertung gemäß § 35a SGB V [Internet]. 2011 [zitiert 20. März 2014]. Verfügbar unter: [http://www.g-ba.de/downloads/92-975-5/2011-01-01-D-001\\_Ticagrelor\\_IQWiG-Nutzenbewertung.pdf](http://www.g-ba.de/downloads/92-975-5/2011-01-01-D-001_Ticagrelor_IQWiG-Nutzenbewertung.pdf)
13. Regier DS, Oetgen M, Tanpaiboon P. Mucopolysaccharidosis Type IVA. In: Pagon RA, Adam MP, Bird TD, Dolan CR, Fong C-T, Smith RJ, u. a., Herausgeber. *GeneReviews*(®) [Internet]. Seattle (WA): University of Washington, Seattle; 1993

- [zitiert 20. März 2014]. Verfügbar unter:  
<http://www.ncbi.nlm.nih.gov/books/NBK148668/>
14. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Aussagekraft von Surrogatendpunkten in der Onkologie [Internet]. Verfügbar unter:  
[https://www.iqwig.de/download/A10-05\\_Rapid\\_Report\\_Version\\_1-1\\_Surrogatendpunkte\\_in\\_der\\_Onkologie.pdf](https://www.iqwig.de/download/A10-05_Rapid_Report_Version_1-1_Surrogatendpunkte_in_der_Onkologie.pdf)
  15. BioMarin Pharmaceutical. Advisory Committee Briefing Materials: Available for Public Release. Vimizim (elosulfase alfa) for the treatment fo Mucopolysaccharidosis Type IVA (Morquio A syndrome) [Internet]. 2013 [zitiert 20. März 2014]. Verfügbar unter:  
[http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drug\\_s/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM375127.pdf](http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drug_s/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM375127.pdf)
  16. 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.
  17. McDonald CM, Henricson EK, Abresch RT, Florence JM, Eagle M, Gappmaier E, u. a. The 6-minute walk test and other endpoints in Duchenne muscular dystrophy: longitudinal natural history observations over 48 weeks from a multicenter study. *Muscle Nerve.* September 2013;48(3):343–56.
  18. Harmatz P, Ketteridge D, Giugliani R, Guffon N, Teles EL, Miranda MCS, u. a. Direct comparison of measures of endurance, mobility, and joint function during enzyme-replacement therapy of mucopolysaccharidosis VI (Maroteaux-Lamy syndrome): results after 48 weeks in a phase 2 open-label clinical study of recombinant human N-acetylgalactosamine 4-sulfatase. *Pediatrics.* Juni 2005;115(6):e681–689.
  19. Harmatz P, Giugliani R, Schwartz IVD, Guffon N, Teles EL, Miranda MCS, u. a. Long-term follow-up of endurance and safety outcomes during enzyme replacement therapy for mucopolysaccharidosis VI: Final results of three clinical studies of recombinant human N-acetylgalactosamine 4-sulfatase. *Mol Genet Metab.* August 2008;94(4):469–75.
  20. Wraith JE, Clarke LA, Beck M, Kolodny EH, Pastores GM, Muenzer J, u. a. Enzyme replacement therapy for mucopolysaccharidosis I: a randomized, double-blinded, placebo-controlled, multinational study of recombinant human alpha-L-iduronidase (laronidase). *J Pediatr.* Mai 2004;144(5):581–8.
  21. European Medicines Agency (EMA) (Vorname). Guideline on the clinical investigation of medicinal products for the treatment of Duchenne and Becker muscular dystrophy. Draft. EMA/CHMP/236981/2011 [Internet]. 2013. Verfügbar unter:  
[http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2013/03/WC500139508.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/03/WC500139508.pdf)
  22. Olsson LG, Swedberg K, Clark AL, Witte KK, Cleland JGF. Six minute corridor walk test as an outcome measure for the assessment of treatment in randomized, blinded intervention trials of chronic heart failure: a systematic review. *Eur Heart J.* April 2005;26(8):778–93.

23. Berger KI, Fagondes SC, Giugliani R, Hardy KA, Lee KS, McArdle C, u. a. Respiratory and sleep disorders in mucopolysaccharidosis. *J Inher Metab Dis.* März 2013;36(2):201–10.
24. Giugliani R, Harmatz P, Wraith JE. Management guidelines for mucopolysaccharidosis VI. *Pediatrics.* August 2007;120(2):405–18.
25. Scarpa M, Almássy Z, Beck M, Bodamer O, Bruce IA, De Meirleir L, u. a. Mucopolysaccharidosis type II: European recommendations for the diagnosis and multidisciplinary management of a rare disease. *Orphanet J Rare Dis.* 2011;6:72.
26. BioMarin. Module 2.7.3. BMN 110. Clinical Summary of Efficacy.
27. Harmatz P, Giugliani R, Schwartz I, Guffon N, Teles EL, Miranda MCS, u. a. Enzyme replacement therapy for mucopolysaccharidosis VI: a phase 3, randomized, double-blind, placebo-controlled, multinational study of recombinant human N-acetylgalactosamine 4-sulfatase (recombinant human arylsulfatase B or rhASB) and follow-on, open-label extension study. *J Pediatr.* April 2006;148(4):533–9.
28. Muenzer J, Wraith JE, Beck M, Giugliani R, Harmatz P, Eng CM, u. a. A phase II/III clinical study of enzyme replacement therapy with idursulfase in mucopolysaccharidosis II (Hunter syndrome). *Genet Med Off J Am Coll Med Genet.* August 2006;8(8):465–73.
29. Noonan V, Dean E. Submaximal exercise testing: clinical application and interpretation. *Phys Ther.* August 2000;80(8):782–807.
30. Hassan J, van der Net J, Helders PJM, Prakken BJ, Takken T. Six-minute walk test in children with chronic conditions. *Br J Sports Med.* 1. März 2010;44(4):270–4.
31. Du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, u. a. Six-minute-walk test in idiopathic pulmonary fibrosis: test validation and minimal clinically important difference. *Am J Respir Crit Care Med.* 1. Mai 2011;183(9):1231–7.
32. Henricson E, Abresch R, Han JJ, Nicorici A, Goude Keller E, de Bie E, u. a. The 6-minute walk test and person-reported outcomes in boys with duchenne muscular dystrophy and typically developing controls: longitudinal comparisons and clinically-meaningful changes over one year. *PLoS Curr.* 2013;5.
33. 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.
34. Dreher M, Waltersbacher S, Sonntag F, Prettin S, Kabitz HJ, Windisch W. Exercise in severe COPD: is walking different from stair-climbing? *Respir Med.* Juni 2008;102(6):912–8.
35. Montañó AM, Tomatsu S, Brusius A, Smith M, Orii T. Growth charts for patients affected with Morquio A disease. *Am J Med Genet A.* 15. Mai 2008;146A(10):1286–95.

