Dokumentvorlage, Version vom 21.02.2019

Dossier zur Nutzenbewertung gemäß § 35a SGB V

Siponimod (Mayzent®)

Novartis Pharma GmbH

Modul 4 A

Patienten mit sekundär progredienter Multipler Sklerose mit Krankheitsaktivität, nachgewiesen durch Schübe oder Bildgebung der entzündlichen Aktivität

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Stand: 04.02.2020

Inhaltsverzeichnis

Seite

Tabellenverzeichnis ... **4**
Abbildungsverzeichnis ... **14**
Abkürzungsverzeichnis ... **18**
4.1 Zusammenfassung der Inhalte von Modul 4 ... 23
4.2 Methodik ... 47
4.2.1 Fragestellung ... 47
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung ... 51
4.2.3 Informationsbeschaffung ... 53
4.2.3.1 Studien des pharmazeutischen Unternehmers ... 53
4.2.3.2 Bibliografische Literaturrecherche ... 54
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken ... 55
4.2.3.4 Suche auf der Internetseite des G-BA ... 56
4.2.3.5 Selektion relevanter Studien ... 57
4.2.4 Bewertung der Aussagekraft der Nachweise ... 58
4.2.5 Informationssynthese und -analyse ... 59
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien ... 59
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien ... 60
4.2.5.3 Meta-Analysen ... 70
4.2.5.4 Sensitivitätsanalysen ... 72
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren ... 72
4.2.5.6 Indirekte Vergleiche ... 76
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen ... 79
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel ... 79
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel ... 79
4.3.1.1.1 Studien des pharmazeutischen Unternehmers ... 79
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche ... 81
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken ... 82
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA ... 84
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel ... 84
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel ... 85
4.3.1.2.1 Studiendesign und Studienpopulationen ... 85
4.3.1.2.2 Verzerrungspotenzial auf Studienebene ... 100
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien ... 101
4.3.1.3.1 Mortalität – RCT ... 102
4.3.1.3.2 Bestätigte Krankheitsschübe – RCT ... 107
4.3.1.3.3 Behinderungsprogression (EDSS) – RCT ... 115
4.3.1.3.4 Behinderungsprogression (MSFC) – RCT ... 125

4.3.1.3.5 Schweregrad der kognitiven Dysfunktion (SDMT) – RCT 137
4.3.1.3.6 Schweregrad der kognitiven Dysfunktion (BVMT-R) – RCT 143
4.3.1.3.7 Schweregrad der Beeinträchtigung des Sehvermögens (LCVA) – RCT 148
4.3.1.3.8 Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12) – RCT 152
4.3.1.3.9 Gesundheitsbezogene Lebensqualität (MSIS-29) – RCT 158
4.3.1.3.10 Gesundheitsbezogene Lebensqualität (EQ-5D) – RCT 167
4.3.1.3.11 Unerwünschte Ereignisse – RCT 177
4.3.1.3.12 Suizidalität (C-SSRS) – RCT 217
4.3.1.3.13 Läsionslast – RCT 222
4.3.1.3.14 Hirnvolumen – RCT 228
4.3.1.3.15 Subgruppenanalysen – RCT 232
4.3.1.3.15.1 Endpunkt *Bestätigte Krankheitsschübe* – RCT (Gesamtpopulation) 257
4.3.1.3.15.2 Endpunkt *Behinderungsprogression (MSFC)* – RCT (Gesamtpopulation) 258
4.3.1.3.15.3 Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* – RCT (Gesamtpopulation) 260
4.3.1.3.15.4 Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* – RCT (Gesamtpopulation) 261
4.3.1.3.15.5 Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* – RCT (Gesamtpopulation) 263
4.3.1.3.15.6 Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* – RCT (Gesamtpopulation) 273
4.3.1.3.15.7 Endpunkt *Unerwünschte Ereignisse* – RCT (Gesamtpopulation) . 276
4.3.1.3.15.8 Endpunkt *Läsionslast* – RCT (Gesamtpopulation) 290
4.3.1.3.15.9 Endpunkt *Behinderungsprogression (EDSS)* – RCT (Subpopulation B) 298
4.3.1.3.15.10 Endpunkt *Behinderungsprogression (MSFC)* – RCT (Subpopulation B) 299
4.3.1.3.15.11 Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* – RCT (Subpopulation B) 300
4.3.1.3.15.12 Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* – RCT (Subpopulation B) 301
4.3.1.3.15.13 Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* – RCT (Subpopulation B) 305
4.3.1.3.15.14 Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* – RCT (Subpopulation B) 307
4.3.1.3.15.15 Endpunkt *Unerwünschte Ereignisse* – RCT (Subpopulation B) 309
4.3.1.3.15.16 Endpunkt *Läsionslast* – RCT (Subpopulation B) 318
4.3.1.4 Liste der eingeschlossenen Studien - RCT 323
4.3.2 Weitere Unterlagen 323
4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien 323
4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche 323
4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche 324

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen ... 324
4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT ... 324
4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT ... 327
4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT ... 327
4.3.2.2 Nicht randomisierte vergleichende Studien ... 327
4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien ... 327
4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien ... 328
4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien ... 329
4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien ... 329
4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien ... 330
4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien ... 330
4.3.2.3 Weitere Untersuchungen ... 330
4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen ... 331
4.3.2.3.2 Charakteristika der weiteren Untersuchungen ... 331
4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen ... 332
4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen ... 332
4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen ... 333
4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen ... 333
4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens ... 333
4.4.1 Beurteilung der Aussagekraft der Nachweise ... 333
4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß ... 334
4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht ... 337
4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte ... 338
4.5.1 Begründung für die Vorlage indirekter Vergleiche ... 338
4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen ... 338
4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen ... 338
4.5.4 Verwendung von Surrogatendpunkten ... 339
4.6 Referenzliste ... 340
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche ... 344
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken ... 348
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche) ... 350
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken) ... 352
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT ... 363
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten ... 383

Tabellenverzeichnis

Seite

Tabelle 4-1: Ein- und Ausschlusskriterien für Studie (Fragestellung für Subpopulation A) ... 24
Tabelle 4-2: Ein- und Ausschlusskriterien für Studie (Fragestellung für Subpopulation B) ... 25
Tabelle 4-3: Zusammenfassung der Ergebnisse aus RCT (Gesamtpopulation) ... 29
Tabelle 4-4: Zusammenfassung der Ergebnisse aus den Subgruppenanalysen (Gesamtpopulation) ... 34
Tabelle 4-5: Zusammenfassung der Ergebnisse aus RCT (Subpopulation B) ... 38
Tabelle 4-6: Zusammenfassung der Ergebnisse aus den Subgruppenanalysen (Subpopulation B) ... 42
Tabelle 4-7: Bewertung des Zusatznutzens bezüglich der untersuchten Endpunkte (Gesamtpopulation und Subpopulation B) ... 43
Tabelle 4-8: Überblick zur Festlegung der zu bewertenden Subpopulationen ... 48
Tabelle 4-9: Auftitrationsschema ... 49
Tabelle 4-10: Ein- und Ausschlusskriterien für Studie (Fragestellung für Subpopulation A) ... 52
Tabelle 4-11: Ein- und Ausschlusskriterien für Studie (Fragestellung für Subpopulation B) ... 53
Tabelle 4-12: Im Dossier dargestellte patientenrelevante Endpunkte ... 61
Tabelle 4-13: Zuordnung der an der EXPAND-Studie teilnehmenden Ländern nach Regionen ... 75
Tabelle 4-14: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel ... 80
Tabelle 4-15: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel ... 80
Tabelle 4-16: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel ... 83
Tabelle 4-17: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel ... 84
Tabelle 4-18: Studienpool – RCT mit dem zu bewertenden Arzneimittel ... 85
Tabelle 4-19: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel ... 86
Tabelle 4-20: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel ... 87
Tabelle 4-21: Patientendisposition – RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 87
Tabelle 4-22: Beobachtungsdauer – RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 89

Tabelle 4-23: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Demographie, Gesamtpopulation) ... 89
Tabelle 4-24: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Erkrankungshistorie, Gesamtpopulation) ... 90
Tabelle 4-25: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (MS-Charakteristika, Gesamtpopulation) ... 92
Tabelle 4-26: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Kriterien der SPMS mit aktiver Erkrankung, Gesamtpopulation) ... 93
Tabelle 4-27: Patientendisposition – RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 93
Tabelle 4-28: Beobachtungsdauer – RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 95
Tabelle 4-29: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Demographie, Subpopulation B) ... 95
Tabelle 4-30: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Erkrankungshistorie, Subpopulation B) ... 96
Tabelle 4-31: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (MS-Charakteristika, Subpopulation B) ... 97
Tabelle 4-32: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel ... 101
Tabelle 4-33: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel ... 101
Tabelle 4-34: Operationalisierung des Endpunkts *Mortalität* ... 104
Tabelle 4-35: Bewertung des Verzerrungspotenzials für den Endpunkt *Mortalität (Todesfälle)* in RCT mit dem zu bewertenden Arzneimittel ... 105
Tabelle 4-36: Ergebnisse für den Endpunkt *Mortalität* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 106
Tabelle 4-37: Ergebnisse für den Endpunkt *Mortalität* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 106
Tabelle 4-38: Operationalisierung des Endpunkts *Bestätigte Krankheitsschübe* ... 108
Tabelle 4-39: Bewertung des Verzerrungspotenzials für den Endpunkt *Bestätigte Krankheitsschübe* in RCT mit dem zu bewertenden Arzneimittel ... 108
Tabelle 4-40: Ergebnisse für den Endpunkt *Jährliche Rate bestätigter Krankheitsschübe* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 109
Tabelle 4-41: Ergebnisse für den Endpunkt *Jährliche Rate bestätigter Krankheitsschübe* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 110
Tabelle 4-42: Ergebnisse für den Endpunkt *Zeit bis zum ersten bestätigten Krankheitsschub* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 111
Tabelle 4-43: Ergebnisse für den Endpunkt *Zeit bis zum ersten bestätigten Krankheitsschub* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 113
Tabelle 4-44: Operationalisierung des Endpunkts *Behinderungsprogression* ... 115

Tabelle 4-45: Bewertung des Verzerrungspotenzials für den Endpunkt *Behinderungsprogression* in RCT mit dem zu bewertenden Arzneimittel 116

Tabelle 4-46: Ergebnisse für den Endpunkt *Inzidenz der bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 117

Tabelle 4-47: Ergebnisse für den Endpunkt *Inzidenz der bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)..... 118

Tabelle 4-48: Ergebnisse für den Endpunkt *Zeit bis zur bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 119

Tabelle 4-49: Ergebnisse für den Endpunkt *Zeit bis zur bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)..... 122

Tabelle 4-50: Operationalisierung des Endpunkts *Behinderungsprogression (MSFC)* 125

Tabelle 4-51: Bewertung des Verzerrungspotenzials für den Endpunkt *Behinderungsprogression (MSFC)* in RCT mit dem zu bewertenden Arzneimittel..... 126

Tabelle 4-52: Rücklaufquoten MSFC-Komponenten (Gesamtpopulation) 127

Tabelle 4-53: Rücklaufquoten MSFC-Komponenten (Subpopulation B)..... 128

Tabelle 4-54: Ergebnisse für den Endpunkt *Behinderungsprogression (MSFC)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 129

Tabelle 4-55: Ergebnisse für den Endpunkt *Behinderungsprogression (MSFC)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)..... 130

Tabelle 4-56: Ergebnisse für den Endpunkt *Inzidenz einer bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 132

Tabelle 4-57: Ergebnisse für den Endpunkt *Inzidenz einer bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)..... 132

Tabelle 4-58: Ergebnisse für den Endpunkt *Zeit bis zur ersten bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 133

Tabelle 4-59: Ergebnisse für den Endpunkt *Zeit bis zur ersten bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 135

Tabelle 4-60: Operationalisierung des Endpunkts *Schweregrad der kognitiven Dysfunktion (SDMT)* 137

Tabelle 4-61: Bewertung des Verzerrungspotenzials für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* in RCT mit dem zu bewertenden Arzneimittel 138

Tabelle 4-62: Rücklaufquoten SDMT (Gesamtpopulation)..... 138

Tabelle 4-63: Rücklaufquoten SDMT (Subpopulation B) 139

Tabelle 4-64: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 139

Tabelle 4-65: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 140

Tabelle 4-66: Ergebnisse für den Endpunkt *Inzidenz der Verbesserung bzw. Verschlechterung im SDMT nach 12 Monaten* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 141

Tabelle 4-67: Ergebnisse für den Endpunkt *Inzidenz der Verbesserung bzw. Verschlechterung im SDMT nach 12 Monaten* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)..... 141

Tabelle 4-68: Operationalisierung des Endpunkts *Schweregrad der kognitiven Dysfunktion (BVMT-R)* 143

Tabelle 4-69: Bewertung des Verzerrungspotenzials für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* in RCT mit dem zu bewertenden Arzneimittel 143

Tabelle 4-70: Rücklaufquoten BVMT-R (Gesamtpopulation) 144

Tabelle 4-71: Rücklaufquoten BVMT-R (Subpopulation B)..... 145

Tabelle 4-72: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 145

Tabelle 4-73: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)..... 146

Tabelle 4-74: Operationalisierung des Endpunkts *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)*..... 148

Tabelle 4-75: Bewertung des Verzerrungspotenzials für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* in RCT mit dem zu bewertenden Arzneimittel..... 148

Tabelle 4-76: Rücklaufquoten LCVA (Gesamtpopulation) 149

Tabelle 4-77: Rücklaufquoten LCVA (Subpopulation B)..... 149

Tabelle 4-78: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 150

Tabelle 4-79: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 150

Tabelle 4-80: Operationalisierung des Endpunkts *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)*..... 152

Tabelle 4-81: Bewertung des Verzerrungspotenzials für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* in RCT mit dem zu bewertenden Arzneimittel..... 152

Tabelle 4-82: Rücklaufquoten MSWS-12 (Gesamtpopulation)..... 153

Tabelle 4-83: Rücklaufquoten MSWS-12 (Subpopulation B) 153

Tabelle 4-84: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 154

Tabelle 4-85: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 155
Tabelle 4-86: Operationalisierung des Endpunkts *Gesundheitsbezogene Lebensqualität (MSIS-29)* ... 158
Tabelle 4-87: Bewertung des Verzerrungspotenzials für *Gesundheitsbezogene Lebensqualität (MSIS-29)* in RCT mit dem zu bewertenden Arzneimittel ... 158
Tabelle 4-88: Rücklaufquoten MSIS-29 (Gesamtpopulation) ... 159
Tabelle 4-89: Rücklaufquoten MSIS-29 (Subpopulation B) ... 160
Tabelle 4-90: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 161
Tabelle 4-91: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 164
Tabelle 4-92: Operationalisierung des Endpunkts *Gesundheitsbezogene Lebensqualität (EQ-5D)* ... 167
Tabelle 4-93: Bewertung des Verzerrungspotenzials für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* in RCT mit dem zu bewertenden Arzneimittel ... 168
Tabelle 4-94: Rücklaufquoten EQ-5D-Komponenten (Gesamtpopulation) ... 169
Tabelle 4-95: Rücklaufquoten EQ-5D-Komponenten (Subpopulation B) ... 170
Tabelle 4-96: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 171
Tabelle 4-97: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 174
Tabelle 4-98: Operationalisierung des Endpunkts *Unerwünschte Ereignisse* ... 177
Tabelle 4-99: Bewertung des Verzerrungspotenzials für den Endpunkt *Unerwünschte Ereignisse* in RCT mit dem zu bewertenden Arzneimittel ... 179
Tabelle 4-100: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 180
Tabelle 4-101: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 181
Tabelle 4-102: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse nach Schweregrad* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 182
Tabelle 4-103: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse nach Schweregrad* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 183
Tabelle 4-104: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse von besonderem Interesse nach Schweregrad* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) ... 185
Tabelle 4-105: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse von besonderem Interesse nach Schweregrad* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) ... 189

Tabelle 4-106: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse (nach SOC und PT)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 194
Tabelle 4-107: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse (nach SOC und PT)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 196
Tabelle 4-108: Ergebnisse für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 197
Tabelle 4-109: Ergebnisse für den Endpunkt *Zeit bis zum Unerwünschten Ereignis* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 207
Tabelle 4-110: Operationalisierung des Endpunkts *Suizidalität (C-SSRS)* 217
Tabelle 4-111: Bewertung des Verzerrungspotenzials für den Endpunkt *Suizidalität (C-SSRS)* in RCT mit dem zu bewertenden Arzneimittel 218
Tabelle 4-112: Ergebnisse für den Endpunkt *Suizidalität (C-SSRS)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 219
Tabelle 4-113: Ergebnisse für den Endpunkt *Suizidalität (C-SSRS)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 220
Tabelle 4-114: Operationalisierung des Endpunkts *Läsionslast* 222
Tabelle 4-115: Bewertung des Verzerrungspotenzials für den Endpunkt *Läsionslast* in RCT mit dem zu bewertenden Arzneimittel 222
Tabelle 4-116: Rücklaufquoten *Läsionslast* (Gesamtpopulation) 223
Tabelle 4-117: Rücklaufquoten *Läsionslast* (Subpopulation B) 224
Tabelle 4-118: Ergebnisse für den Endpunkt *Läsionslast (T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 224
Tabelle 4-119: Ergebnisse für den Endpunkt *Läsionslast (T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 225
Tabelle 4-120: Ergebnisse für den Endpunkt *Läsionslast (Anzahl neuer bzw. sich vergrößernder T2-Läsionen und Gd-anreichernder T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 226
Tabelle 4-121: Ergebnisse für den Endpunkt *Läsionslast (Anzahl neuer bzw. sich vergrößernder T2-Läsionen und Gd-anreichernder T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 227
Tabelle 4-122: Operationalisierung des Endpunkts *Hirnvolumen* 228
Tabelle 4-123: Bewertung des Verzerrungspotenzials für den Endpunkt *Läsionslast* in RCT mit dem zu bewertenden Arzneimittel 229
Tabelle 4-124: Rücklaufquoten *Hirnvolumen* (Gesamtpopulation) 229
Tabelle 4-125: Rücklaufquoten *Hirnvolumen* (Subpopulation B) 230
Tabelle 4-126: Ergebnisse für den Endpunkt *Hirnvolumen* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 230
Tabelle 4-127: Ergebnisse für den Endpunkt *Hirnvolumen* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 231
Tabelle 4 -128 Matrix der durchgeführten Subgruppenanalysen 233

Tabelle 4-129: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für die Studie EXPAND (Gesamtpopulation) 235
Tabelle 4-130: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für die Studie EXPAND (Subpopulation B) 244
Tabelle 4-131: Zusammenfassung der Ergebnisse aus den Subgruppenanalysen (Gesamtpopulation) 253
Tabelle 4-132: Zusammenfassung der Ergebnisse aus den Subgruppenanalysen (Subpopulation B) 255
Tabelle 4-133: Ergebnisse für den Endpunkt *Bestätigte Krankheitsschübe (jährliche Rate bestätigter Schübe)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation) 257
Tabelle 4-134: Ergebnisse für den Endpunkt *Behinderungsprogression (MSFC)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation) 258
Tabelle 4-135: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation) 260
Tabelle 4-136: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation) 261
Tabelle 4-137: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation) 262
Tabelle 4-138: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* aus RCT mit dem zu bewertenden Arzneimittel nach Alter (Gesamtpopulation) 263
Tabelle 4-139: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* aus RCT mit dem zu bewertenden Arzneimittel nach Erkrankungsverlauf (Gesamtpopulation) 274
Tabelle 4-140: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation) 277
Tabelle 4-141: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Alter (Gesamtpopulation) 278
Tabelle 4-142: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach schnell fortschreitender Erkrankung (Gesamtpopulation) 283
Tabelle 4-143: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation) 287
Tabelle 4-144: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation) 288
Tabelle 4-145: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Gesamtpopulation) 289
Tabelle 4-146: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Erkrankungsverlauf (Gesamtpopulation) 290

Tabelle 4-147: Ergebnisse für den Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel nach Alter (Gesamtpopulation) ... 291

Tabelle 4-148: Ergebnisse für den Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation) ... 292

Tabelle 4-149: Ergebnisse für den Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel nach Erkrankungsverlauf (Gesamtpopulation) ... 293

Tabelle 4-150: Ergebnisse für den Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Gesamtpopulation) ... 294

Tabelle 4-151: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation) ... 295

Tabelle 4-152: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Gesamtpopulation) ... 296

Tabelle 4-153: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Gesamtpopulation) ... 296

Tabelle 4-154: Ergebnisse für den Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Gesamtpopulation) ... 297

Tabelle 4-155: Ergebnisse für den Endpunkt *Behinderungsprogression (EDSS)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B) ... 298

Tabelle 4-156: Ergebnisse für den Endpunkt *Behinderungsprogression (MSFC)* aus RCT mit dem zu bewertenden Arzneimittel bei schnell fortschreitender Erkrankung (Subpopulation B) ... 299

Tabelle 4-157: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B) ... 300

Tabelle 4-158: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B) ... 301

Tabelle 4-159: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel nach Alter (Subpopulation B) ... 303

Tabelle 4-160: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B) ... 304

Tabelle 4-161: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Subpopulation B) ... 305

Tabelle 4-162: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Subpopulation B) 306
Tabelle 4-163: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* aus RCT mit dem zu bewertenden Arzneimittel nach Erkrankungsverlauf (Subpopulation B) 307
Tabelle 4-164: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B) 310
Tabelle 4-165: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B) 311
Tabelle 4-166: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B) 314
Tabelle 4-167: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B) 317
Tabelle 4-168: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Subpopulation B) 317
Tabelle 4-169: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Subpopulation B) 319
Tabelle 4-170: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Subpopulation B) 319
Tabelle 4-171: Ergebnisse für den Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Subpopulation B) 321
Tabelle 4-172: Ergebnisse für den Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Subpopulation B) 322
Tabelle 4-173: Ergebnisse für den Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Subpopulation B) 322
Tabelle 4-174: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche 324
Tabelle 4-175: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden 325
Tabelle 4-176: Operationalisierung von <Endpunkt xxx> 325
Tabelle 4-177: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche 326
Tabelle 4-178: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche 326
Tabelle 4-179: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien 328
Tabelle 4-180: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien 329

Tabelle 4-181: Operationalisierung von <Endpunkt xxx> 329
Tabelle 4-182: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien 329
Tabelle 4-183: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen ... 332
Tabelle 4-184: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen 332
Tabelle 4-185: Bewertung des Zusatznutzens bezüglich der untersuchten Endpunkte (Gesamtpopulation und Subpopulation B) 334
Tabelle 4-186: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens 338
Tabelle 4-187 (Anhang): Studiendesign und -methodik für Studie EXPAND 364
Tabelle 4-190 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie EXPAND 384

Abbildungsverzeichnis

Seite

Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel 82

Abbildung 2: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten bestätigten Krankheitsschub* (Studienende) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 112

Abbildung 3: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten bestätigten Krankheitsschub* (Studienende) aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 114

Abbildung 4: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur nach 3 Monaten bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 120

Abbildung 5: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 121

Abbildung 6: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur nach 3 Monaten bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 123

Abbildung 7: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 124

Abbildung 8: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur ersten nach 3 Monaten bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 134

Abbildung 9: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur ersten nach 3 Monaten bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 136

Abbildung 10: Zeitverlauf für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 155

Abbildung 11: Zeitverlauf für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)..... 156

Abbildung 12: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 162

Abbildung 13: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)..... 163

Abbildung 14: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 165

Abbildung 15: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 166

Abbildung 16: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-Indexwert)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 172

Abbildung 17: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-VAS)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 173

Abbildung 18: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-Indexwert)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 175

Abbildung 19: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-VAS)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 176

Abbildung 20: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 200

Abbildung 21: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 200

Abbildung 22: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 201

Abbildung 23: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 202

Abbildung 24: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten schweren UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 203

Abbildung 24: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten schweren UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 204

Abbildung 25: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum Studienabbruch aufgrund von UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 205

Abbildung 26: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum Behandlungsabbruch aufgrund von UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation) 206

Abbildung 27: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 209

Abbildung 28: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 210

Abbildung 29: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 211

Abbildung 30: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 212

Abbildung 31: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten schweren UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 213

Abbildung 31: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten schweren UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 214

Abbildung 32: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum Studienabbruch aufgrund von UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 215

Abbildung 33: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum Behandlungsabbruch aufgrund von UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B) 216

Abbildung 36: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 18 – 30 Jahre) 266

Abbildung 37: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 31 – 40 Jahre) 267

Abbildung 38: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 41 – 55 Jahre) 268

Abbildung 39: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter > 55 Jahre) 269

Abbildung 40: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 18 – 30 Jahre) 270

Abbildung 41: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 31 – 40 Jahre) 271

Abbildung 42: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 41 – 55 Jahre) 272

Abbildung 43: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter > 55 Jahre) 273

Abbildung 44: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-VAS)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – moderater/schwerer Erkrankungsverlauf: Nein) 275

Abbildung 45: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-VAS)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – moderater/schwerer Erkrankungsverlauf: Ja) 276

Abbildung 46: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 18 bis 30 Jahre) 279

Abbildung 47: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 31 bis 40 Jahre) 280

Abbildung 48: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 41 bis 55 Jahre) 281

Abbildung 49: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter > 55 Jahre) 282

Abbildung 50: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Schnell fortschreitende Erkrankung: Nein) 284

Abbildung 51: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Schnell fortschreitende Erkrankung: Ja) 285

Abbildung 52: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-Index)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Schnell fortschreitende Erkrankung: Nein) 308

Abbildung 53: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-Index)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Schnell fortschreitende Erkrankung: Ja) 309

Abbildung 54: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Männliche Patienten) 312

Abbildung 55: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Weibliche Patientinnen) 313

Abbildung 56: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Männliche Patienten) 315

Abbildung 57: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Weibliche Patientinnen) 316

Abkürzungsverzeichnis

Abkürzung	Bedeutung
9-HPT	9 Hole Peg Test
AESI	Adverse Events of Special Interest [*Unerwünschte Ereignisse von besonderem Interesse*]
AIDS	Acquired Immune Deficiency Syndrome [*Erworbenes Immunschwächesyndrom*]
ALP	Alkalische Phosphatase
ALT	Alanin-Aminotransferase
AMIS	Arzneimittel-Informationssystem
ARR	Annualized Relapse Rate [*Jährliche Schubrate*]
AST	Aspartat-Aminotransferase
AV-Block	Atrioventrikulärer Block
BAF312	Siponimod
BMI	Body-Mass-Index
bpm	beats per minute [*Schläge pro Minute*]
BSC	Best Supportive Care [*Bestmögliche unterstützende Behandlung*]
BVMT-R	Brief Visuospatial Memory Test – Revised
C	Celsius
CDP	Confirmed Disability Progression [*Bestätigte Behinderungsprogression*]
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CONSORT	Consolidated Standards of Reporting Trials
COPD	Chronic Obstructive Pulmonary Disease [*Chronisch obstruktive Lungenerkrankung*]
CRO	Contract Research Organisation
CSR	Clinical Study Report [*Studienbericht*]
C-SSRS	Columbia Suicide Severity Rating Scale
CTCAE	Common Terminology Criteria for Adverse Events [*Allgemeine Terminologie und Merkmale unerwünschter Ereignisse*]
CYP	Cytochrome P450
DIMDI	Deutsches Institut für Medizinische Dokumentation
DMC	Data Monitoring Committee [*Datenüberwachungsausschuss*]
DMT	Disease Modifying Therapy [*Krankheitsmodifizierende Therapie*]

eCRF	electronic Case Report Form [*elektronischer Prüfbogen*]
eC-SSRS	electronic Columbia Suicide Severity Rating Scale
EDSS	Expanded Disability Status Scale
EKG	Elektrokardiogramm
EMBASE	Excerpta Medica Database
EOS	End of Study [*Ende der Studie*]
EQ-5D	EuroQol – 5 Dimensionen
EU	Europäische Union
EU-CTR	EU Clinical Trial Register
EudraCT	European Clinical Trials Database
FAS	Full Analysis Set
FS	Funktionelles System
G-BA	Gemeinsamer Bundesausschuss
GCP	Good Clinical Practice
Gd	Gadolinium
GGT	Gamma-Glutamyltransferase
HR	Hazard Ratio
ICTRP	International Clinical Trials Registry Platform
IFN	Interferon
IgG	Immunglobulin G
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRS	Interactive Response System
IRT	Interactive Response Technology
ITT	Intention to Treat
k. A.	keine Angabe
Kat.	Kategorie
KI	Konfidenzintervall
KM-Schätzer	Kaplan-Meier-Schätzer
LCVA	Low Contrast Visual Acuity
LOCF	Last observation carried forward
MedDRA	Medical Dictionary for Regulatory Activities
MEDLINE	Medical Literature Analysis and Retrieval System Online
MMRM	Mixed Model for Repeated Measures