36. Stein R, Selvadurai H, Coates A, Wilkes DL, Schneiderman-Walker J, Corey M. Determination of maximal voluntary ventilation in children with cystic fibrosis. *Pediatr Pulmonol.* Juni 2003;35(6):467–71.
37. Harmatz P, Yu Z-F, Giugliani R, Schwartz IVD, Guffon N, Teles EL, u. a. Enzyme replacement therapy for mucopolysaccharidosis VI: evaluation of long-term pulmonary function in patients treated with recombinant human N-acetylgalactosamine 4-sulfatase. *J Inher Metab Dis.* Februar 2010;33(1):51–60.
38. Tomatsu S, Okamura K, Taketani T, Orii KO, Nishioka T, Gutierrez MA, u. a. Development and testing of new screening method for keratan sulfate in mucopolysaccharidosis IVA. *Pediatr Res.* April 2004;55(4):592–7.
39. BioMarin Pharmaceutical. A Double-Blind Study to Evaluate the Efficacy and Safety of BMN 110 in Patients With Mucopolysaccharidosis IVA (Morquio A Syndrome) [online]. [Internet]. 2013. Verfügbar unter: <http://clinicaltrials.gov/ct2/show/study/NCT01275066?term=BMN+110&rank=2>
40. BioMarin Pharmaceutical. A Phase 3, Randomized, Double-Blind, Placebo-Controlled, Multinational Clinical Study to Evaluate the Efficacy and Safety of 2.0 mg/kg/week and 2.0 mg/kg/every other week BMN 110 in Patients with Mucopolysaccharidosis IVA (Morquio A Syndrome) [online] [Internet]. 2012. Verfügbar unter: <http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-020198-18-GB>
41. Rahlfs VW, Zimmermann H, Lees KR. Effect size measures and their relationships in stroke studies. *Stroke J Cereb Circ.* Februar 2014;45(2):627–33.
42. Wei LJ, Lachin JM. Two-Sample Asymptotically Distribution-Free Tests for Incomplete Multivariate Observations. *J Am Stat Assoc.* 1. September 1984;79(387):653–61.
43. Thall PF, Lachin JM. Analysis of Recurrent Events: Nonparametric Methods for Random-Interval Count Data. *J Am Stat Assoc.* Juni 1988;83(402):339.
44. Lachin JM. Some large-sample distribution-free estimators and tests for multivariate partially incomplete data from two populations. *Stat Med.* 30. Juni 1992;11(9):1151–70.
45. Daly LE. Confidence limits made easy: interval estimation using a substitution method. *Am J Epidemiol.* 15. April 1998;147(8):783–90.
46. Zimmermann H, Rahlfs VW. Comments on number-needed-to-treat derived from ordinal scales. *Stat Methods Med Res.* Februar 2014;23(1):107–10.
47. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, u. a. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol.* Dezember 2011;64(12):1311–6.
48. Snapinn S, Jiang Q. Responder analyses and the assessment of a clinically relevant treatment effect. *Trials.* 25. Oktober 2007;8:31.

49. Kleist P. Wie sinnvoll sind Responderanalysen in klinischen Studien? Responderanalysen sollen kritisch hinterfragt werden. Schweiz Med-Forum. 2010;10(9):169.
50. Office of the Commissioner C for BE and R. Guidances - Collection of Race and Ethnicity Data in Clinical Trials [Internet]. [zitiert 20. März 2014]. Verfügbar unter: <http://www.fda.gov/RegulatoryInformation/Guidances/ucm126340.htm>
51. European Medicines Agency (EMA). E 5 (R1) Ethnic Factors in the Acceptability of Foreign Clinical Data [Internet]. 1998 [zitiert 20. März 2014]. Verfügbar unter: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500002842.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002842.pdf)
52. Dummer W, Slasor P, Boyd H, Lounsbury D, Glasscock B, Gagnon B. A Multicenter, Multinational, Extension Study to Evaluate the Long-Term Efficacy and Safety of BMN 110 in Patients with Mucopolysaccharidosis IVA (Morquio A Syndrome). 2013.
53. BioMarin Pharmaceutical. Long-Term Efficacy and Safety Extension Study of BMN 110 in Patients With Mucopolysaccharidosis IVA (Morquio A Syndrome) - Full Text View - ClinicalTrials.gov [Internet]. 2012 [zitiert 20. März 2014]. Verfügbar unter: [http://clinicaltrials.gov/ct2/results/displayOpt?flds=a&flds=b&flds=q&submit\\_fld\\_opt=on&term=mor-005&show\\_flds=Y](http://clinicaltrials.gov/ct2/results/displayOpt?flds=a&flds=b&flds=q&submit_fld_opt=on&term=mor-005&show_flds=Y)
54. Committee for Medicinal Products for Human Use. Assessment report. Vimizim. CHMP; 2014 Feb S. 1–67. Report No.: EMA/145808/2014.

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

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

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

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

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

Nicht zutreffend

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

Nicht zutreffend

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

Nicht zutreffend

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

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

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

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

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

**Studienregister** clinicaltrials.gov  
**Internetadresse** <http://www.clinicaltrials.gov>  
**Datum der Suche** 10.02.2014  
**Suchstrategie** BMN 110 OR GALNS OR Morquio OR N-acetylgalactosamine-6-sulfatase OR N-acetylgalactosamine-6-sulfate OR sulfatase galactose-6-sulfatase [ALL FIELDS]  
**Treffer** 15

**Studienregister** WHO International Clinical Trials Registry Platform (ICRTP)  
**Internetadresse** <http://apps.who.int/trialsearch/>  
**Datum der Suche** 10.02.2014  
**Suchstrategie** BMN 110 OR GALNS OR Morquio OR N-acetylgalactosamine-6-sulfatase OR N-acetylgalactosamine-6-sulfate OR sulfatase galactose-6-sulfatase [ALL FIELDS]  
**Treffer** 37 Einträge (21 Studien)

**Studienregister** EU Clinical Trials Register  
**Internetadresse** <https://www.clinicaltrialsregister.eu/>

**Datum der Suche** 10.02.2014  
**Suchstrategie** BMN 100 OR Mucopolysaccharidosis Type IVA [ALL FIELDS]  
**Treffer** 7 Einträge

**Studienregister** PharmNet.Bund  
**Internetadresse** [https://portal.dimdi.de/websearch/servlet/FlowController/Search#\\_\\_DEFANCHOR\\_\\_](https://portal.dimdi.de/websearch/servlet/FlowController/Search#__DEFANCHOR__)  
**Datum der Suche** 10.02.2014  
**Suchstrategie** BMN?110 [Produktname / Code] ODER N-acetylgalactosamine-6-sulfatase [Aktive Substanz] ODER N-acetylgalactosamine-6-sulfate ODER sulfatase galactose-6-sulfatase [Aktive Substanz] ODER GALNS [Textfeld]  
**Treffer** 4 Einträge

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

Nicht zutreffend

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

Es fand eine uneingeschränkte Suche statt, bei der auch nicht randomisierte, vergleichende Studien mit dem zu bewertenden Arzneimittel identifiziert wurden.

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

Es fand eine uneingeschränkte Suche statt, bei der auch nicht randomisierte, vergleichende Studien mit dem zu bewertenden Arzneimittel identifiziert wurden.

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

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

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

Nicht zutreffend

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

Nicht zutreffend

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

Nicht zutreffend

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

#### 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

Nr.	Referenz	Ausschlussgrund
1.	<p>BioMarin Pharmaceutical. Safety and Exercise Study of BMN 110 for Morquio A Syndrome. NCT01609062. MOR-008 [online]. 2012 Aug [cited 2013 Aug 30]. Available from: URL: <a href="http://clinicaltrials.gov/ct2/show/NCT01609062?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=2">http://clinicaltrials.gov/ct2/show/NCT01609062?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=2</a></p> <p>URL:<a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01609062">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01609062</a></p> <p>URL:<a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-005682-20-DE">http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-005682-20-DE</a></p> <p>URL:<a href="https://portal.dimdi.de/websearch/servlet/FlowController/DisplayDocuments#__DEFANCHOR__">https://portal.dimdi.de/websearch/servlet/FlowController/DisplayDocuments#__DEFANCHOR__</a></p> <p>URL:<a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA">https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA</a></p>	Vergleich von 2 Dosen von BMN 110, Studie nicht abgeschlossen
2.	<p>BioMarin Pharmaceutical. A Phase 4 Multi-center, Multi-national, Open-label, Randomized, Two Dose Level Study of Naglazyme (galsulfase) in Infants with Maroteaux-Lamy Syndrome (MPS VI) - N/A. EUCTR2005-003512-30-FR [online]. 2012 Mar 19 [cited 2013 Aug 30]. Available from: URL:<a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2005-003512-30-FR">http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2005-003512-30-FR</a></p>	BMN 110 nicht getestet, abweichende Indikation

Nr.	Referenz	Ausschlussgrund
3.	BioMarin Pharmaceutical. Study of Recombinant Human N-acetylgalactosamine 4-sulfatase (rhASB) in Patients With MPS VI. NCT00067470. ASB-03-05 [online]. 2012 Oct 17 [cited 2013 Aug 30]. Available from: URL: <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00067470">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00067470</a>	BMN 110 nicht getestet, abweichende Indikation
4.	BioMarin Pharmaceutical. Open-Label Study of Efficacy and Safety of Recombinant Human N-acetylgalactosamine 4-sulfatase in Patients With MPS VI. NCT00048711. ASB-01-04 [online]. 2012 Oct 17 [cited 2013 Aug 30]: Available from: URL: <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00048711">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00048711</a>	BMN 110 nicht getestet, abweichende Indikation
5.	BioMarin Pharmaceutical. Double-Blind, 2 Dose Group Study of Recombinant Human N-Acetylgalactosamine 4-Sulfatase in Patients With MPS VI. NCT00048620. ASB-00-01 [online]. 2012 Oct 17 [cited 2013 Aug 30]. Available from: URL: <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00048620">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00048620</a>	BMN 110 nicht getestet, abweichende Indikation

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

Nicht zutreffend

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

Es wurden keine nicht randomisierten vergleichenden Studien identifiziert.