MRT	Magnetresonanztomographie
MS	Multiple Sklerose
MSFC	Multiple Sclerosis Functional Composite
MSIS	Multiple Sclerosis Impact Scale
MSSS	Multiple Sclerosis Severity Score
MSWS	Multiple Sclerosis Walking Scale
MTC	Mixed Treatment Comparison
MW	Mittelwert
N	Patienten gesamt
n	Patienten in der Kategorie
N'	Anzahl der in die Auswertung eingegangenen Patienten
n. b.	nicht bestimmbar
NCT	National Clinical Trial
n. d.	nicht durchgeführt
NESC	Neurostatus eScoring
NYHA	New York Heart Association
n. v.	nicht verfügbar
OR	Odds Ratio
PASAT-3	Paced Auditory Serial Addition Test
PID	Patientenidentifikation
PK	Phamakokinetik
PPD	Purified Protein Derivative
PT	Preferred Term nach MedDRA
QoL	Quality of Life [*Lebensqualität*]
QTcF	Corrected QT interval by Fredericia
RAN	Randomisierte Patienten
RCT	Randomized Controlled Trial *[Randomisierte, kontrollierte Studie]*
RD	Risikodifferenz
RMP	Risk Management Plan
RR	Relatives Risiko
RRMS	Relapsing Remitting Multiple Sclerosis [*Schubförmig remittierende Multiple Sklerose*]
SAF	Safety Set

SD	Standard deviation [*Standardabweichung*]
SDMT	Symbol Digit Modalities Test
SE	Standard error [*Standardfehler*]
SGB	Sozialgesetzbuch
SMQs	Standardised MedDRA Queries
SOC	System Organ Class nach MedDRA [*Systemorganklasse*]
SPMS	Sekundär progrediente Multiple Sklerose
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegende Unerwünschte Ereignisse
SUSAR	Suspected Unexpected Serious Reaction
S1P	Sphingosin-1-phosphat
TIA	Transitorische ischämische Attacke
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
T25W	Timed 25 Foot Walking Test
UE	Unerwünschtes Ereignis
ULN	Upper Limit of Normal [*Oberer Normwert*]
UN	United Nations [*Vereinte Nationen*]
VAS	Visuelle Analogskala
VZV	Varizella-Zoster-Virus
WBC	White Blood Cell [*Leukozyten*]
WHO	World Health Organization
WT	Wildtyp

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

– Zusammenfassung (Abschnitt 4.1)

– Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)

– Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)

– eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)

– ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

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

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

4.1 Zusammenfassung der Inhalte von Modul 4

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

Fragestellung

Mit dem vorliegenden Dossier wird der medizinische Zusatznutzen von Siponimod (Mayzent®) bei erwachsenen Patienten mit sekundär progredienter Multipler Sklerose (SPMS) mit Krankheitsaktivität, nachgewiesen durch Schübe oder Bildgebung der entzündlichen Aktivität, gemäß Zulassung bewertet.

Der Gemeinsame Bundesausschuss (G-BA) hat zwei Patientenpopulationen definiert:

- Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, mit aufgesetzten Schüben (Subpopulation A).

- Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe (Subpopulation B).

Datenquellen

Zur Beantwortung der Fragestellung für die Subpopulation A liegt keine Evidenz zum Vergleich von Siponimod mit den festgelegten zweckmäßigen Vergleichstherapien Interferon- (IFN-) beta 1a oder 1b oder Ocrelizumab vor.

Die Beantwortung der Fragestellung erfolgt für die Subpopulation B im Vergleich zur zweckmäßigen Vergleichstherapie Best Supportive Care (BSC) auf der Grundlage der Evidenz der EXPAND-Studie (NCT01665144/CBAF312A2304).

Es wurden zusätzlich eine bibliographische Literaturrecherche, eine umfassende Suche in öffentlichen Studienregistern und eine Suche auf der Internetseite des G-BA vorgenommen. Es wurde keine zusätzliche relevante Studie identifiziert.

Ein-/Ausschlusskriterien für Studien

Studien wurden nach folgenden Kriterien für die Bewertung des medizinischen Nutzens und Zusatznutzens ein- bzw. ausgeschlossen:

Tabelle 4-1: Ein- und Ausschlusskriterien für Studie (Fragestellung für Subpopulation A)

	Einschlusskriterien	**Ausschlusskriterien**
Patientenpopulation	Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, mit aufgesetzten Schüben.	Andere Populationen.
Intervention	Siponimod 2 mg (bzw. 1 mg bei Patienten mit CYP2C9-Polymorphismus), 1 × täglich oral.	Andere Interventionen.
Vergleichstherapie	IFN-beta 1a oder 1b oder Ocrelizumab.	Andere oder keine Vergleichstherapie.
Endpunkte	Mindestens einer der patientenrelevanten Endpunkte wird berichtet: • Mortalität • Morbidität (z. B. Krankheitsschübe, Behinderungsprogression, Schweregrad der Behinderung) • Gesundheitsbezogene Lebensqualität • Sicherheit und Verträglichkeit (z. B. Unerwünschte Ereignisse)	Studien ohne Ergebnisse zu mindestens einem der genannten relevanten Endpunkte.
Studientypen	RCT, Meta-Analysen von RCT.	Andere Studientypen.
Studiendauer	Zwölf Monate und länger.	Studien mit kürzerer Behandlungsdauer.
Dokumentationstyp	Studienbericht, Volltextpublikation, ausführlicher Ergebnisbericht aus einem Studienregister / einer Studienergebnisdatenbank.	Ausschließlich andere Dokumenttypen als Studienbericht und Volltextpublikation (z. B. Konferenzabstracts, Editorials, Letters to the editor).

Tabelle 4-2: Ein- und Ausschlusskriterien für Studie (Fragestellung für Subpopulation B)

	Einschlusskriterien	**Ausschlusskriterien**
Patientenpopulation	Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe.	Andere Populationen
Intervention	Siponimod 2 mg (bzw. 1 mg bei Patienten mit CYP2C9-Polymorphismus), 1 × täglich oral.	Andere Interventionen.
Vergleichstherapie	BSC	Andere oder keine Vergleichstherapien.
Endpunkte	Mindestens einer der patientenrelevanten Endpunkte wird berichtet: • Mortalität • Morbidität (z. B. Krankheitsschübe, Behinderungsprogression, Schweregrad der Behinderung) • Gesundheitsbezogene Lebensqualität • Sicherheit und Verträglichkeit (z. B. Unerwünschte Ereignisse)	Studien ohne Ergebnisse zu mindestens einem der genannten relevanten Endpunkte.
Studientypen	RCT, Meta-Analysen von RCT.	Andere Studientypen.
Studiendauer	Zwölf Monate und länger.	Studien mit kürzerer Behandlungsdauer.
Dokumentationstyp	Studienbericht, Volltextpublikation, ausführlicher Ergebnisbericht aus einem Studienregister / einer Studienergebnisdatenbank.	Ausschließlich andere Dokumenttypen als Studienbericht und Volltextpublikation (z. B. Konferenzabstracts, Editorials, Letters to the editor).

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

Das Verzerrungspotenzial der für die Bewertung identifizierten und herangezogenen Studie wurde anhand der endpunktübergreifenden und endpunktspezifischen Aspekte beurteilt. Details zum Verzerrungspotenzial auf Studien- und Endpunktebene sind dem Bewertungsbogen in Anhang 4-F zu entnehmen.

Die Beschreibung des Designs und der Methodik der randomisierten kontrollierten Studie (Randomized Controlled Trial; RCT) erfolgte anhand der Kriterien des CONSORT-Statements. Die Informationen zu den Items wurden, soweit möglich, dem Studienbericht und dem Studienprotokoll entnommen (siehe Anhang 4-E).

Für die relevante Studie EXPAND wurden folgende Patientencharakteristika zu Studienbeginn dargestellt, soweit diese Informationen verfügbar waren:

- Alter
- Geschlecht
- Gewicht
- Körpergröße
- Body-Mass-Index (BMI)
- Rasse
- Ethnie
- Erkrankungsdauer seit Diagnose
- Erkrankungsdauer seit Auftreten der ersten Symptome
- Zeit seit Konversion zur SPMS
- Anzahl der Schübe in den letzten 12 bzw. 24 Monaten
- Schweregrad der Behinderung (EDSS)
- Anzahl der Gd-anreichernden T1-Läsionen
- T2-Läsionsvolumen
- Hirnvolumen

Aus der Auswertung der direkt-vergleichenden Studie EXPAND werden Ergebnisse zur Gesamtpopulation sowie zur Subpopulation B für die patientenrelevanten Endpunkte Mortalität (Todesfälle), Morbidität (Krankheitsschübe, Behinderungsprogression, Schweregrad der kognitiven Dysfunktion, Schweregrad der Beeinträchtigung des Sehvermögens, Schweregrad der Beeinträchtigung der Gehfähigkeit, Unerwünschte Ereignisse, Suizidalität) und Gesundheitsbezogene Lebensqualität dargestellt. Ergänzend werden Ergebnisse aus der MRT-Bildgebung zur Krankheitsaktivität präsentiert.

Für das vorliegende Dossier wurden alle Endpunkte aus dem direkten Vergleich in den folgenden Subgruppen analysiert:

Präspezifizierte Subgruppen:
- Geschlecht (männlich; weiblich),
- Vorherige krankheitsmodifizierende Therapie (ja; nein),
- Vorherige Behandlung mit IFN-beta 1b (ja; nein),
- Schnell fortschreitende Erkrankung, d. h. EDSS-Anstieg um mindestens 1,5 Punkte in den zwei Jahren vor oder zu Studienbeginn (ja; nein),
- Moderater oder schwerer Erkrankungsverlauf (ja; nein), d. h. Wert von ≥ 4 auf der MS Status Scale (ja; nein).

Nachträglich definierte Subgruppen:
- Alter (18 – 30 Jahre; 31 – 40 Jahre bzw. 18 – 40 Jahre; 41 – 55 Jahre; > 55 Jahre),
- Region (Europa; andere).

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

In die Bewertung des Zusatznutzens im direkten Vergleich wurde die RCT EXPAND (NCT01665144/CBAF312A2304; Evidenzstufe Ib) eingeschlossen. In der EXPAND-Studie wurden patientenrelevante Endpunkte zur Mortalität, Morbidität, Gesundheitsbezogenen Lebensqualität sowie zur Sicherheit und Verträglichkeit erhoben. Die für die Bewertung herangezogenen Endpunkte sind valide.

Die EXPAND-Studie ist für die zu bewertende Indikation und die relevante Patientengruppe aufgrund der Population, Fallzahl, Studiendauer, Art der Intervention und der Wahl der Endpunkte relevant. Die Auswertung der entsprechenden Teilpopulation der Studie erlaubt eine valide Beurteilung des Nutzens und Zusatznutzens für die Subpopulation B (Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe).

Das Verzerrungspotenzial war studienspezifisch niedrig, die Studienqualität somit hoch. Das Verzerrungspotenzial war auf Endpunktebene ebenfalls niedrig. Ausnahmen hiervon bilden zumeist die Auswertungen zu Monat 24, da hier die Nichtberücksichtigungsanteile hoch waren. Grund für die hohen Nichtberücksichtigungsanteile war das ereignisgetriebene Design der Studie. Die Kernphase wurde etwa drei Jahre nach Beginn der Randomisierung beendet, da angenommen wurde, dass zu diesem Zeitpunkt die für die Auswertung der Behinderungsprogression notwendigen 374 Ereignisse eingetreten waren und die Randomisierung für mehr als 95 % der Patienten mindestens ein Jahr zurücklag. Die mediane Beobachtungsdauer lag bei Ende der Kernphase bei 639 Tagen. 88,1 % der Patienten wiesen eine Beobachtungsdauer von mindestens 365 Tagen auf. Die Auswertungen zu Monat 12 bzw. über die gesamte Studiendauer lassen demnach eine valide Bewertung des Zusatznutzens zu.

Die Aussagekraft der RCT ist somit sehr hoch und damit für eine Beurteilung des medizinischen Zusatznutzens uneingeschränkt geeignet.

Im vorliegenden Modul 4A wird der medizinische Zusatznutzen von Siponimod im Vergleich zu BSC in der Subpopulation B (Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe) dargestellt. Zusätzlich erfolgt die Darstellung der Ergebnisse für die Gesamtpopulation der Studie. Für die Subpopulation A (Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, mit aufgesetzten Schübe) liegen keine bewertungsrelevanten Daten zum Vergleich mit der zweckmäßigen Vergleichstherapie vor.

Auswertung der Gesamtpopulation

Statistisch signifikante Unterschiede zwischen Siponimod und BSC zeigten sich in den folgenden Endpunkten (Tabelle 4-3):

- Bestätigte Krankheitsschübe:
  Unter Therapie mit Siponimod traten statistisch signifikant weniger bestätigte

Krankheitsschübe pro Jahr auf als unter BSC (Rate Ratio 0,45; p < 0,001). Außerdem war das Schubrisiko unter Siponimod um 46 % reduziert (HR 0,54; p < 0,001).

- Behinderungsprogression gemäß EDSS:
  Das Risiko für eine nach drei bzw. sechs Monaten bestätigte Behinderungsprogression war unter Siponimod um 21 % (3-CDP) bzw. 26 % (6-CDP) niedriger (HR 0,79; p = 0,013 bzw. HR 0,74; p = 0,006). Dies entspricht einem relativen Risiko von 0,83 (p = 0,021) für Siponimod bzw. 0,78 (p = 0,010).

- Schweregrad der kognitiven Dysfunktion (SDMT):
  In der Auswertung der Veränderung des SDMT (Symbol Digit Modalities Test) seit Studienbeginn zeigte sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten von Siponimod zu Monat 12 (MW-Differenz 1,09; p = 0,013). Dabei verschlechterte sich der SDMT-Wert unter BSC, unter Siponimod hingegen blieb der Wert weitgehend stabil. Insgesamt konnten unter Siponimod 1,3-mal mehr Patienten eine klinisch relevante Verbesserung im SDMT zu Monat 12 erreichen (RR 1,30; p = 0,003), der Anteil der Patienten mit einer klinsich relevanten Verschlechterung war zu diesem Zeitpunkt unter Siponimod statistisch signifikant niedriger (RR 0,78; p = 0,002). Die Auswertung der Veränderung zu Monat 24 zeigte zwar ebenfalls einen statistisch signifikanten Unterschied, ist wegen hoher Nichtberücksichtigungsanteile aber nicht zuverlässig interpretierbar.