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

Nr.	Referenz	Ausschlussgrund
6.	<p>BioMarin Pharmaceutical. Long-Term Efficacy and Safety Extension Study of BMN 110 in Patients With Mucopolysaccharidosis IVA (Morquio A Syndrome). NCT01415427. MOR-005 [online]. 2012 Sep [cited 2013 Aug 30]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT01415427?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=1">http://clinicaltrials.gov/ct2/show/NCT01415427?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=1</a>            URL:<a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-020199-45-GB">http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-020199-45-GB</a>            URL:<a href="https://portal.dimdi.de/websearch/servlet/FlowController/DisplayDocuments#__DEFANCHOR__">https://portal.dimdi.de/websearch/servlet/FlowController/DisplayDocuments#__DEFANCHOR__</a>            URL:<a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA">https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA</a></p>	Keine RCT, Studie nicht abgeschlossen
7.	<p>BioMarin Pharmaceutical. Study of BMN 110 in Pediatric Patients &lt; 5 Years of Age With Mucopolysaccharidosis IVA (Morquio A Syndrome). NCT01515956. MOR-007 [online]. 2012 Aug [cited 2013 Aug 30]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT01515956?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=3">http://clinicaltrials.gov/ct2/show/NCT01515956?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=3</a>            URL:  <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01515956">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01515956</a>            URL:  <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-003197-84-GB">http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-003197-84-GB</a>            URL:<a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA">https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA</a></p>	Keine RCT, Studie nicht abgeschlossen

Nr.	Referenz	Ausschlussgrund
8.	<p>BioMarin Pharmaceutical. Efficacy and Safety Study of BMN 110 for Morquio A Syndrome Patients Who Have Limited Ambulation. NCT01697319. MOR-006 [online]. 2013 Apr [cited 2013 Aug 30]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT01697319?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=4">http://clinicaltrials.gov/ct2/show/NCT01697319?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=4</a>            URL:<a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-005703-33-DE">http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2011-005703-33-DE</a>            URL:<a href="https://portal.dimdi.de/websearch/servlet/FlowController/DisplayDocuments#_DEFANCHOR">https://portal.dimdi.de/websearch/servlet/FlowController/DisplayDocuments#_DEFANCHOR</a>            URL:<a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA">https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA</a></p>	Keine RCT, Rekrutierung nicht abgeschlossen
9.	<p>BioMarin Pharmaceutical. A Study to Evaluate the Safety, Tolerability and Efficacy of BMN 110 in Subjects With Mucopolysaccharidosis IVA. NCT00884949. MOR-002 [online]. 2011 Jul [2013 Aug 30]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT00884949?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=6">http://clinicaltrials.gov/ct2/show/NCT00884949?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=6</a>            URL:<a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2008-007365-23-GB">http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2008-007365-23-GB</a>            URL:<a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00884949">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00884949</a>            URL:<a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA">https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA</a></p>	Keine RCT

Nr.	Referenz	Ausschlussgrund
10.	<p>BioMarin Pharmaceutical. A Study to Evaluate the Long-Term Efficacy and Safety of BMN 110 in Patients With Mucopolysaccharidosis IVA (Morquio A Syndrome). NCT01242111. MOR-100 [online]. 2013 Jul [cited 2013 Aug 30]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT01242111?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=7">http://clinicaltrials.gov/ct2/show/NCT01242111?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=7</a>            URL:<a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-021048-16-GB">http://apps.who.int/trialsearch/Trial.aspx?TrialID=EUCTR2010-021048-16-GB</a>            URL:<a href="https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA">https://www.clinicaltrialsregister.eu/ctr-search/search?query=BMN+100+OR+Mucopolysaccharidosis+Type+IVA</a></p>	Keine RCT, Studie nicht abgeschlossen
11.	<p>BioMarin Pharmaceutical. BMN 110 US Expanded Access Program. NCT01858103[online]. 2013 Aug [cited 2013 Aug 30]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT01858103?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=8">http://clinicaltrials.gov/ct2/show/NCT01858103?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=8</a>            URL:  <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01858103">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01858103</a></p>	Keine RCT, Programm nicht abgeschlossen
12.	<p>BioMarin Pharmaceutical. A Clinical Assessment Study of Subjects With Mucopolysaccharidosis IVA (Morquio Syndrome). NCT00787995. MOR-001 [online]. 2011 Jan [cited 2013 Aug 30]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT00787995?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=9">http://clinicaltrials.gov/ct2/show/NCT00787995?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=9</a>            URL:  <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00787995">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00787995</a></p>	Keine RCT

Nr.	Referenz	Ausschlussgrund
13.	<p>Clínica de Hidroterapia e Recuperacao Funcional. Morquio's Syndrome: a Case Study. NCT00609440 [online]. 2008 Jan [cited 2013 Aug 30]. Available from: URL: <a href="http://clinicaltrials.gov/ct2/show/NCT00609440?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=10">http://clinicaltrials.gov/ct2/show/NCT00609440?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=10</a></p> <p>URL: <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00609440">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00609440</a></p>	Keine RCT, BMN 110 nicht getestet
14.	<p>Nadia Ali, BioMarin Pharmaceutical. Psychological Concomitants of Morquio Syndrome (The MAP Study). NCT01752296 [online]. 2013 Jul [cited 2013 Aug 30]. Available from: URL: <a href="http://clinicaltrials.gov/ct2/show/NCT01752296?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=11">http://clinicaltrials.gov/ct2/show/NCT01752296?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=11</a></p> <p>URL: <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01752296">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01752296</a></p>	Keine RCT, BMN 110 nicht getestet, Studie rekrutierend
15.	<p>University of Rostock. Biomarker for Morquio Disease (BioMorquio). NCT01457456 [online]. 2013 Mar [cited 2013 Aug 30]. Available from: URL: <a href="http://clinicaltrials.gov/ct2/show/NCT01457456?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=12">http://clinicaltrials.gov/ct2/show/NCT01457456?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=12</a></p> <p>URL: <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01457456">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01457456</a></p>	Keine RCT, BMN 110 nicht getestet, Studie rekrutierend
16.	<p>Emory University, BioMarin Pharmaceutical. Discovering New Biomarkers For Monitoring Disease Progression in Patients With Mucopolysaccharidosis IVA. NCT01733615 [online]. 2012 Nov [2013 Aug 30]. Available from: URL: <a href="http://clinicaltrials.gov/ct2/show/NCT01733615?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=13">http://clinicaltrials.gov/ct2/show/NCT01733615?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=13</a></p>	Keine RCT, BMN 110 nicht getestet, Studie rekrutierend

Nr.	Referenz	Ausschlussgrund
17.	<p>Oregon Health and Science University, University of Utah, BioMarin Pharmaceutical. Gait Analysis in MPS IVA. NCT01920828 [online]. 2013 Aug [cited 2013 Aug 30]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT01920828?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=14">http://clinicaltrials.gov/ct2/show/NCT01920828?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=14</a>            URL:  <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01920828">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT01920828</a></p>	Keine RCT, BMN 110 nicht getestet, Studie rekrutierend
18.	<p>Children's Hospitals and Clinics of Minnesota, BioMarin Pharmaceutical, Greenwood Genetic Center, Gillette Children's Specialty Healthcare. Diagnosis of Mucopolysaccharidosis Disorders in Patients Presenting With Bilateral Hip Disease. NCT01707433 [online]. 2013 Aug [cited 2013 Aug]. Available from: URL:  <a href="http://clinicaltrials.gov/ct2/show/NCT01707433?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=15">http://clinicaltrials.gov/ct2/show/NCT01707433?term=BMN+110+OR+GALNS+OR+Morquio+OR+N-acetylgalactosamine-6-sulfatase+OR+N-acetylgalactosamine-6-sulfate+OR+sulfatase+galactose-6-sulfatase&amp;rank=15</a></p>	Keine RCT, BMN 110 nicht getestet, Studie rekrutierend
19.	<p>BioMarin Pharmaceutical. A Multicenter, Multinational Open-Label Extension Study of Recombinant Human N-acetylgalactosamine 4-sulfatase (rhASB) in Patients With Mucopolysaccharidosis VI. NCT00104234 [online]. 2012 Oct 17 [cited 2013 Aug 30]. Available from: URL:  <a href="http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00104234">http://apps.who.int/trialsearch/Trial.aspx?TrialID=NCT00104234</a></p>	Keine RCT, BMN 110 nicht getestet, abweichende Diagnose

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

*Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.4 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-85 inklusive eines Flow-Charts für den Patientenfluss.*

*Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-85 inklusive eines Flow-Charts für den Patientenfluss.*