- Gesundheitsbezogene Lebensqualität (MSIS-29):
  Die Auswertungen der MSIS-29 (Multiple Sclerosis Impact Scale) zeigten einen statistisch signifikanten Vorteil von Siponimod in der Physical Impact Scale zu Monat 12. Unter Siponimod stieg dabei die Beeinträchtigung durch die Erkrankung weniger stark an als unter BSC (MW-Differenz -2,89; p = 0,003).

- Gesundheitsbezogene Lebensqualität (EQ-5D):
  Die Auswertungen des EQ-5D zeigten einen statistisch signifikanten Vorteil von Siponimod im Indexwert zu Monat 12. Unter Siponimod verschlechterte sich der Gesundheitszustand weniger stark als unter BSC (MW-Differenz 0,03; p = 0,037).

- Unerwünschte Ereignisse:
  Unerwünschte Ereignisse (UE) waren meist nur von leichter bis moderater Ausprägung (CTCAE-Grad < 3). Sie traten insgesamt unter Siponimod statistisch signifikant häufiger auf, der Gruppenunterschied war aber gering (RR 1,09; p < 0,001). Bei Nichtberücksichtigung krankheitsspezifischer UE war die Häufigkeit von UE und schwerwiegenden UE (SUE) insgesamt unter Siponimod höher als unter BSC (RR 1,09; p = 0,001 bzw. RR 1,37; p = 0,038). Geringgradige UE (CTCAE-Grad < 3) traten unter Siponimod geringfügig häufiger auf (RR 1,09; p < 0,001). Schwere UE (CTCAE-Grad ≥ 3) traten nur bei rund 10 % der Patienten insgesamt auf und die Häufigkeit war hierbei zwischen den Behandlungsgruppen nicht statistisch signifikant unterschiedlich. In der Betrachtung der SUE nach System Organ Class (SOC) und Preferred Term (PT) zeigte sich lediglich in der SOC ‚Untersuchung' ein statistisch signifikanter Unterschied

zwischen den Behandlungsgruppen zugunsten von BSC. Außerdem waren Behandlungsabbrüche aufgrund von UE unter Siponimod nicht statistisch signifikant häufiger als unter BSC. Insgesamt wird daher anhand der Gesamtraten der unerwünschten Ereignisse nicht von einem höheren Schadenspotenzial von Siponimod ausgegangen. Unter den UE von besonderem Interesse, die auch in der Fachinformation bereits abgebildet sind, traten unter einer Therapie mit Siponimod Varizella-Zoster-Infektionen, Makulaödeme, erhöhte Leberenzymwerte, Hypertonie, Bradykardie, Krampfanfälle und periphere Ödeme bzw. Schwellungen statistisch signifikant häufiger als unter BSC auf. Dabei handelte es sich überwiegend um leichte und moderate UE (CTCAE-Grad < 3). Nur in wenigen Fällen wurden diese Ereignisse als SUE klassifiziert, lediglich bei den Krampfanfällen handelte es sich in etwa bei der Hälfte der Fälle um SUE.

- Läsionslast:
  Unter Siponimod nahm das T2-Läsionsvolumen weniger zu als unter BSC. Der Unterschied zwischen den Behandlungsarmen war zu Monat 12 statistisch signifikant (Monat 12: MW-Differenz -613,09 mm^3; $p < 0,001$). Unter Siponimod bildeten sich außerdem weniger neue bzw. sich vergrößernde T2-Läsionen (Monat 12: Rate Ratio 0,27; $p < 0,001$) und Gd-anreichernde Läsionen (Monat 12: Rate Ratio 0,13; $p < 0,001$). Die Auswertungen zu Monat 24 sind wegen hoher Nichtberücksichtigungsanteile nicht zuverlässig interpretierbar.

- Hirnvolumen:
  Die Auswertung der prozentualen Veränderung des Hirnvolumens zeigte einen Vorteil von Siponimod. Unter Siponimod verringerte sich das Hirnvolumen statistisch signifikant weniger als unter BSC (Monat 12: MW-Differenz 0,17 %; $p < 0,001$). Der Unterschied zwischen den Behandlungsarmen war zu beiden Auswertungszeitpunkten statistisch signifikant. Die Auswertung zu Monat 24 ist wegen hoher Nichtberücksichtigungsanteile nicht zuverlässig interpretierbar.

Tabelle 4-3: Zusammenfassung der Ergebnisse aus RCT (Gesamtpopulation)

Endpunkte	**Siponimod vs. BSC**	
	Effektschätzer [95%-KI]	**p-Wert**
Mortalität:		
• Todesfälle gesamt	RR 0,50 [0,12;1,98]	0,321

Endpunkte	**Siponimod vs. BSC**	
	Effektschätzer [95%-KI]	**p-Wert**
Morbidität:		
Bestätigte Krankheitsschübe		
• Jährliche Rate bestätigter Krankheitsschübe	**Rate Ratio 0,45 [0,34;0,59]**	**< 0,001**
• Zeit bis zum ersten bestätigten Schub	**HR 0,54 [0,41; 0,70]**	**< 0,001**
Behinderungsprogression (EDSS)		
• Inzidenz der nach 3 Monaten bestätigten Behinderungsprogression	**RR 0,83 [0,71; 0,97]**	**0,021**
• Inzidenz der nach 6 Monaten bestätigten Behinderungsprogression	**RR 0,78 [0,65; 0,94]**	**0,010**
• Zeit bis zur nach 3 Monaten bestätigten Behinderungsprogression	**HR 0,79 [0,65; 0,95]**	**0,013**
• Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression	**HR 0,74 [0,60; 0,92]**	**0,006**
Behinderungsprogression (MSFC)		
• MSFC-z – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 0,02 [-0,03; 0,07]	0,373
• MSFC-z – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -0,01 [-0,08; 0,07]	0,856
• PASAT-3 – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 0,45 [-0,57; 1,47]	0,387
• PASAT-3 – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 0,59 [-0,79; 1,98]	0,401
• 9-HPT – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 0,00 [-1,50; 1,51]	0,996
• 9-HPT – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -1,20 [-3,18; 0,78]	0,233
• T25W – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 0,03 [-2,24; 2,31]	0,976
• T25W – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 2.38 [-1,25; 6,01]	0,199
• Inzidenz der Verschlechterung im T25W	RR 0,96 [0,85; 1,08]	0,501
• Zeit bis zur Verschlechterung im T25W	HR 0,94 [0,80; 1,10]	0,440
Schweregrad der kognitiven Dysfunktion (SDMT)		
• SDMT – Veränderung seit Studienbeginn zu Monat 12	**MW-Differenz 1,09 [0,23; 1,95]**	**0,013**

Endpunkte	**Siponimod vs. BSC**	
	Effektschätzer [95%-KI]	**p-Wert**
• SDMT – Veränderung seit Studienbeginn zu Monat 24	**MW-Differenz 2,48 [1,11; 3,85]**	**< 0,001**
• Inzidenz der Verbesserung im SDMT	**RR 1,30 [1,09; 1,55]**	**0,003**
• Inzidenz der Verschlechterung im SDMT	**RR 0,78 [0,67; 0,91]**	**0,002**
Schweregrad der kognitiven Dysfunktion (BVMT-R)		
• BVMT-R Delayed Recall Score – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -0,15 [-0,42; 0,12]	0,284
• BVMT-R Delayed Recall Score – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 0,01 [-0,43; 0,45]	0,964
• BVMT-R Total Recall Score – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -0,35 [-1,01; 0,32]	0,307
• BVMT-R Total Recall Score – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 0,45 [-0,59; 1,49]	0,394
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)		
• LCVA – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 0,01 [-0,01; 0,04]	0,366
• LCVA – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 0,00 [-0,03; 0,04]	0,877
Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)		
• MSWS-12 – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -1,83 [-3,85; 0,19]	0,076
• MSWS-12 – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -1,23 [-3,89; 1,44]	0,367
Gesundheitsbezogene Lebensqualität (MSIS-29)		
• MSIS-29 Physical Impact Scale – Veränderung seit Studienbeginn zu Monat 12	**MW-Differenz -2,89 [-4,83; -0,96]**	**0,003**
• MSIS-29 Physical Impact Scale – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -1,36 [-3,93; 1,21]	0,300
• MSIS-29 Psychological Impact Scale – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -2,03 [-4,15; 0,09]	0,060
• MSIS-29 Psychological Impact Scale – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -0,66 [-3,73; 2,40]	0,670
Gesundheitsbezogene Lebensqualität (EQ-5D)		
• EQ-5D-Indexwert – Veränderung seit Studienbeginn zu Monat 12	**MW-Differenz 0,03 [0,00; 0,05]**	**0,037**

Endpunkte	**Siponimod vs. BSC**	
	Effektschätzer [95%-KI]	**p-Wert**
• EQ-5D-Indexwert – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 0,03 [0,00; 0,07]	0,088
• EQ-5D-VAS – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 1,79 [-0,16; 3,75]	0,072
• EQ-5D-VAS – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 0,94 [-2,18; 4,07]	0,555
Sicherheit und Verträglichkeit		
Unerwünschte Ereignisse (UE)		
• UE	**1,09 [1,04; 1,14]**	**< 0,001**
• UE ohne krankheitsspezifische UE	**1,09 [1,03; 1,15]**	**0,001**
• Schwerwiegende Unerwünschte Ereignisse (SUE)	1,25 [0,97; 1,60]	0,080
• SUE ohne krankheitsspezifische UE	**1,37 [1,02; 1,84]**	**0,038**
• Studienabbrüche aufgrund von UE	1,43 [0,82; 2,50]	0,212
• Behandlungsabbrüche aufgrund von UE	1,49 [0,98; 2,26]	0,060
• Schwere UE (CTCAE-Grad ≥ 3)	1,13 [0,85; 1,50]	0,390
• Nicht schwere UE (CTCAE-Grad < 3)	**1,09 [1,04; 1,15]**	**<0,001**
• UE Ereignisse von besonderem Interesse[a]:		
Varizella-Zoster-Infektion (alle Grade)	**4,10 [1,46; 11,51]**	**0,007**
Varizella-Zoster-Infektion (CTCAE-Grad < 3)	**3,73 [1,32; 10,52]**	**0,013**
Makulaödem (alle Grade)	**9,94 [1,34; 73,84]**	**0,025**
Erhöhte Leberenzymwerte (alle Grade)	**3,19 [2,04; 5,00]**	**< 0,001**
Erhöhte Leberenzymwerte (CTCAE-Grad < 3)	**3,27 [2,04; 5,24]**	**< 0,001**

Endpunkte	**Siponimod vs. BSC**	
	Effektschätzer [95%-KI]	**p-Wert**
Hypertonie (alle Grade)	**1,36 [1,00; 1,85]**	**0,049**
Hypertonie (CTCAE-Grad < 3)	**1,39 [1,02; 1,89]**	**0,038**
Krampfanfälle (alle Grade)	**4,72 [1,10; 20,19]**	**0,036**
Periphere Ödeme / Schwellungen (alle Grade)	**2,04 [1,23; 3,38]**	**0,006**
Periphere Ödeme / Schwellungen (CTCAE-Grad < 3)	**2,04 [1,23; 3,38]**	**0,006**
Bradyarrhythmie/Bradykardie (alle Grade)	**1,65 [1,12; 2,43]**	**0,011**
Bradyarrhythmie/Bradykardie (CTCAE-Grad < 3)	**1,67 [1,13; 2,48]**	**0,011**
Suizidalität		
• Suizidgedanken / suizidales Verhalten	1,05 [0,79; 1,40]	0,719
• Selbstverletzung ohne Suizidabsicht	1,51 [0,16; 14,45]	0,722
Ergänzend dargestellt:		
Läsionslast		
• T2-Läsionsvolumen – Veränderung seit Studienbeginn zu Monat 12	**MW-Differenz -613,09 [-800,20; -425,98]**	**< 0,001**
• T2-Läsionsvolumen – Veränderung seit Studienbeginn zu Monat 24	**MW-Differenz -777,52 [-990,63; -564,40]**	**< 0,001**
• Anzahl neuer oder sich vergrößernder T2-Läsionen zu Monat 12	**Rate Ratio 0,27 [0,21; 0,33]**	**< 0,001**
• Anzahl neuer oder sich vergrößernder T2-Läsionen zu Monat 24	**Rate Ratio 0,14 [0,10; 0,20]**	**< 0,001**
• Anzahl Gd-anreichernder T1-Läsionen zu Monat 12	**Rate Ratio 0,13 [0,08; 0,19]**	**< 0,001**
• Anzahl Gd-anreichernder T1-Läsionen zu Monat 24	**Rate Ratio 0,18 [0,09; 0,36]**	**< 0,001**
Hirnvolumen		
• Hirnvolumen – Prozentuale Veränderung seit Studienbeginn zu Monat 12	**MW-Differenz 0,17 [0,10; 0,25]**	**< 0,001**
• Hirnvolumen – Prozentuale Veränderung seit Studienbeginn zu Monat 24	**MW-Differenz 0,13 [0,02; 0,24]**	**0,020**

Endpunkte	Siponimod vs. BSC	
	Effektschätzer [95%-KI]	p-Wert
9-HPT: 9 Hole Peg Test; BVMT-R: Brief Visuospatial Memory Test – Revised; EDSS: Expanded Disability Status Scale; EQ-5D: EuroQol – 5 Dimensionen; HR: Hazard Ratio; KI: Konfidenzintervall; LCVA: Low Contrast Visual Acuity; MSFC: Multiple Sclerosis Functional Composite; MSIS: Multiple Sclerosis Impact Scale; MSWS: Multiple Sclerosis Walking Scale; MW: Mittelwert; PASAT-3: Paced Auditory Serial Addition Test; RR: Relatives Risiko; SDMT: Symbol Digital Modalities Test; SUE: Schwerwiegendes Unerwünschtes Ereignis; T25W: Timed 25 Foot Walking Test; VAS: Visuelle Analogskala. a: Für die Unerwünschten Ereignisse von besonderem Interesse werden nur diejenigen Ereignisse mit statistisch signifikantem relativem Risiko im Vergleich der Behandlungsgruppen dargestellt.		

Subgruppenanalysen zur Gesamtpopulation

In den Subgruppenanalysen ergaben die Interaktionstests keine systematischen und über alle Endpunkte hinweg konsistenten Belege für eine mögliche Modifikation des Behandlungseffekts durch Alter, Geschlecht, DMT-Vorbehandlung, IFN-beta-1b-Vorbehandlung, eine unterschiedlich schnell fortschreitende Erkrankung, Schwere des Erkrankungsverlaufs oder durch die Region (Tabelle 4-4). Bei der Interpretation sind insbesondere die teilweise geringen Fallzahlen in bestimmten Subgruppen und die erhöhte Wahrscheinlichkeit falsch-positiver Belege ($p < 0,05$) durch unadjustierte multiple Tests zu berücksichtigen.