Tabelle 4-85 (Anhang): Studiendesign und -methodik für Studie MOR-004

Item <sup>a</sup>	Charakteristikum	Studieninformation
<b>Studienziel</b>		
2 b	Genaue Ziele, Fragestellung und Hypothesen	<p><u>Primäres Ziel:</u> Fähigkeit von BMN 110 2.0 mg/kg/Woche und 2.0 mg/kg jede zweite Woche im Vergleich zu Placebo, die Ausdauer von Patienten mit MPS IVA zu erhöhen, gemessen mittels 6MWT vom Ausgangswert bis Woche 24</p> <p><u>Sekundäre Ziele:</u> 3MSCT und KS-Spiegel im Urin, jeweils Baseline bis Woche 24 Für tertiäre Ziele s. Zeile 6. Sicherheit und Verträglichkeit von BMN 110 2.0 mg/kg KG/Woche oder jede zweite Woche über 24 Wochen</p>
<b>Methoden</b>		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Randomisierte, doppelblinde, Placebo-kontrollierte, multinationale Parallelgruppen-Studie, Zuteilungsverhältnis 1:1:1 (BMN 110 2.0 mg/kg/Woche oder qow oder Placebo)
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p><u>Amendment 1 (04.10.2010, vor Studienbeginn):</u> Einschluss eines dritten Behandlungsarms (BMN 110 2.0 mg/kg KG jede 2. Woche) mit einem möglicherweise besseren Sicherheitsprofil und höherer Akzeptanz verbunden mit einer möglichen höheren Compliance, Dosisregime ausgewählt auf Basis der intralysosomalen Halbwertszeit von BMN 110 von 5-7 d in Morquio-Fibroblasten, das Dosisregime ähnelt demjenigen anderer ERTs mit vergleichbarer intralysosomaler oder Gewebe-Halbwertszeit</p> <p>Zusätzliche Stratifizierung nach Altersgruppen (5-11, 12-18 und <math>\geq 19</math> Jahre) und 6MWT-Kategorien (<math>\leq 200</math> m und <math>&gt; 200</math> m)</p> <p>Implementierung eines zusätzlichen ARRB, um schwere und schwerwiegende Infusionsreaktionen zu bewerten</p> <p>Auf Empfehlung der EMA und FDA wurden einige der im Studienprotokoll geplanten Analysen im SAP Version 1.0 (24.08.2012) vor Entblindung des Datensatzes (19.10.2012) modifiziert oder ergänzt:</p> <ul style="list-style-type: none"> <li>– Änderung der Imputationsmethoden, um fehlende Werte bei den Ausdauer tests (6MWT und 3MSC) zu ersetzen</li> </ul>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> <li>– Änderung einer Kovariate in der primären Analyse des 6MWT (6MWT-Stratifizierung [<math>\leq 200</math> m, <math>&gt; 200</math> m] anstelle 6MWT Ausgangswert), da der 6MWT ein Stratifizierungskriterium war (bereits in Amendment 1 des Studienprotokolls geändert, 04.10.2010)</li> <li>– Zusätzliche, unterstützende Responder-Analyse der Ausdauer tests (6MWT, 3MSC) basierend auf der KDF</li> <li>– 3MSCT: Die Zahl der Stufen (in 3 min) sowie die Änderung und prozentuale Änderung vom Ausgangswert im Protokoll wurde durch Stufen/min (sekundärer Endpunkt) sowie die absolute und prozentuale Änderung im SAP ersetzt. Einige Patienten erreichten das Ende der Treppe vor 3 min, so dass die Rate als akkurateres Maß angesehen wurde.</li> <li>– Zusätzlicher kombinierter Endpunkt bestehend aus 6MWT, 3MSCT und MVV</li> <li>– Änderung der Definition der ITT-Population: Alle randomisierten Patienten, die mindestens eine Dosis der Prüfpräparate erhalten hatten (SAP) anstelle alle randomisierten Patienten (Protokoll)</li> </ul> <p>Nach Finalisierung des SAPs wurde beschlossen statt der prozentualen Änderung die absolute Änderung des MVV vom Ausgangswert als Komponente des kombinierten Endpunktes auszuwerten, um eine gleiche Wichtung der Komponenten zu ermöglichen.</p> <p>Die Auswertung der Wachstumsrate wurde spezifiziert, um eine genauere Schätzung zu ermöglichen. Die Länge der Wirbelkörper wurde nicht wie im SAP vorgegeben ausgewertet, da nach der Eingangsuntersuchung keine Werte vorlagen.</p> <p>Im Studienprotokoll wurde nicht festgelegt, wer die Echokardiogramme auswerten sollte. Entsprechend des SAPs sollten die Echokardiogramme lokal und zentral ausgewertet werden, wobei die unabhängige zentrale Auswertung Priorität hatte. Da diese Entscheidung nach Einschluss einiger Patienten getroffen wurde, lagen für diese Patienten keine standardisierten Auswertungen vor.</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		Die Sensitivitätsanalysen des 3MSCT wurden im SAP spezifiziert (Stufen/min als Zielgröße).
<b>4</b>	Probanden / Patienten	Patienten
<b>4a</b>	Ein-/Ausschlusskriterien der Probanden / Patienten	<p><u>Einschlusskriterien:</u></p> <p>Alter <math>\geq 5</math> Jahre</p> <p>Dokumentierte, klinische Diagnose von MPS IVA auf Basis klinischer Zeichen und Symptome und dokumentierter, reduzierter Fibroblasten- oder Leukozyten GALNS-Enzymaktivität oder genetischer Testung</p> <p>Schriftliche Einverständniserklärung des Patienten oder im Fall jüngerer Patienten (&lt;18 oder &lt;16 Jahre je nach Region) des legal autorisierten Stellvertreters</p> <p>Mittlerer Screening-6MWT zwischen <math>\geq 30</math> m und <math>\leq 325</math> m</p> <p>Akzeptable Methode der Empfängnisverhütung bei sexuell aktiven Patienten</p> <p>Negativer Schwangerschaftstest bei der Screening-Untersuchung und Bereitschaft zu weiteren Schwangerschaftstests während der Studie</p> <p><u>Ausschlusskriterien:</u></p> <p>Frühere hämatopoetische Stammzelltransplantation</p> <p>Frühere Behandlung mit BMN 110</p> <p>Bekannte Überempfindlichkeit gegenüber Komponenten von BMN 110</p> <p>Größere Operation innerhalb von 3 Monaten vor Studieneintritt oder geplante schwere Operationen während der 24-wöchigen Behandlung</p> <p>Schwangere oder stillende Mütter (Screening) oder Patienten, die planen, während der Studie schwanger zu werden</p> <p>Anwendung eines in der Erprobung befindlichen Produkts oder Medizinprodukts innerhalb von 30 d vor der Screening-Untersuchung oder Notwendigkeit eines investigativen Stoffes jeder Art vor Abschluss aller geplanten Studierhebungen</p> <p>Begleiterkrankungen oder -bedingungen einschließlich aber nicht begrenzt auf Wirbelsäuleninstabilität, klinisch signifikante Kompression des Rückenmarks oder schwere kardiale Erkrankungen, die mit der Teilnahme an der Studie oder der Sicherheit aus Sicht des Prüfarztes interferieren könnten</p> <p>Jedweder Zustand, der aus Sicht des Prüfarztes den Patienten einem höheren</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		Risiko schlechter Compliance oder eines vorzeitigen Abbruchs der Studie aussetzt
4b	Studienorganisation und Ort der Studiendurchführung	33 Studienzentren in 17 Ländern (Argentinien, Brasilien, Kanada, Kolumbien, Dänemark, Frankreich, Deutschland, Italien, Japan, Portugal, Katar, Saudi Arabien, Südkorea, Taiwan, Niederlande, UK, USA)
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p>BMN 110 2.0 mg/kg KG/Woche            BMN 110 0.2 mg/kg KG jede 2. Woche alternierend mit Placebo            Placebo (Vehikel von BMN 110 verdünnt in 0.9% NaCl)</p> <p>jeweils einmal wöchentlich (7±3 d) i.v. im Abstand von mindestens 4 d über etwa 4 h, zu Beginn niedrige Flussrate (3 oder 6 ml/h) mit gradueller Steigerung alle 15 min bis der maximale Fluss (36 bzw. 72 ml/h bei Verdünnung in 100 bzw. 250 ml Kochsalzlösung) erreicht wird, um das Risiko möglicher Infusionsreaktionen zu minimieren</p> <p>Vor jeder Infusion Gabe von Antihistaminika (bevorzugt nicht sedierend) aufgrund möglicher Überempfindlichkeitsreaktionen, bei Patienten mit bekannter Infusionsreaktion oder anderen Risikofaktoren (z.B. Historie allergischer Reaktionen) sedierende Antihistaminika und Prämedikation mit zusätzlichen Arzneimitteln wie H2-Blocker, Montelukast-Natrium oder Steroiden möglich, Gabe von Antipyretika im Ermessen des Arztes</p>
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p><u>Primäres Zielkriterium:</u>            Veränderung des 6MWT vom Ausgangspunkt bis Woche 24 (entsprechend ATS-Richtlinie [10])</p> <p><u>Sekundäre Zielkriterien:</u>            Veränderung des 3MSCT (Stufen/min, supportiv: Zahl der erklommenen Stufen)            KS-Spiegel im Urin (normalisiert auf Kreatinin), jeweils Veränderung vom Ausgangswert bis Woche 24</p> <p><u>Supportive Wirksamkeitsendpunkte:</u>            MVV            Kombinationsendpunkt aus 6MWT, 3MSCT und MVV, jeweils Veränderung vom Ausgangswert bis Woche 24</p> <p><u>Tertiäre Zielkriterien:</u>            PK-Parameters von BMN 110 in einer Subgruppe von Patienten</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Respiratorische Funktionen (FET, FEV<sub>1</sub>, FIVC, FVC, MVV, entsprechend der ATS-Richtlinie)</p> <p>Biochemische Inflammationsmarker (TNF<math>\alpha</math>) und Marker des Knochen- und Knorpel-metabolismus (CTX1 und PIIANP)</p> <p>MPS HAQ</p> <p>Anthropometrische Messungen (Körperlänge, Stehgröße, Größe im Sitzen, Gewicht, z-Scores der Stehgröße und Wachstumsrate)</p> <p>Röntgenbilder der unteren Extremitäten</p> <p>Hörvermögen gemessen mittels Audiometrie (3 Zentren)</p> <p>Echokardiogramm</p> <p>Hornhauttrübung evaluiert mittels klinischer Untersuchung</p> <p><u>Sicherheitsparameter:</u></p> <p>UEs, klinische Routine-Labortests (Serum-chemie, Hämatologie und Urinanalyse), Vitalzeichen, Echokardiogramme, EKGs, körperliche Untersuchung einschließlich Neurologie, Begleitmedikationen, Immunogenitätsstests</p> <p><u>Training der Prüfarzte:</u></p> <p>Das für die Ausdauerstests (6MWT und 3MSCT) verantwortliche Klinikpersonal unterzog sich einem Qualifizierungs- und Trainingsprogramm, das aus einem initialen Prüfertreffen und mindestens zwei Trainingseinheiten vor Ort bestand. Die Zentren und das zuständige Personal wurden dabei zertifiziert. In den meisten Zentren fand das Programm vor Einschluss des ersten Patienten statt.</p> <p>Den Studienzentren wurde ein SMP zur Einschätzung und Bewertung von Überempfindlichkeitsreaktionen und IARs zur Verfügung gestellt. Entsprechend des SMP, fanden monatliche Telefonkonferenzen zwischen dem klinischen Monitor und den Prüfarzten statt, wie IAR zu überwachen, zu handhaben und zu berichten sind.</p>
<b>6b</b>	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Die audiometrischen Untersuchungen, EKGs und Untersuchungen zur Hornhauttrübung wurden in Amendment 1 (vor Studienbeginn) ergänzt.
<b>7</b>	Fallzahl	
<b>7a</b>	Wie wurden die Fallzahlen bestimmt?	Unter der Annahme einer üblichen SD der Laufstrecke im 6MWT von 65 m bei einem 2-seitigen Signifikanzniveau von 0.05, wurde geschätzt, dass etwa 162 Patienten

Item <sup>a</sup>	Charakteristikum	Studieninformation
		(54/Gruppe) ausreichen, um einen Unterschied von 40 m der mittleren Gehstrecke im 6MWT bei einer Teststärke von >90% nachweisen zu können. Zur Adjustierung der Multiplizität wurde die Hochberg-Prozedur eingesetzt. Verglichen mit der SD in einer ähnlichen Population in einer früheren Studie (45-61 m in Woche 24-48) wurde die SD von 65 m als konservativ eingestuft.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Eine formale Zwischenanalyse war nicht geplant.</p> <p>Ein unabhängiges Statistikzentrum führte etwa alle 4 Monate Analysen durch, die einem unabhängigen Datenüberwachungskomitee (DMC) zugeleitet wurden. Die Vertreter des DMCs waren unabhängig vom Sponsor und nicht an klinischen Studien mit BMN 110 beteiligt. Weder Vertreter des Statistikzentrums noch des DMCs hatten direkten Kontakt mit dem Studienpersonal oder den Patienten. Die Aufgabe des DMCs bestand hauptsächlich darin, im Falle eventuell auftretender schwerer Sicherheitsprobleme Empfehlungen hinsichtlich des weiteren Verlaufs der Studie gegenüber dem Sponsor abzugeben. Für einen Studienabbruch war ein einstimmiges Votum des DMCs erforderlich. Während der geschlossenen Sitzungen des DMCs und eines unabhängigen Statistikers durfte kein Vertreter des Sponsors anwesend sein. Die während der geschlossenen Sitzungen diskutierten Zwischenberichte waren verblindet und waren nur dem geschlossenen Kreis der DMC-Mitglieder zugänglich. Sollte ein DMC-Mitglied die Entblindung der Behandlungsarme fordern, sollte diese Frage direkt und vertraulich an den unabhängigen Statistiker gerichtet werden, der niemanden außerhalb des DMCs von dieser Frage informieren sollte. Der Vorsitzende des DMCs behielt einen Datensatz pro geschlossenem Expertentreffen.</p> <p>Ein unabhängiges ARRB war im Falle schwerer oder schwerwiegender Infusionsreaktionen verfügbar. Die Mitglieder des ARRB waren nicht direkt in die Studie involviert. Die Vertreter des ARRB kommunizierten direkt mit dem klinischen Monitor oder Mitgliedern des Studienteams.</p> <p><u>Kriterien für einen vorzeitigen Studienabbruch:</u></p> <p>Sollten zu irgendeinem Zeitpunkt 20% der Teilnehmer der Studie eine nicht akzeptable</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Toxizität hinsichtlich Überempfindlichkeitsreaktionen oder Reaktionen allergischer Art<sup>1</sup> entwickeln, sollte der Sponsor das DMC konsultieren, um das Nutzen-Risiko-Verhältnis einer Fortsetzung der Studie zu evaluieren.</p> <p>Hinsichtlich UEs anderer Art, sollte das DMC konsultiert werden, wenn zwei Patienten dasselbe schwere UE oder SUE entwickelten und das Ereignis nach Meinung des Prüfarztes mit der Behandlung in Zusammenhang stand. In diesem Fall sollte das DMC das Nutzen-Risiko-Verhältnis einer Fortsetzung der Studie bewerten.</p>
<b>8</b>	Randomisierung, Erzeugung der Behandlungsfolge	
<b>8a</b>	Methode zur Generierung der zufälligen Zuteilung	“Interactive Web Response System” (IWRS) oder “Interactive Voice Response System” (IVRS)
<b>8b</b>	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Randomisierung erfolgte stratifiziert nach 6MWT Kategorien ( $\leq 200$ m und $> 200$ m bei der Screening-Untersuchung) und Altersgruppen (5-11, 12-18 und $\geq 19$ Jahre).
<b>9</b>	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Das Randomisierungsschema wurde durch einen unabhängigen Drittanbieter erstellt, sodass die Patienten, die Prüfarzte, das Studienpersonal und die Mitarbeiter des Sponsors verblindet waren.
<b>10</b>	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	<p>Das Randomisierungsschema wurde durch unabhängige Dritte erstellt und die Patienten wurden den Behandlungen innerhalb der Studienzentren entsprechend des Randomisierungsschemas zugeteilt.</p> <p>Ein interner, nicht verblindeter Mitarbeiter des Sponsors, der an der Studie nicht direkt beteiligt war, überprüfte vor der ersten Dosierung die Gruppenzuteilung, um sicher zu stellen, dass alle Patienten die zugeteilte Behandlung erhielten.</p>
<b>11</b>	Verblindung	
<b>11a</b>	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Die Studie war doppelblind und die Prüfpräparate waren äußerlich nicht unterscheidbar. Die Patienten, Prüfarzte und das Studienpersonal waren verblindet und hatten bis zur finalen Auswertung der Studie keine Kenntnis der Gruppenzuteilung. Im Falle eines schweren oder lebensbedrohlichen UEs konnte die Verblindung des Prüfarztes nach formaler, schriftlicher Genehmigung durch einen Monitor des Sponsors gebrochen werden, was in der Studie allerdings nicht

Item <sup>a</sup>	Charakteristikum	Studieninformation
		vorgekommen ist.
<b>11b</b>	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Nicht relevant
<b>12</b>	Statistische Methoden	
<b>12a</b>	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<p><u>Primäre, supportive und sekundäre Zielparameter:</u></p> <p>ANCOVA-Modell (Mittelwert aus zwei Messungen) mit den Kovariaten Gehstrecke im 6MWT zum Ausgangspunkt und Altersklassen, Hochberg-Methode zur Adjustierung für Multiplizität (Typ I-Fehler von 0.05)</p> <p>Für sekundäre und andere Zielparameter wurde der jeweilige Baseline-Wert als zusätzliche Kovariate in das ANCOVA-Modell eingeschlossen.</p> <p>Fehlende Werte:</p> <p>Imputation des Einzelwertes, wenn nur ein 6MWT- oder 3MSC-Ergebnis pro Zeitpunkt vorlag</p> <p>Patienten, die verstarben oder physisch nicht in der Lage waren, an den Tests teilzunehmen: Imputation des Wertes 0 bei 6MWT und 3 MSC, Imputation des schlechtesten Wertes im Datensatz für respiratorische Funktion</p> <p>Alle übrigen fehlenden Werte (nicht Tod oder physisches Unvermögen): Multiple Imputation</p> <p><u>Tertiäre Zielparameter:</u></p> <p>Deskriptive Statistik</p>
<b>12b</b>	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p><u>Supportive Analyse:</u></p> <p>ANCOVA-Modell mit Messwiederholungen (nicht MVV und kombinierter Endpunkt), Kovariaten analog Primäranalyse</p> <p><u>Sensitivitätsanalysen: 6MWT, 3MSC, uKS, MVV (jeweils ANCOVA):</u></p> <p>Gepoolter Placebomittelwert</p> <p>LOCF (nicht für MVV und kombinierten Endpunkt)</p> <p>Effekt der Auswertung von Mittelwerten der replikativen Tests (6MWT und 3MSC):</p> <p>Imputation der Einzelwerte (1. bzw. 2. Test) und des jeweils besten Wertes</p> <p>Verletzung der Annahmen zur Verteilung (z.B. Normalverteilung, Ausreißer):</p> <p>Gewichtete Änderung bis Woche 24</p> <p>Effekt von Ausreißern: Analyse ohne größte Ausreißer</p> <p>Heterogenität des Behandlungseffekts:</p>

Item <sup>a</sup>	Charakteristikum	Studieninformation
		<p>Analyse mit jeweils einer zusätzlichen Kovariate pro Endpunkt (Interaktion zwischen Behandlung und Baseline 6MWT, ≤200 m, &gt;200 m, Behandlung und Baseline 3MSCT, Behandlung und Baseline uKS, Behandlung und MVV, Behandlung und kombinierter Endpunkt Baseline)</p> <p>Kontinuierliche Baseline-Werte: Austausch der kategorialen 6MWT-Stratifizierung gegen kontinuierliche 6MWT-Werte</p> <p>Evaluation der Behandlung: Behandlung als alleiniger Faktor im Modell (ANOVA)</p> <p>Zusätzliche Kovariaten</p> <p><u>Subgruppenanalysen</u></p> <p>6MWT-Kategorien (Baseline): ≤200 m, &gt;200 m</p> <p>Altersgruppen (Baseline): 5-11, 12-18, ≥19 Jahre</p> <p>Geschlecht: weiblich / männlich</p> <p>Rasse: weiß / nicht weiß</p> <p>Region: Nordamerika (Canada, USA), Europa (Dänemark, Frankreich, Deutschland, UK, Italien, Niederlande, Portugal) und Andere (Brasilien, Kolumbien, Japan, Korea, Katar, Saudi-Arabien, Taiwan)</p>
<b>Resultate</b>		
<b>13</b>	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
<b>13a</b>	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p>Placebo: a) 60 (100%), b) 59 (98.3%), c) 59 (98.3%)</p> <p>BMN 110 2.0 mg/kg KG qow: a) 59 (100%), b) 59 (100%), c) 59 (100%)</p> <p>BMN 110 2.0 mg/kg KG/Woche: a) 58 (100%), b) 57 (98.3%), c) 58 (100%)</p>
<b>13b</b>	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Placebo: n=1 Patient (1.7%) erhielt keine Infusion, da die Diagnose MPS IVA nicht bestätigt werden konnte (protokollgemäß aus der ITT-Population ausgeschlossen)</p> <p>BMN 110 2.0 mg/kg KG/Woche: n=1 Patientin (1.7%) widerrief die Einwilligungserklärung, da die Patientin sich außerstande fühlte, sich an das einmal wöchentliche Infusionsregime zu halten</p>
<b>14</b>	Aufnahme / Rekrutierung	
<b>14a</b>	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die Studie wurde zwischen dem 01.02.2011 (erste Aufnahme eines Patienten) und dem 16.08.2012 (letzte Dosis verabreicht) verabreicht. Die Teilnehmer der Studie hatten

Item <sup>a</sup>	Charakteristikum	Studieninformation
		die Möglichkeit an einer nachfolgenden offenen Studie teilzunehmen.
<b>14b</b>	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde wie geplant beendet.