Tabelle 4-4: Zusammenfassung der Ergebnisse aus den Subgruppenanalysen (Gesamtpopulation)

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen	
Bestätigte Krankheitsschübe (Jährliche Rate)			
Geschlecht		**Weiblich**	**Männlich**
	p = 0,037	+	+
Behinderungsprogression (MSFC)			
Geschlecht		**Weiblich**	**Männlich**
Veränderung im MSFC-z	p = 0,004	+ / = [a]	= / = [a]
Veränderung im PASAT-3	p = 0,012	+ / = [a]	= / = [a]
Veränderung im T25W	p = 0,011	= / = [a]	– / = [a]

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen			
Schweregrad der kognitiven Dysfunktion (SDMT)					
Geschlecht		**Weiblich**		**Männlich**	
Inzidenz einer Verschlechterung	p = 0,003	+		=	
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)					
Geschlecht		**Weiblich**		**Männlich**	
Veränderung im LCVA	p = 0,016	= / = [a]		= / = [a]	
DMT-Vorbehandlung		**Nein**		**Ja**	
Veränderung im LCVA	p = 0,012	= / = [a]		= / = [a]	
Gesundheitsbezogene Lebensqualität (MSIS-29)					
Alter		**18 – 30 J.**	**31 – 40 J.**	**41 – 55 J.**	**> 55 J.**
Physical Impact Scale	p = 0,023	= / + [a]	= / = [a]	+ / = [a]	= / = [a]
Psychological Impact Scale	p = 0,044	= / = [a]	= / = [a]	= / = [a]	= / = [a]
Gesundheitsbezogene Lebensqualität (EQ-5D)					
Moderater/schwerer Erkrankungsverlauf		**Nein**		**Ja**	
EQ-5D-VAS	p = 0,036	= / = [a]		= / = [a]	
Unerwünschte Ereignisse					
Alter		**18 – 30 J.**	**31 – 40 J.**	**41 – 55 J.**	**> 55 J.**
SUE (Zeit bis zum Ereignis)	p = 0,040	=	=	=	–
Geschlecht		**Weiblich**		**Männlich**	
Periphere Ödeme	p = 0,006	–		=	
Bradyarrhythmien/ Bradykardien	p = 0,031	=		–	
DMT-Vorbehandlung		**Nein**		**Ja**	
UE (ohne krankheitsspezifische UE)	p = 0,036	–		=	
Herpes-simplex-Infektionen	p = 0,014	+		=	
Atemwegserkrankungen	p = 0,019	+		=	
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**	
Malignitäten	p = 0,044	=		+	
Bradyarrhythmien/ Bradykardien	p = 0,049	=		–	
Schnell fortschreitende Erkrankung		**Nein**		**Ja**	
UE ohne krankheitsspezifische UE (Zeit bis zum Ereignis)	p = 0,050	–		=	
Moderater/schwerer Erkrankungsverlauf		**Nein**		**Ja**	
Herpes-simplex-Infektionen	p = 0,046	=		=	

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen			
Läsionslast					
Alter		**18 – 30 J.**	**31 – 40 J.**	**41 – 55 J.**	**> 55 J.**
T2-Läsionsvolumen	p < 0,001	+ / + [a]	+ / + [a]	+ / + [a]	= / + [a]
DMT-Vorbehandlung		**Nein**		**Ja**	
T2-Läsionsvolumen	p = 0,009	= / + [a]		+ / + [a]	
Neue/größere T2-Läsionen	p = 0,030	+ / + [a]		+ / + [a]	
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**	
Neue/größere T2-Läsionen	p = 0,010	+ / + [a]		+ / + [a]	
Moderater/schwerer Erkrankungsverlauf		**Nein**		**Ja**	
T2-Läsionsvolumen	p = 0,035	= / = [a]		+ / + [a]	
Region		**Europa**		**Andere Regionen**	
T2-Läsionsvolumen	p = 0,009	+/ + [a]		= / + [a]	
Neue/größere T2-Läsionen	p = 0,045	+ / + [a]		+ / + [a]	
Gd-anreichernde T1-Läsionen	p = 0,001	+ / + [a]		+ / = [a]	

Dargestellt sind die Ergebnisse von Subgruppenanalysen nur für diejenigen Subgruppenmerkmale, bei denen der Interaktionstest einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Subgruppenmerkmal ergab.

a: Monat 12 / Monat 24

\+ Siponimod signifikant überlegen ($p < 0,05$).

= Kein Unterschied zwischen den Behandlungen ($p \geq 0,05$).

– Siponimod signifikant unterlegen ($p < 0,05$).

Auswertung der Subpopulation B

Statistisch signifikante Unterschiede zwischen Siponimod und BSC zeigten sich in den folgenden Endpunkten (Tabelle 4-5):

- Bestätigte Krankheitsschübe:
  Unter Therapie mit Siponimod traten statistisch signifikant weniger bestätigte Krankheitsschübe pro Jahr auf als unter BSC (Rate Ratio 0,41; p = 0,031). Außerdem war das Schubrisiko unter Siponimod um 62 % reduziert (HR 0,38; p = 0,026).

- Schweregrad der kognitiven Dysfunktion (SDMT):
  In der Auswertung der Veränderung des SDMT seit Studienbeginn zeigte sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten von Siponimod zu Monat 12 (SDMT-Score MW-Differenz 2,89; p = 0,027). Dabei zeigte sich unter BSC eine Verschlechterung im SDMT, unter Siponimod hingegen blieb der Wert weitestgehend stabil.

- Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12):
  In der Auswertung der Veränderung des MSWS-12 seit Studienbeginn zeigte sich zu Monat 24 ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zuungunsten von Siponimod. Die Auswertung der Veränderung zu Monat 24 ist jedoch wegen sehr geringer Rücklaufquoten nicht zuverlässig interpretierbar.

- Gesundheitsbezogene Lebensqualität (MSIS-29):
  Die Auswertungen der MSIS-29 zeigten einen statistisch signifikanten Vorteil von BSC sowohl in der Physical Impact Scale als auch in der Psychological Impact Scale zu Monat 24. Die Auswertung zu Monat 24 ist jedoch wegen hoher Nicht-berücksichtigungsanteile nicht zuverlässig interpretierbar.

- Unerwünschte Ereignisse:
  Unerwünschte Ereignisse waren meist nur von milder bis moderater Ausprägung. Sie traten insgesamt unter Siponimod statistisch signifikant häufiger auf, der Unterschied zwischen den Gruppen war aber gering (RR 1,17; p = 0,045). Schwere UE (CTCAE-Grad ≥3) traten bei weniger als 10 % der Patienten in der Siponimod-Gruppe auf und die Häufigkeit war hierbei zwischen den Behandlungsgruppen nicht statistisch signifikant unterschiedlich. Auch SUE und Behandlungsabbrüche aufgrund von UE waren unter Siponimod nicht statistisch signifikant häufiger als unter BSC. Unter den UE von besonderem Interesse waren Leberwerterhöhungen (CTCAE-Grad < 3) unter Siponimod statistisch signifikant häufiger. Hierbei handelt es sich um Laborparameter, die in der Nutzenbewertung als nicht patientenrelevant eingestuft werden.

- Läsionslast:
  Unter Siponimod stabilisierte sich das T2-Läsionsvolumen, während es unter BSC weiter zunahm. Der Unterschied zwischen den Behandlungsarmen war zu Monat 12 statistisch signifikant (zu Monat 12: MW-Differenz -977,83 mm^3; p = 0,002). Unter Siponimod bildeten sich außerdem statistisch signifikant weniger neue bzw. sich vergrößernde T2-Läsionen (zu Monat 12: Rate Ratio 0,35; p < 0,001) und Gd-anreichernde Läsionen (zu Monat 12: Rate Ratio 0,15; p < 0,001). Die Auswertungen zu Monat 24 sind wegen hoher Nichtberücksichtigungsanteile nicht zuverlässig interpretierbar.

- Hirnvolumen:
  Die Auswertung der prozentualen Veränderung des Hirnvolumens zeigt einen Vorteil von Siponimod. Unter Siponimod verringerte sich das Hirnvolumen statistisch signifikant weniger als unter BSC (zu Monat 12: MW-Differenz 0,27 %; p = 0,009). Die Auswertung zu Monat 24 ist wegen hoher Nichtberücksichtigungsanteile nicht zuverlässig interpretierbar.

Tabelle 4-5: Zusammenfassung der Ergebnisse aus RCT (Subpopulation B)

Endpunkte	**Siponimod vs. BSC**	
	Effektschätzer [95%-KI]	**p-Wert**
Mortalität:		
• Todesfälle gesamt	RR 0,48 [0,03; 7,55]	0,602
Morbidität:		
Bestätigte Krankheitsschübe		
• Jährliche Rate bestätigter Krankheitsschübe	**Rate Ratio 0,41 [0,18; 0,92]**	**0,031**
• Zeit bis zum ersten bestätigten Schub	**HR 0,38 [0,16; 0,89]**	**0,026**
Behinderungsprogression (EDSS)		
• Inzidenz der nach 3 Monaten bestätigten Behinderungsprogression	RR 0,76 [0,47; 1,23]	0,266
• Inzidenz der nach 6 Monaten bestätigten Behinderungsprogression	RR 0,72 [0,41; 1,25]	0,246
• Zeit bis zur nach 3 Monaten bestätigten Behinderungsprogression	HR 0,65 [0,34; 1,23]	0,185
• Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression	HR 0,57 [0,28; 1,16]	0,121
Behinderungsprogression (MSFC)		
• MSFC-z – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 0,02 [-0,11; 0,15]	0,743
• MSFC-z – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 0,03 [-0,17; 0,22]	0,789
• PASAT-3 – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -0,31 [-3,39; 2,77]	0,842
• PASAT-3 – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 1,31 [-3,16; 5,78]	0,563
• 9-HPT – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -1,03 [-5,13; 3,07]	0,620
• 9-HPT – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -2,89 [-11,28; 5,49]	0,495
• T25W – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -1,76 [-8,05; 4,54]	0,579
• T25W – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -0,74 [-11,55; 10,08]	0,891
• Inzidenz der Verschlechterung im T25W	RR 1,06 [0,77; 1,47]	0,711

Endpunkte	**Siponimod vs. BSC**	
	Effektschätzer [95%-KI]	**p-Wert**
• Zeit bis zur Verschlechterung im T25W	HR 1,27 [0,76; 2,11]	0,367
Schweregrad der kognitiven Dysfunktion (SDMT)		
• SDMT – Veränderung seit Studienbeginn zu Monat 12	**MW-Differenz 2,89 [0,33; 5,44]**	**0,027**
• SDMT – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 3,77 [-0,70; 8,24]	0,097
• Inzidenz der Verbesserung im SDMT	RR 1,14 [0,69; 1,88]	0,604
• Inzidenz der Verschlechterung im SDMT	RR 0,86 [0,57; 1,29]	0,457
Schweregrad der kognitiven Dysfunktion (BVMT-R)		
• BVMT-R Delayed Recall Score – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -0,89 [-1,79; 0,01]	0,052
• BVMT-R Delayed Recall Score – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -0,41 [-1,84; 1,03]	0,574
• BVMT-R Total Recall Score – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -0,52 [-2,55; 1,52]	0,616
• BVMT-R Total Recall Score – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -0,13 [-3,43; 3,16]	0,937
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)		
• LCVA – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 0,02 [-0,06; 0,09]	0,632
• LCVA – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -0,04 [-0,15; 0,06]	0,401
Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)		
• MSWS-12 – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -0,26 [-6,89; 6,36]	0,938
• MSWS-12 – Veränderung seit Studienbeginn zu Monat 24	**MW-Differenz 10,13 [1,96; 18,29]**	**0,016**
Gesundheitsbezogene Lebensqualität (MSIS-29)		
• MSIS-29 Physical Impact Scale – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 1,12 [-5,22; 7,46]	0,727
• MSIS-29 Physical Impact Scale – Veränderung seit Studienbeginn zu Monat 24	**MW-Differenz 8,16 [0,17; 16,16]**	**0,045**
• MSIS-29 Psychological Impact Scale – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 1,98 [-4,77; 8,73]	0,563

Endpunkte	**Siponimod vs. BSC**	
	Effektschätzer [95%-KI]	**p-Wert**
• MSIS-29 Psychological Impact Scale – Veränderung seit Studienbeginn zu Monat 24	**MW-Differenz 10,75 [1,34; 20,16]**	**0,026**
Gesundheitsbezogene Lebensqualität (EQ-5D)		
• EQ-5D-Indexwert – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz 0,04 [-0,02; 0,11]	0,188
• EQ-5D-Indexwert – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz 0,05 [-0,07; 0,17]	0,413
• EQ-5D-VAS – Veränderung seit Studienbeginn zu Monat 12	MW-Differenz -2,11 [-8,02; 3,80]	0,482
• EQ-5D-VAS – Veränderung seit Studienbeginn zu Monat 24	MW-Differenz -4,33 [-12,61; 3,96]	0,302
Sicherheit und Verträglichkeit		
Unerwünschte Ereignisse (UE)		
• UE	**RR 1,17 [1,00; 1,35]**	**0,045**
• UE ohne krankheitsspezifische UE	RR 1,13 [0,96; 1,32]	0,143
• Schwerwiegende Unerwünschte Ereignisse (SUE)	RR 0,70 [0,35; 1,41]	0,318
• SUE ohne krankheitsspezifische UE	RR 0,72 [0,31; 1,67]	0,445
• Studienabbrüche aufgrund von UE	RR 0,32 [0,05; 1,87]	0,205
• Behandlungsabbrüche aufgrund von UE	RR 0,72 [0,21; 2,46]	0,601
• Schwere UE (Grad ≥ 3)	RR 0,53 [0,24; 1,18]	0,118
• Nicht schwere UE (Grad < 3)	RR 1,15 [0,99; 1,34]	0,060
• Unerwünschte Ereignisse von besonderem Interesse[a]:		
Erhöhte Leberenzymwerte (CTCAE-Grad < 3)	**RR 4,32 [1,04; 18,04]**	**0,045**
Suizidalität		
• Suizidgedanken / suizidales Verhalten	RR 1,63 [0,69; 3,84]	0,267
• Selbstverletzung ohne Suizidabsicht	RR nicht bestimmbar, da keine Ereignisse	