**a: nach CONSORT 2010.**

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

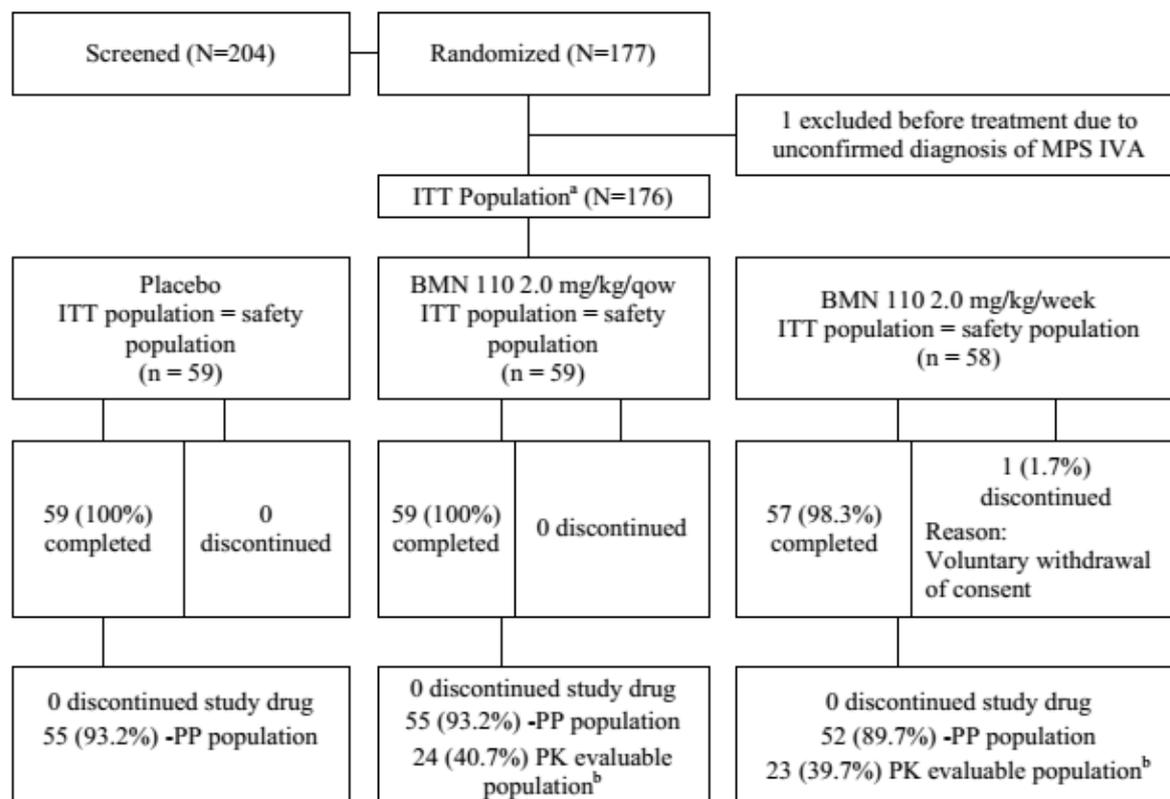


Abbildung 24: Patientenfluss in Studie MOR-004

Die Nenner basieren auf der ITT-Population. Proben zur Bestimmung der PK-Parameter wurden bei einer Subgruppe von 65 Patienten in ausgewählten Zentren entnommen.

a ITT-Population ohne bedeutende Protokollverletzungen. Ein Patient der Placebogruppe, der nicht behandelt worden war, da die Diagnose MPS IVA nicht bestätigt werden konnte, wurde aus allen Analysen ausgeschlossen.

b PK-Parameter waren bei den Placebo-Patienten nicht bestimmbar. Bei einem Patienten der BMN 110-Gruppe (2.0 mg/kg/Woche) wurde nur eine Blutprobe vor der Infusion in Woche 0 entnommen und die PK-Parameter konnten nicht bestimmt werden.

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

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

Der Bogen gliedert sich in zwei Teile:

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

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

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

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

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

Tabelle 4-86 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie MOR-004

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

Genauere Benennung der Quelle	Kürzel
MOR-004 Studienprotokoll	A
MOR-004 SAP	B
MOR-004 Studienbericht	C

**A: Verzerrungsaspekte auf Studienebene:**

**Einstufung als randomisierte Studie**

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

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

Angaben zum Kriterium:

---

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

---

1.

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

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

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

---

Das Randomisierungsschema wurde durch einen unabhängigen Dritten erstellt. Die Gruppenzuteilung erfolgte randomisiert mittels IWRS oder IVRS und die Randomisierung ist detailliert im Protokoll und SAP beschrieben (A,B).

---

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

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

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

---



---

2.

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

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

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

---

Die Gruppenzuteilung erfolgte zentral mittels IWRS oder IVRS und die Prüfpräparate waren äußerlich nicht unterscheidbar (A,B).

---

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

ja     unklar     nein

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

---

---

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

ja     unklar     nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

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

---

**behandelnde bzw. weiterbehandelnde Personen:**

ja     unklar     nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

---

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

---

**4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte**

ja     unklar     nein

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

---

Es existieren keine Hinweise auf eine ergebnisabhängige Berichterstattung; die Auswertung der Studie, die Darstellung der Endpunkte, sowie Umfang und Detailreichtum der Darstellung lassen keine Diskrepanzen zwischen dem primär geplanten Vorgehen und den im Studienbericht dargestellten Ergebnissen erkennen (A,B,C).

---

**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,B,C). Auf Empfehlungen der EMA und der FDA wurden nach Beginn der Studie einige Ergänzungen und Modifikationen hinsichtlich der Zielkriterien und deren Auswertung wie z.B. der Imputationsmethoden für fehlende Werte oder Auswertung der Zahl der erklommenen Stufen pro Minute anstelle Zahl der Stufen in drei Minuten im 3MSC-Test im SAP vorgenommen (B). Da diese Änderungen bereits vor der Entblindung der Studie implementiert wurden, hatten diese Modifikationen keinen erkennbaren Einfluss auf die Berichterstattung oder die Interpretation der Studienergebnisse.

**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: Gehstrecke im 6MWT**

**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:

Die primäre Analyse bezog sich auf die ITT-Population d.h. alle randomisierten Patienten, die mindestens eine Dosis der Prüfmedikationen erhalten hatten. Bis auf einen Patienten wurden alle randomisierten Patienten in die ITT-Analyse eingeschlossen (n=176/177, Nichtberücksichtigungsanteil: 0.6%). Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (z.B. Imputation für fehlende Werte) ist adäquat und verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung.

**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. Im Hauptteil des Studienberichts wurde der primäre Endpunkt mit den Kovariaten Baseline 6MWT und Altersgruppen berichtet, was nicht dem SAP entsprach (zusätzliche Kovariate: Behandlung). Das Ergebnis dieser Analyse entsprach dem Ergebnis der im SAP festgelegten Analyse, sodass sich daraus keine Konsequenzen für die Nutzenbewertung ergeben.