Endpunkte	Siponimod vs. BSC	
	Effektschätzer [95%-KI]	**p-Wert**
Ergänzend dargestellt:		
Läsionslast		
• T2-Läsionsvolumen – Veränderung seit Studienbeginn zu Monat 12	**MW-Differenz -977,83 [-1.604,94; -350,73]**	**0,002**
• T2-Läsionsvolumen – Veränderung seit Studienbeginn zu Monat 24	**MW-Differenz -1.603,96 [-2.284,99; -922,93]**	**< 0,001**
• Anzahl neuer oder sich vergrößernder T2-Läsionen zu Monat 12	**Rate Ratio 0,35 [0,20; 0,60]**	**< 0,001**
• Anzahl neuer oder sich vergrößernder T2-Läsionen zu Monat 24	**Rate Ratio 0,08 [0,03; 0,21]**	**< 0,001**
• Anzahl Gd-anreichernder T1-Läsionen zu Monat 12	**Rate Ratio 0,15 [0,06; 0,38]**	**< 0,001**
• Anzahl Gd-anreichernder T1-Läsionen zu Monat 24	Rate Ratio 0,30 [0,08; 1,10]	0,069
Hirnvolumen		
• Hirnvolumen – Prozentuale Veränderung seit Studienbeginn zu Monat 12	**MW-Differenz 0,27 [0,07; 0,47]**	**0,009**
• Hirnvolumen – Prozentuale Veränderung seit Studienbeginn zu Monat 24	**MW-Differenz 0,37 [0,08; 0,66]**	**0,014**

9-HPT: 9 Hole Peg Test; BVMT-R: Brief Visuospatial Memory Test – Revised; EDSS: Expanded Disability Status Scale; EQ-5D: EuroQol – 5 Dimensionen; HR: Hazard Ratio; KI: Konfidenzintervall; LCVA: Low Contrast Visual Acuity; MSFC: Multiple Sclerosis Functional Composite; MSIS: Multiple Sclerosis Impact Scale; MSWS: Multiple Sclerosis Walking Scale; MW: Mittelwert; PASAT-3: Paced Auditory Serial Addition Test; RR: Relatives Risiko; SDMT: Symbol Digital Modalities Test; T25W: Timed 25 Foot Walking Test; VAS: Visuelle Analogskala.

a: Für die Unerwünschten Ereignisse von besonderem Interesse werden nur diejenigen Ereignisse mit statistisch signifikantem relativem Risiko im Vergleich der Behandlungsgruppen dargestellt.

Subgruppenanalysen zur Subpopulation B

In den Subgruppenanalysen ergaben die Interaktionstests keine systematischen und über alle Endpunkte hinweg konsistenten Belege für eine mögliche Modifikation des Behandlungseffekts durch Alter, Geschlecht, DMT-Vorbehandlung, IFN-beta-1b-Vorbehandlung, durch eine unterschiedlich schnell fortschreitende Erkrankung, die Schwere des Erkrankungsverlaufs oder die Region (Tabelle 4-6). Die Alterskategorien 18 – 30 und 31 – 40 Jahre wurden für die Auswertung der Subpopulation B zusammengefasst, da die unterste Altersklasse zu gering besetzt war und keine Analyse erlaubt hatte.

Bei der Interpretation sind insbesondere die teilweise geringen Fallzahlen in bestimmten Subgruppen und die erhöhte Wahrscheinlichkeit falsch-positiver Belege ($p < 0,05$) durch unadjustierte multiple Tests zu berücksichtigen.

Tabelle 4-6: Zusammenfassung der Ergebnisse aus den Subgruppenanalysen (Subpopulation B)

Endpunkt/ Subgruppenmerkmal	**Interaktionstest**	**Subgruppen**		
Behinderungsprogression (EDSS)				
Geschlecht		**Weiblich**		**Männlich**
Nach 6 Monaten bestätigte Behinderungsprogression	p = 0,047	+		=
Behinderungsprogression (MSFC)				
Schnell fortschreitende Erkrankung		**Nein**		**Ja**
Veränderung im T25W	p = 0,024	= / = [a]		= / = [a]
Schweregrad der kognitiven Dysfunktion (SDMT)				
Alter		**18– 40 J.**	**41 – 55 J.**	**> 55 J.**
Inzidenz der Verschlechterung	p = 0,034	=	=	=
Schweregrad der kognitiven Dysfunktion (BVMT-R)				
Geschlecht		**Weiblich**		**Männlich**
Delayed Recall Score	p = 0,036	= / = [a]		– / = [a]
Total Recall Score	p = 0,014	= / = [a]		– / = [a]
Alter		**18 – 40 J.**	**41 – 55 J.**	**> 55 J.**
Total Recall Score	p = 0,016	= / – [a]	= / = [a]	= / = [a]
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**
Total Recall Score	p = 0,034	= / = [a]		= / + [a]
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)				
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**
Veränderung im LCVA	p = 0,038	= / = [a]		= / – [a]
Gesundheitsbezogene Lebensqualität (EQ-5D)				
Schnell fortschreitende Erkrankung		**Nein**		**Ja**
EQ-5D-Index	p = 0,037	= / = [a]		= / + [a]
Unerwünschte Ereignisse				
Geschlecht		**Weiblich**		**Männlich**
UE	p = 0,028	–		=
UE (Zeit bis zum Ereignis)	p = 0,028	–		=

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen	
UE ohne krankheitsspezifische UE (Zeit bis zum Ereignis)	p = 0,033	–	=
Hypertonie	p = 0,050	=	=
DMT-Vorbehandlung		**Nein**	**Ja**
Atemwegserkrankungen	p = 0,038	=	=
Läsionslast			
DMT-Vorbehandlung		**Nein**	**Ja**
Neue/größere T2-Läsionen	p < 0,001	+ / + [a]	= / + [a]
Gd-anreichernde T1-Läsionen	p = 0,006	+ / + [a]	+ / = [a]
IFN-beta-1b-Vorbehandlung		**Nein**	**Ja**
Gd-anreichernde T1-Läsionen	p = 0,027	+ / + [a]	= / = [a]
Region		**Europa**	**Andere Regionen**
Neue/größere T2-Läsionen	p < 0,001	+ / + [a]	= / = [a]
Gd-anreichernde T1-Läsionen	p < 0,001	+ / + [a]	= / = [a]

Dargestellt sind die Ergebnisse von Subgruppenanalysen nur für diejenigen Subgruppenmerkmale, bei denen der Interaktionstest einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Subgruppenmerkmal ergab.

a: Monat 12 / Monat 24

\+ Siponimod signifikant überlegen (p < 0,05).

= Kein Unterschied zwischen den Behandlungen (p > 0,05).

– Siponimod signifikant unterlegen (p < 0,05).

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Es wurden Daten zum Vergleich von Siponimod mit BSC in der Gesamtpopulation der Studie (ergänzend dargestellt) sowie die durch den G-BA definierte bewertungsrelevante Subpopulation B vorgelegt.

Tabelle 4-7: Bewertung des Zusatznutzens bezüglich der untersuchten Endpunkte (Gesamtpopulation und Subpopulation B)

Endpunkt	Zusatznutzen Siponimod vs. BSC	
	Gesamtpopulation	**Subpopulation B**
Mortalität		
• Todesfälle gesamt	=	=

Morbidität		
• Bestätigte Krankheitsschübe	+	+
• Behinderungsprogression (EDSS)	+	=
• Behinderungsprogression (MSFC)	=	=
• Schweregrad der kognitiven Dysfunktion (SDMT)	+	+
• Schweregrad der kognitiven Dysfunktion (BVMT-R)	=	=
• Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)	=	=
• Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)	=	=
Gesundheitsbezogene Lebensqualität		
• MSIS-29	+	=
• EQ-5D-Index	+	=
Sicherheit und Verträglichkeit		
• Unerwünschte Ereignisse (UE / SUE / Abbrüche)	– / = / =	– / = / =
• Unerwünschte Ereignisse von besonderem Interesse	−	=
• Suizidalität	=	=
Ergänzend dargestellt:		
Bildgebung		
• Läsionslast	+	+
• Hirnvolumen	+	+

BSC: Best Supportive Care; BVMT-R: Brief Visuospatial Memory Test – Revised; EDSS: Expanded Disability Status Scale; EQ-5D: EuroQol – 5 Dimensionen; LCVA: Low Contrast Visual Acuity; MSFC: Multiple Sclerosis Functional Composite; MSIS: Multiple Sclerosis Impact Scale; MSWS: Multiple Sclerosis Walking Scale; SDMT: Symbol Digital Modalities Test; SUE: Schwerwiegendes Unerwünschtes Ereignis; UE: Unerwünschtes Ereignis.

\+ Vorteil von Siponimod.

= kein Unterschied zwischen den Behandlungen.

– Nachteil für Siponimod.

Hinsichtlich der Mortalität war kein Unterschied zwischen Siponimod und BSC zu beobachten.

Siponimod zeigte in der Gesamtpopulation sowie in der Subpopulation B eine statistisch signifikante Reduktion der jährlichen Rate *bestätigter Krankheitsschübe* und eine statistisch signifikante Reduktion des Schubrisikos.

Im Endpunkt *Behinderungsprogression (EDSS)* zeigte sich, dass die Inzidenz einer Behinderungsprogression in der Gesamtpopulation signifikant reduziert wurde bei gleichzeitig signifikant reduziertem Risiko einer Behinderungsprogression (Zeit bis zur ersten bestätigten Progression). In der Subpopulation B konnte in diesem Endpunkt kein statistisch signifikanter Unterschied festgestellt werden.

Es konnten sowohl in der Gesamtpopulation als auch in der Subpopulation B jeweils eine statistisch signifikante Reduktion *kognitiver Beeinträchtigungen* (gemessen anhand des SDMT) unter Siponimod beobachtet werden.

In der *Gesundheitsbezogenen Lebensqualität* zeigte sich in der Gesamtpopulation ein statistisch signifikanter Vorteil von Siponimod, gemessen anhand des validierten MSIS-29-Fragenbogens sowie des EQ-5D. In der Subpopulation B konnte in diesem Endpunkt kein statistisch signifikanter Unterschied festgestellt werden.

Bei den sicherheitsrelevanten Endpunkten (*Unerwünschte Ereignisse*) zeigte sich in der Gesamtpopulation und in der Subpopulation B eine statistisch signifikant höhere Inzidenz unter Siponimod, die aber vor allem auf Ereignisse mit geringem Schweregrad zurückzuführen ist. Schwere UE waren nicht statistisch signifikant häufiger. SUE, Behandlungsabbrüche aufgrund von UE und patientenrelevante UE von besonderem Interesse traten in der Subpopulation B unter einer Therapie mit Siponimod nicht statistisch signifikant häufiger auf als unter BSC. In der Gesamtpopulation waren UE von besonderem Interesse unter Siponimod statistisch signifikant häufiger, meist aber von milder bis moderater Ausprägung.

Daneben konnte Siponimod sowohl in der Gesamtpopulation als auch in der Subpopulation B die *Läsionslast* (T2- sowie Gd-anreichernde T1-Läsionen) sowie die Hirnatrophie (*Verringerung des Hirnvolumens)* statistisch signifikant reduzieren. Beides sind zentrale Parameter in der Überwachung des Therapieerfolgs in der klinischen Praxis und werden daher ergänzend dargestellt.

Die Aussagekraft der randomisierten, kontrollierten Studie, die der Bewertung zugrunde liegt, ist sehr hoch; damit ist die Studie für eine valide Beurteilung des medizinischen Nutzens und Zusatznutzens uneingeschränkt geeignet. Ebenso lassen sich die Ergebnisse der Studie uneingeschränkt auf den deutschen Versorgungskontext übertragen. Die Subgruppenanalysen ergaben keinen eindeutigen Beleg für eine mögliche Modifikation des therapeutischen Effekts durch Alter, Geschlecht, DMT-Vorbehandlung, IFN-beta-1b-Vorbehandlung, eine unterschiedlich schnell fortschreitende Erkrankung, die Schwere des Erkrankungsverlaufs oder durch die Region.

Ein geringer Zusatznutzen liegt gemäß Verfahrensordnung §5 Absatz 7 Ziffer 3 vor, wenn eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapierelevanten Nutzens erreicht wird, insbesondere eine Verringerung von nicht schwerwiegenden Symptomen der Erkrankung. Auf Basis der beobachteten Reduktion von Krankheitsschüben und einer Reduktion kognitiver Beeinträchtigungen im Vergleich zur zweckmäßigen Vergleichstherapie ergibt sich daher ein

Hinweis auf einen geringen patientenrelevanten Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie in der Subpopulation B (Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe).

Der Zusatznutzen in der Subpopulation B ist aufgrund der Studie der Evidenzstufe Ib, die der Bewertung zugrunde liegt, als gesichert anzusehen.

Für die Subpopulation A (Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, mit aufgesetzten Schüben) liegen keine Daten gegenüber der zweckmäßigen Vergleichstherapie vor. Ein Zusatznutzen ist für die Subpopulation A daher nicht belegt.

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).

Die Benennung der Vergleichstherapie in Modul 4 muss zur Auswahl der zweckmäßigen Vergleichstherapie im zugehörigen Modul 3 konsistent sein.

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.

Zielsetzung

Mit dem vorliegenden Dossier wird der medizinische Zusatznutzen von Siponimod bei erwachsenen Patienten mit sekundär progredienter Multipler Sklerose (SPMS) mit Krankheitsaktivität, nachgewiesen durch Schübe oder Bildgebung der entzündlichen Aktivität, gemäß Zulassung bewertet.

Der medizinische Nutzen von Siponimod bei erwachsenen Patienten mit SPMS wurde in der großen placebokontrollierten Studie EXPAND (CBAF312A2304) nachgewiesen und durch die Zulassung durch die europäischen Behörden bestätigt.

Population

Das für die Fragestellung betrachtete Patientenkollektiv umfasst gemäß Anwendungsgebiet erwachsene Patienten mit SPMS mit Krankheitsaktivität, nachgewiesen durch Schübe oder Bildgebung der entzündlichen Aktivität (1). Gemäß G-BA-Beratungen und Rückmeldung nach

positiver CHMP-Empfehlung wird dieses Kollektiv für die Nutzenbewertung in zwei Patientenpopulationen aufgeteilt (Tabelle 4-8) (2-5).

Tabelle 4-8: Überblick zur Festlegung der zu bewertenden Subpopulationen

Zu bewertende Subpopulationen	Zweckmäßige Vergleichstherapie	Vorhandene Evidenz
Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, mit aufgesetzten Schüben (Subpopulation A)	IFN-beta 1a oder 1b oder Ocrelizumab	Keine
Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe (Subpopulation B)	Best Supportive Care (BSC)	RCT gegen BSC (EXPAND)

Zur Beantwortung der Fragestellung für die Subpopulation A liegt keine Evidenz zum Vergleich von Siponimod mit der festgelegten zweckmäßigen Vergleichstherapie vor.