---

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

ja       nein

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

---

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

---

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

niedrig       hoch

Begründung für die Einstufung:

---

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

---

**B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt**

**Endpunkt: 6MWT Responderanalyse**

**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:

---

---

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

---

---

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

---

---

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

Begründung für die Einstufung:

---

---

**Endpunkt: Ausdauer im 3MSCT****1. Verblindung der Endpunkterheber** ja       unklar       neinAngaben zum Kriterium; obligate Begründung für die Einstufung

---

Die Erhebung und Auswertung des Endpunktes erfolgte verblindet.

---

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

---

Die primäre Analyse bezog sich auf die ITT-Population d.h. alle randomisierten Patienten, die mindestens eine Dosis der Prüfmedikationen erhalten hatten. Bis auf einen Patienten wurden alle randomisierten Patienten in die ITT-Analyse eingeschlossen (n=176/177, Nichtberücksichtigungsanteil: 0.6%). Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (z.B. Imputation für fehlende Werte) ist adäquat und verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung.

---

**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. Im Hauptteil des Studienberichts wurde der primäre Endpunkt mit den Kovariaten Baseline 6MWT und Altersgruppen berichtet, was nicht dem SAP entspricht (zusätzliche Kovariate: Behandlung). Das Ergebnis dieser Analyse entspricht dem Ergebnis der im SAP festgelegten Analyse, sodass sich daraus keine Konsequenzen für die Nutzenbewertung ergeben.

---

**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: uKS**

**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:

---

Die primäre Analyse bezog sich auf die ITT-Population d.h. alle randomisierten Patienten, die mindestens eine Dosis der Prüfmedikationen erhalten hatten. Bis auf einen Patienten wurden alle randomisierten Patienten in die ITT-Analyse eingeschlossen (n=176/177, Nichtberücksichtigungsanteil: 0.6%). Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (z.B. Imputation für fehlende Werte) ist adäquat und verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung.

---

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

---

Es existieren keine Hinweise auf eine ergebnisgesteuerte Berichterstattung dieses Endpunkts. Im Hauptteil des Studienberichts wurde der primäre Endpunkt mit den Kovariaten Baseline 6MWT und Altersgruppen berichtet, was nicht dem SAP entsprach (zusätzliche Kovariate: Behandlung). Das Ergebnis dieser Analyse entsprach dem Ergebnis der im SAP festgelegten Analyse, sodass sich daraus keine Konsequenzen für die Nutzenbewertung ergeben.

---

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

---

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

---

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

Begründung für die Einstufung:

---

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

---

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

---

Die Erhebung und Auswertung des Endpunktes erfolgte verblindet.

---

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

---

Die primäre Analyse bezog sich auf die ITT-Population d.h. alle randomisierten Patienten, die mindestens eine Dosis der Prüfmedikationen erhalten hatten. Bis auf einen Patienten wurden alle randomisierten Patienten in die ITT-Analyse eingeschlossen (n=176/177, Nichtberücksichtigungsanteil: 0.6%). Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (z.B. Imputation für fehlende Werte) ist adäquat und verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung.

---

**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. Im Hauptteil des Studienberichts wurde der primäre Endpunkt mit den Kovariaten Baseline 6MWT und Altersgruppen berichtet, was nicht dem SAP entsprach (zusätzliche Kovariate: Behandlung). Das Ergebnis dieser Analyse entsprach dem Ergebnis der im SAP festgelegten Analyse, sodass sich daraus keine Konsequenzen für die Nutzenbewertung ergeben.

---

**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: Respiratorische Parameter****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:

---

Die primäre Analyse bezog sich auf die ITT-Population d.h. alle randomisierten Patienten, die mindestens eine Dosis der Prüfmedikationen erhalten hatten. Bis auf einen Patienten wurden alle randomisierten Patienten in die ITT-Analyse eingeschlossen (n=176/177,

Nichtberücksichtigungsanteil: 0.6%). Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (z.B. Imputation für fehlende Werte) ist adäquat und verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung.

---

### 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. Im Hauptteil des Studienberichts wurde der primäre Endpunkt mit den Kovariaten Baseline 6MWT und Altersgruppen berichtet, was nicht dem SAP entspricht (zusätzliche Kovariate: Behandlung). Das Ergebnis dieser Analyse entspricht dem Ergebnis der im SAP festgelegten Analyse, sodass sich daraus keine Konsequenzen für die Nutzenbewertung ergeben.

---

### 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: Anthropometrie

#### 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:

Die primären Analysen bezogen sich auf die ITT-Population, allerdings waren diese Analysen auf die jüngeren Patienten beschränkt. Daher handelt es sich um Subgruppenanalysen.

---

### 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. Im Hauptteil des Studienberichts wurde der primäre Endpunkt mit den Kovariaten Baseline 6MWT und Altersgruppen berichtet, was nicht dem SAP entsprach (zusätzliche Kovariate: Behandlung). Das Ergebnis dieser Analyse entsprach dem Ergebnis der im SAP festgelegten Analyse, sodass sich daraus keine Konsequenzen für die Nutzenbewertung ergeben.

---

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

ja     nein

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

---

In die primären Analysen wurde die Subgruppe der jüngeren Patienten eingeschlossen.

---

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

niedrig     hoch

Begründung für die Einstufung:

---

Das Verzerrungspotential wird als hoch eingeschätzt, da es sich um Subgruppenanalysen handelt.

---

### Endpunkt: Gesamte UEs und gesamte SUEs

#### 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:

---

Die primäre Analyse bezog sich auf die ITT-Population d.h. alle randomisierten Patienten,

---

die mindestens eine Dosis der Prüfmedikationen erhalten hatten. Bis auf einen Patienten wurden alle randomisierten Patienten in die ITT-Analyse eingeschlossen (n=176/177, Nichtberücksichtigungsanteil: 0.6%). Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (z.B. Imputation für fehlende Werte) ist adäquat und verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung.

### 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. Im Hauptteil des Studienberichts wurde der primäre Endpunkt mit den Kovariaten Baseline 6MWT und Altersgruppen berichtet, was nicht dem SAP entspricht (zusätzliche Kovariate: Behandlung). Das Ergebnis dieser Analyse entsprach dem Ergebnis der im SAP festgelegten Analyse, sodass sich daraus keine Konsequenzen für die Nutzenbewertung ergeben.

### 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: Überempfindlichkeitsreaktionen und IARs

#### 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:

---

Die primäre Analyse bezog sich auf die ITT-Population d.h. alle randomisierten Patienten, die mindestens eine Dosis der Prüfmedikationen erhalten hatten. Bis auf einen Patienten wurden alle randomisierten Patienten in die ITT-Analyse eingeschlossen (n=176/177, Nichtberücksichtigungsanteil: 0.6%). Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (z.B. Imputation für fehlende Werte) ist adäquat und verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung.

---

**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. Im Hauptteil des Studienberichts wurde der primäre Endpunkt mit den Kovariaten Baseline 6MWT und Altersgruppen berichtet, was nicht dem SAP entspricht (zusätzliche Kovariate: Behandlung). Das Ergebnis dieser Analyse entsprach dem Ergebnis der im SAP festgelegten Analyse, sodass sich daraus keine Konsequenzen für die Nutzenbewertung ergeben.

---

**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 zu Verzerrung anderer Endpunkte führt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

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

---



---

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

z. B.

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

- ja**
- nein**

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

---



---

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

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

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

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

Begründung für die Einstufung:

---



---

### B: Verzerrungsaspekte auf Endpunktebene pro Endpunkt

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

**Endpunkt:** \_\_\_\_\_

#### 1. Verblindung der Endpunkterheber

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

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

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

Angaben zum Kriterium; obligate Begründung für die Einstufung

---



---

#### 2. Adäquate Umsetzung des ITT-Prinzips

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

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

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

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

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

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

---



---

### 3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

*Beachte die Hinweise zu Punkt A.4!*

**ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

**unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

**nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

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

---



---

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

*z. B.*

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

**ja**

**nein**

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

---



---

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

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

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

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

Begründung für die Einstufung:

---

---