Die Beantwortung der Fragestellung erfolgt für die Subpopulation B im Vergleich zur zweckmäßigen Vergleichstherapie Best Supportive Care (BSC) auf der Grundlage der Evidenz der EXPAND-Studie (CBAF312A2304, NCT01665144).

Von einer aktiven Erkrankung wird im Anwendungsgebiet von Siponimod gesprochen, wenn klinische Schübe vorliegen oder eine entzündliche Aktivität in der Bildgebung (MRT) nachgewiesen werden kann (1). In diesem Sinne wird eine aktive Erkrankung beispielsweise auch in den zugelassenen Anwendungsgebieten von Ocrelizumab („[...] mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung“) konkretisiert (6). Auch Lublin et al. definieren in der 2014 veröffentlichten neuen Kategorisierung der MS eine aktive Erkrankung durch das Vorhandensein klinischer Schübe oder MRT-Aktivität in Form von Kontrastmittel-anreichernden Läsionen oder neuen bzw. sich vergrößernden T2-Läsionen (7).

Die Zielpopulation im zugelassenen Anwendungsgebiet sind daher ausschließlich Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinische Schübe oder Aktivität in der Bildgebung (MRT-Aktivität). Zur Beantwortung der Fragestellung zur Subpopulation B werden aus der EXPAND-Studie daher die Patienten ausgewertet, die diese Kriterien erfüllen. Aufgrund der Notwendigkeit eines mind. ein Jahr zurückliegenden, gemäß Studienkriterien standardisierten Referenz-MRTs bei Studienbeginn, können keine Angaben zu T2-Veränderungen gemacht werden. Somit liegen liegen keine Information zu neuen bzw. sich vergrößernden T2-Läsionen zu Studienbeginn vor. Die Selektion der Subpopulation B aus der EXPAND-Studie erfolgte deshalb anhand folgender Charakteristika:

- kein klinischer Schub in den zwei Jahren vor Studieneinschluss, aber
- Aktivität in der Bildgebung (MRT) in Form von Kontrastmittel- (Gd-) anreichernden Läsionen.

Mit Patienten, die während der Studie einen Schub hatten, wurde ein Gespräch geführt und sie mussten einem Verbleib in der Studie erneut zustimmen. Damit wurde sichergestellt, dass die weitere Behandlung mit Studienmedikation für diese Patienten angemessen war und kein Wechsel auf eine Schubprophylaxe mit IFN-beta angezeigt war.

Die Ergebnisse der Gesamtpopulation der EXPAND-Studie werden ergänzend dargestellt.

Intervention

Gegenstand der Nutzenbewertung und damit hier relevante Intervention ist der S1P-Rezeptormodulator Siponimod. Laut Fachinformation wird Siponimod nach einer Auftitrationsphase von fünf Tagen ab dem sechsten Tag in einer Erhaltungsdosis von 2 mg einmal täglich eingesetzt. Patienten mit intermediärer CYP2C9-Enzymaktivität (Genotypen CYP2C9*2*3 oder -*1*3) erhalten eine niedrigere Erhaltungsdosis von 1 mg. Die Einnahme erfolgt oral. Die Auftitration erfolgt nach folgendem Schema:

Tabelle 4-9: Auftitrationsschema

	Zieldosis 2 mg	**Zieldosis 1 mg**
Tag 1	0,25 mg	0,25 mg
Tag 2	0,25 mg	0,25 mg
Tag 3	0,5 mg	0,5 mg
Tag 4	0,75 mg	0,75 mg
Tag 5	1,25 mg	1,25 mg
Ab Tag 6	2 mg	1 mg

Vergleichstherapie

Die zweckmäßige Vergleichstherapie für das Anwendungsgebiet von Siponimod wurde durch den G-BA bestimmt (4).

Die zweckmäßige Vergleichstherapie zur Behandlung von erwachsenen Patienten mit sekundär progredienter Multipler Sklerose (SPMS) mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ist:

- bei Patienten mit aufgesetzten Schüben (im Dossier bezeichnet als Subpopulation A) Interferon-beta 1 a oder 1 b oder Ocrelizumab;

- bei Patienten ohne aufgesetzte Schübe (im Dossier bezeichnet als Subpopulation B) Best-Supportive-Care.
  Als BSC wird die Therapie verstanden, die eine bestmögliche, patientenindividuell optimierte, unterstützende Behandlung zur Linderung von Symptomen und Verbesserung der Lebensqualität gewährleistet.

In diesem Dossier werden Daten im Vergleich zu BSC (EXPAND-Studie) gezeigt. BSC wird in der EXPAND-Studie als Placebo-Vergleich mit zusätzlicher Möglichkeit der Anwendung pharmakologischer sowie nicht-pharmakologischer Interventionen zur Behandlung der Symptome umgesetzt. Ausgeschlossen waren Behandlungen, die immunsuppressiv oder immunmodulierend wirken. Zu BSC wurde im Beratungsgespräch ausgeführt, „dass für die verwendeten unterstützenden Maßnahmen, wie z. B. bei möglicher Begleitmedikation der Zulassungsstatus zu berücksichtigen ist. Im deutschen Versorgungskontext werden Patienten mit körperlichen Einschränkungen in der Regel auch physikalische Therapie oder krankengymnastische Maßnahmen angeboten. Dies sollte bei Best-Supportive-Care berücksichtigt werden (4). Die Operationalisierung von BSC in der Studie EXPAND entspricht damit der Definition des G-BA.

Endpunkte

Die Bewertung des Zusatznutzens erfolgt auf Basis von Endpunkten zu Mortalität, Morbidität, Gesundheitsbezogener Lebensqualität sowie Sicherheit und Verträglichkeit. Ergänzend werden Ergebnisse zur Krankheitsaktivität basierend auf der Bildgebung mittels MRT präsentiert.

- Mortalität
    - Todesfälle gesamt
- Morbidität
    - Bestätigte Krankheitsschübe
    - Behinderungsprogression (EDSS)
    - Behinderungsprogression (MSFC)
    - Schweregrad der kognitiven Dysfunktion (SDMT)
    - Schweregrad der kognitiven Dysfunktion (BVMT-R)
    - Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)
    - Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)
- Gesundheitsbezogene Lebensqualität
    - MSIS-29
    - EQ-5D
- Sicherheit und Verträglichkeit
    - Unerwünschte Ereignisse
    - Schwerwiegende Unerwünschte Ereignisse
    - Schwere Unerwünschte Ereignisse
    - Unerwünschte Ereignisse, die zum Abbruch der Studie führten
    - Unerwünschte Ereignisse, die zum Absetzen der Studienmedikation führten

  - Unerwünschte Ereignisse nach Schweregrad
  - Unerwünschte Ereignisse von besonderem Interesse
  - Columbia-Beurteilungsskala zur Suizidalität (Columbia Suicide Severity Rating Scale; C-SSRS)

- Bildgebung
  - Läsionslast im MRT (T2-Läsionen, Gd-anreichernde T1-Läsionen)
  - Hirnvolumen

Studientypen

Zur Beantwortung der Fragestellung werden nur RCT (Evidenzstufe 1b) oder Meta-Analysen aus RCT (Evidenzstufe 1a) herangezogen.

Zur Beantwortung der Fragestellung für die Subpopulation A liegt keine entsprechende Evidenz zum Vergleich von Siponimod mit der festgelegten zweckmäßigen Vergleichstherapie vor.

Die Beantwortung der Fragestellung erfolgt für die Subpopulation B im Vergleich zur zweckmäßigen Vergleichstherapie BSC auf der Grundlage der Evidenz der EXPAND-Studie, einer doppelblinden, randomisierten, placebokontrollierten klinischen Studie (NCT01665144/CBAF312A2304).

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/ einer Studienergebnisdatenbank erfolgen, während ein Kongressabstrakt 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. Erstellen Sie dabei für unterschiedliche Themen der Recherche (z. B. unterschiedliche Fragestellungen) jeweils eine separate Übersicht.

Studien wurden nach folgenden Kriterien für die Bewertung des medizinischen Nutzens und Zusatznutzens ein- bzw. ausgeschlossen:

Tabelle 4-10: Ein- und Ausschlusskriterien für Studie (Fragestellung für Subpopulation A)

	Einschlusskriterien	**Ausschlusskriterien**
Patientenpopulation	Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, mit aufgesetzten Schüben.	Andere Populationen.
Intervention	Siponimod 2 mg (bzw. 1 mg bei Patienten mit CYP2C9-Polymorphismus), 1 × täglich oral.	Andere Interventionen.
Vergleichstherapie	IFN-beta 1a oder 1b oder Ocrelizumab.	Andere oder keine Vergleichstherapie.
Endpunkte	Mindestens einer der patientenrelevanten Endpunkte wird berichtet: • Mortalität • Morbidität (z. B. Krankheitsschübe, Behinderungsprogression, Schweregrad der Behinderung) • Gesundheitsbezogene Lebensqualität • Sicherheit und Verträglichkeit (z. B. Unerwünschte Ereignisse)	Studien ohne Ergebnisse zu mindestens einem der genannten relevanten Endpunkte.
Studientypen	RCT, Meta-Analysen von RCT.	Andere Studientypen.
Studiendauer	Zwölf Monate und länger.	Studien mit kürzerer Behandlungsdauer.
Dokumentationstyp	Studienbericht, Volltextpublikation, ausführlicher Ergebnisbericht aus einem Studienregister / einer Studienergebnisdatenbank.	Ausschließlich andere Dokumenttypen als Studienbericht und Volltextpublikation (z. B. Konferenzabstracts, Editorials, Letters to the editor).

Tabelle 4-11: Ein- und Ausschlusskriterien für Studie (Fragestellung für Subpopulation B)

	Einschlusskriterien	**Ausschlusskriterien**
Patientenpopulation	Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe.	Andere Populationen
Intervention	Siponimod 2 mg (bzw. 1 mg bei Patienten mit CYP2C9-Polymorphismus), 1 × täglich oral.	Andere Interventionen.
Vergleichstherapie	BSC	Andere oder keine Vergleichstherapien.
Endpunkte	Mindestens einer der patientenrelevanten Endpunkte wird berichtet: • Mortalität • Morbidität (z. B. Krankheitsschübe, Behinderungsprogression, Schweregrad der Behinderung) • Gesundheitsbezogene Lebensqualität • Sicherheit und Verträglichkeit (z. B. Unerwünschte Ereignisse)	Studien ohne Ergebnisse zu mindestens einem der genannten relevanten Endpunkte.
Studientypen	RCT, Meta-Analysen von RCT.	Andere Studientypen.
Studiendauer	Zwölf Monate und länger.	Studien mit kürzerer Behandlungsdauer.
Dokumentationstyp	Studienbericht, Volltextpublikation, ausführlicher Ergebnisbericht aus einem Studienregister / einer Studienergebnisdatenbank.	Ausschließlich andere Dokumenttypen als Studienbericht und Volltextpublikation (z. B. Konferenzabstracts, Editorials, Letters to the editor).

4.2.3 Informationsbeschaffung

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

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1

und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

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

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

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

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE (inklusive „in-process & other non-indexed citations) 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.

Um relevante Studien für den Vergleich von Siponimod gegen die zweckmäßige Vergleichstherapie zu identifizieren, wurden die Datenbanken EMBASE, MEDLINE und Cochrane Central Register of Clinical Trials durchsucht. Die Publikationssprache wurde nicht eingeschränkt. Eine zeitliche Restriktion wurde nicht vorgenommen. Die zeitlichen Angaben beziehen sich auf den maximal verfügbaren Zeitraum (MEDLINE ab 1946, EMBASE ab 1974). Das Cochrane Register of Clinical Trials wurde bis einschließlich Oktober 2019 vollständig durchsucht. Die Suche erfolgte am 18. November 2019. Die Suchstrategien sind in Anhang 4-A dokumentiert.

4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken

Eine Suche in öffentlich zugänglichen Studienregistern/ Studienergebnisdatenbanken ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern / Studienergebnisdatenbanken vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern/ Studienergebnisdatenbanken muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern sowie Studienergebnisdatenbanken 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/ Studienergebnisdatenbanken clinicaltrials.gov (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal), Suchportal der WHO, Clinical Data Suchportal der European Medicines Agency (https://clinicaldata.ema.europa.eu) sowie dem Arzneimittel-Informationssystem (AMIS, https://www.pharmnet-bund.de/dynamic/de/arzneimittel-informationssystem/index.html) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern / Studienergebnisdatenbanken (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern/ Studienergebnisdatenbanken anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

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

Für Clinical Data (Suchportal der European Medicines Agency) und das Arzneimittel-Informationssystem (AMIS) genügt hingegen die Suche nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern/ Studienergebnisdatenbanken 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.

In folgenden öffentlichen Studienregistern/Studienergebnisdatenbanken wurde nach klinischen Studien mit Siponimod analog zur bibliografischen Recherche gesucht:

- clinicaltrials.gov (www.clinicaltrials.gov),
- EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu),
- International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal, Suchportal der WHO: http//apps.woho.int/trialsearch/)

Bei den Recherchen wurden keine generellen Einschränkungen vorgenommen. Die durchgeführten Suchstrategien entsprechen den obigen Vorgaben und sind in Anhang 4-B dokumentiert.

In folgenden Datenbanken wurde nach Einträgen mit Ergebnisberichten zu Studien gesucht, die bereits anderweitig (z. B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden:

- Clinical Data Suchportal der European Medicines Agency (https://clinicaldata.ema.europa.eu)
- Arzneimittel-Informationssystem (AMIS, www.pharmnet-bund.de/dyna-mic/de/am-info-system/index.html)

Eine Dokumentation der zugehörigen Suchstrategie erfolgte hier gemäß Vorgabe zu Abschnitt 4.2.3.3 nicht.

4.2.3.4 Suche auf der Internetseite des G-BA

Die Internetseite des G-BA ist grundsätzlich zu durchsuchen, um sicherzustellen, dass alle vorliegenden Daten zu Studienmethodik und –ergebnissen von relevanten Studien in die Bewertung einfließen.

Auf der Internetseite des G-BA werden Dokumente zur frühen Nutzenbewertung nach §35a SGB V veröffentlicht. Diese enthalten teilweise anderweitig nicht veröffentlichte Daten zu Studienmethodik und –ergebnissen[1]. Solche Daten sind dabei insbesondere in den Modulen 4 der Dossiers pharmazeutischer Unternehmer, in IQWiG-Nutzenbewertungen sowie dem

[1] Köhler M, Haag S, Biester K, Brockhaus AC, McGauran N, Grouven U, Kölsch H, Seay U, Hörn H, Moritz G, Staeck K, Wieseler B. Information on new drugs at market entry: retrospective analysis of health technology assessment reports, journal publications, and registry reports. BMJ 2015;350:h796

Beschluss des G-BA einschließlich der Tragenden Gründe und der Zusammenfassenden Dokumentation zu erwarten.

Die Suche auf der Internetseite des G-BA muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche auf der G-BA Internetseite immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird. Die Suche ist dann sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie durchzuführen. Es genügt die Suche nach Einträgen zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

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

Beschreiben Sie nachfolgend das Vorgehen für die Suche. Benennen Sie die Wirkstoffe und die auf der Internetseite des G-BA genannten zugehörigen Vorgangsnummern, zu denen Sie eine Suche durchgeführt haben.

Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Die Internetseite des G-BA (Dokumente zur frühen Nutzenbewertung) wurde nach Informationen zu RCT mit dem zu bewertenden Arzneimittel Siponimod bzw. mit den zweckmäßigen Vergleichstherapien IFN-beta und Ocrelizumab im relevanten Anwendungsgebiet durchsucht. Für Siponimod ist das vorliegende Dossier Gegenstand des ersten Verfahrens, es ist daher noch kein Vorgang zur frühen Nutzenbewertung aufgeführt. Für IFN-beta liegen keine Nutzenbewertungen zur SPMS vor. Für Ocrelizumab ist ein Verfahren unter der Vorgangsnummer 2018-02-01-D-332 gelistet, das jedoch nicht die Indikation SPMS umfasst. Eine Dokumentation der zugehörigen Suchstrategie erfolgte hier gemäß Verfahrensordnung nicht.

4.2.3.5 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2, 4.2.3.3 und 4.2.3.4 beschriebenen Rechercheschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die über die bibliografische Literaturrecherche und Suche in Studienregistern identifizierten Studien wurden anhand ihres Titels und, sofern vorhanden, ihres Abstracts von zwei Reviewern unabhängig voneinander hinsichtlich ihrer Relevanz für den Einschluss in die vorliegende Bewertung anhand der in Abschnitt 4.2.2 beschriebenen Kriterien beurteilt. Publikationen, die

beide Reviewer als potenziell relevant erachteten, wurden anhand ihres Volltextes geprüft und bei Relevanz in das Dossier zur Nutzenbewertung eingeschlossen. Publikationen, die nur ein Reviewer als potenziell relevant einstufte, wurden von beiden Reviewern nochmals gesichtet und anschließend nach Diskussion und Volltextprüfung bei Relevanz in das Dossier zur Nutzenbewertung eingeschlossen.

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.

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise wurden Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie endpunktübergreifend und endpunktspezifisch beschrieben. Für randomisierte Studien wurde das Verzerrungspotenzial bewertet und als ‚niedrig‘ oder ‚hoch‘ eingestuft. Ein niedriges Verzerrungspotenzial lag dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden konnte, 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.

Für die Bewertung eines Endpunkts wurde für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend als ‚niedrig‘ oder ‚hoch‘ eingestuft. Falls die Einstufung als ‚hoch‘ erfolgte, wurde das Verzerrungspotenzial für den Endpunkt in der Regel auch als ‚hoch‘ bewertet, etwaige Abweichungen hiervon wurden begründet. Ansonsten wurden endpunktspezifische Aspekte, wie Verblindung der Endpunkterheber, Umsetzung des ITT-Prinzips, Hinweise auf ergebnisgesteuerte Berichterstattung und sonstige Aspekte, berücksichtigt. Wurde das Verzerrungspotenzial auf Endpunktebene als ‚hoch‘ bewertet, führte dies nicht zum Ausschluss der Daten zu diesem Endpunkt. Die Bewertung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Details zum Verzerrungspotenzial sind den Bewertungsbögen in Anhang 4-F zu entnehmen. Die eingesetzte Methodik weicht somit nicht von der oben beschriebenen Methodik ab.

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)[2]. Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-[3] bzw. STROBE-Statements[4] 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.

Die Beschreibung des Designs und der Methodik der RCT erfolgte anhand der Kriterien des CONSORT-Statements. Informationen zu den Items wurden jeweils dem Studienprotokoll und dessen Amendments bzw. dem Studienbericht und den relevanten Publikationen der eingeschlossenen RCT entnommen. Die ausführliche Darstellung pro Item findet sich in Anhang 4-E.

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.

[2] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.
[3] 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.
[4] 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.

Sofern zur Berechnung von Ergebnissen von Standardverfahren und –software abgewichen wird (insbesondere beim Einsatz spezieller Software oder individueller Programmierung), sind die Berechnungsschritte und ggf. verwendete Software explizit abzubilden. Insbesondere der Programmcode ist in lesbarer Form anzugeben.

Patientencharakteristika:

Die im Folgenden beschriebenen Studienergebnisse wurden an erwachsenen Patienten mit SPMS erhoben. Dieses Studienkollektiv ist für die Beurteilung des medizinischen Nutzens und Zusatznutzens von Siponimod im hier relevanten Anwendungsgebiet geeignet. Es entspricht der Patientengruppe, die für die Behandlung infrage kommt.

Für die in die Nutzenbewertung eingeschlossene Studie wurden folgende Baseline-Patientencharakteristika dargestellt:

- Alter
- Geschlecht
- Gewicht
- Körpergröße
- Body-Mass-Index (BMI)
- Rasse
- Ethnie
- Erkrankungsdauer seit Diagnose
- Erkrankungsdauer seit Auftreten der ersten Symptome
- Zeit seit Konversion zur SPMS
- Anzahl der Schübe in den letzten 12 bzw. 24 Monaten
- Schweregrad der Behinderung (EDSS)
- Anzahl der Gd-anreichernden Läsionen
- T2-Läsionsvolumen
- Hirnvolumen

Endpunkte:

Tabelle 4-12: Im Dossier dargestellte patientenrelevante Endpunkte

Endpunkt	Auswertung[a]
Mortalität	
Todesfälle gesamt	• Inzidenz
Morbidität	
Bestätigte Krankheitsschübe	• Jährliche Rate bestätigter Krankheitsschübe • Zeit bis zum ersten bestätigten Schub
Behinderungsprogression – EDSS	• Inzidenz der nach 3 bzw. 6 Monaten bestätigten Behinderungsprogression • Zeit bis zur ersten nach 3 bzw. 6 Monaten bestätigten Behinderungsprogression

Behinderungsprogression – MSFC	• Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn im MSFC-z / PASAT-3 / 9-HPT / T25W • Bestätigte (nach 3 Monaten) Verschlechterung im T25W um mindestens 20 % • Zeit bis zur ersten bestätigten (nach 3 Monaten) Verschlechterung im T25W um mindestens 20 % • Bestätigte (nach 3 Monaten) Verbesserung im T25W um mindestens 20 % • Zeit bis zur ersten bestätigten (nach 3 Monaten) Verbesserung im T25W um mindestens 20 %
Schweregrad der kognitiven Dysfunktion – SDMT	• Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn • Inzidenz einer Verbesserung des SDMT-Score um mindestens 4 Punkte • Inzidenz einer Verschlechterung des SDMT-Score um mindestens 4 Punkte
Schweregrad der kognitiven Dysfunktion – BVMT-R	• Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn im Delayed Recall Score bzw. im Total Recall Score
Schweregrad der Beeinträchtigung der Gehfähigkeit – MSWS-12	• Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn im MSWS-12 Converted Score
Schweregrad der Beeinträchtigung des Sehvermögens – LCVA	• Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn
Gesundheitsbezogene Lebensqualität	
MSIS-29	• Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn in der MSIS-29 Physical Impact Scale und der MSIS-29 Psychological Impact Scale
EQ-5D	• Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn im EQ-5D-Indexwert und in der EQ-5D-VAS
Sicherheit und Verträglichkeit	
Unerwünschte Ereignisse	• Inzidenz und Zeit bis zum ersten Ereignis für: UE / SUE / Schwere UE / Studien- bzw. Behandlungsabbruch wegen UE / UE von besonderem Interesse gemäß RMP
Suizidalität (C-SSRS)	• Inzidenz von suizidalen Gedanken / suizidalem Verhalten
Ergänzend dargestellt: Bildgebung	
Läsionslast im MRT	• Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn im T2-Läsionsvolumen • Anzahl neuer oder sich vergrößernder T2-Läsionen zu Monat 12 bzw. 24 • Anzahl Gd-anreichernder Läsionen zu Monat 12 bzw. 24
Hirnvolumen	• Prozentuale Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn im Hirnvolumen

a: Details zur Operationalisierung siehe Abschnitt 4.3.1.3

Zur Bewertung des medizinischen Nutzens und Zusatznutzens von Siponimod wurden ausschließlich Endpunkte gewählt, die für die Studienziele geeignet und darüber hinaus patientenrelevant sind (Tabelle 4-12). Alle in der vorliegenden Bewertung betrachteten Endpunkte stellen im Anwendungsgebiet etablierte Endpunkte dar, die auch bei der Zulassung

neuer krankheitsmodifizierender Therapeutika als Parameter des Wirksamkeitsnachweises gefordert werden (8).

Mortalität

Bei der Mortalität handelt es sich um einen in der Verfahrensordnung definierten patientenrelevanten Endpunkt. Im vorliegenden Dossier wurde die Häufigkeit der aufgetretenen Todesfälle berichtet.

Morbidität

Entsprechend den Vorgaben nach § 3 der Verfahrensordnung gehören zu den patientenrelevanten Endpunkten die Verbesserung des Gesundheitszustandes, die Verkürzung der Krankheitsdauer und die Verringerung von Nebenwirkungen. Folgende Endpunkte zur Morbidität wurden im Dossier dargestellt:

Bestätigte Krankheitsschübe

Je nach klinischem Bild unterscheidet man zwischen einer SPMS mit aufgesetzten Schüben und einer SPMS ohne aufgesetzte Schübe. Dabei kann jeweils nur retrospektiv beurteilt werden, ob ein Patient in der Vergangenheit, üblicherweise im Zeitraum der vergangenen zwei Jahre, noch Schubaktivität gezeigt hat. Allerdings ist auch bei Patienten, die über einen längeren Zeitraum von z. B. zwei Jahren keine Schubaktivität gezeigt hatten, ein Wiederauftreten von Schüben nicht vollständig auszuschließen. So hatten z. B. in der EXPAND-Studie 19 % der BSC-Patienten, die in den zwei Jahren vor Studienbeginn ohne Schübe waren, während der Studie einen bestätigten Krankheitsschub (9). Da Schübe bei SPMS-Patienten im Vergleich zur RRMS eine geringere Rückbildungstendenz haben (10), ist die Schubprophylaxe bei allen SPMS-Patienten essenziell zur Verringerung des Fortschreitens der Erkrankung. Dies gilt auch für alle Patienten, die längere Zeit keinen Schub hatten (in der EXPAND-Studie zwei Jahre vor Studieneinschluss) und daher die formale Definition einer SPMS ohne aufgesetzte Schübe erfüllen.

Ein Krankheitsschub war im Prüfplan der EXPAND-Studie definiert als neue neurologische Symptome oder Verschlechterung von zuvor stabilen oder sich verbessernden vorbestehenden Symptomen, die mindestens 30 Tage nach einem vorhergehenden demyelinisierenden Ereignis auftraten. Die Symptome mussten für mindestens 24 Stunden andauern und ohne Fieber (< 37,5° C) oder eine Infektion aufgetreten sein. Das Auftreten von Symptomen eines Krankheitsschubs wurde bei jeder Visite durch den Arzt erfragt. Zudem konnten Symptome auch jederzeit durch den Patienten selbst spontan berichtet werden.

Ein bestätigter Krankheitsschub lag vor, wenn die genannten Kriterien eines Krankheitsschubs erfüllt waren und zudem ein klinisch relevanter Anstieg im Wert der Expanded Disability Status Scale (EDSS-Wert) zu verzeichnen war. Dabei musste es sich um einen Anstieg um 0,5 Punkte im EDSS-Wert insgesamt oder um einen Anstieg um einen Punkt auf zwei verschiedenen „Funktionellen Systemen“ (FS) der EDSS oder um einem Anstieg um zwei Punkte auf einem FS (außer Darm/Blase oder zerebrales FS) handeln. Der Anstieg musste durch einen

unabhängigen EDSS-Bewerter festgestellt werden. Die EDSS-Bewertung sollte binnen sieben Tagen nach Einsetzen der Symptome geschehen.

Das Auftreten von Krankheitsschüben hat der G-BA im Beratungsgespräch am 9. Mai 2019 als patientenrelevanten Endpunkt definiert (4).

Behinderungsprogression, gemessen anhand des EDSS

Die EDSS (Expanded Disability Status Scale) ist eine etablierte Skala zur Bewertung von neurologischen Beeinträchtigungen bei MS (11, 12). Sie umfasst eine Serie von Punktwerten in acht verschiedenen „Funktionellen Systemen“ (FS) sowie die EDSS-Stufen von 0 (normal) bis 10 (Tod aufgrund von MS). Funktionelle Systeme sind (mögliche Symptome in Klammern): Pyramidenbahn (z. B. Lähmungen), Kleinhirn (z. B. Ataxie, Tremor), Hirnstamm (z. B. Sprach-/Schluckstörungen), Sensorium (z. B. Verminderung des Berührungssinns), Blasen- und Mastdarmfunktionen (z. B. Harn- bzw. Stuhlinkontinenz), Sehfunktionen (z. B. eingeschränktes Gesichtsfeld, Skotom), zerebrale Funktionen (z. B. Wesensveränderung, Demenz) und andere Funktionen (andere neurologische Bereiche, die auf MS zurückzuführen sind).

Bei SPMS-Patienten hat sich bis zum Konversionszeitpunkt zur SPMS bereits ein deutlich spürbares Maß an Behinderung akkumuliert. Der EDSS-Wert liegt bei Konversion zur SPMS bei 75 % der Patienten bereits über dem Wert von 3,0 (13). Ein EDSS-Wert von 3,0 bedeutet eine mäßige Behinderung bei noch uneingeschränkter Gehfähigkeit. Bei einem EDSS-Wert von 4,0 hingegen liegt bereits eine relativ schwere Behinderung vor mit einer Gehfähigkeit ohne Hilfsmittel von nur 500 Metern. Eine weitere kontinuierliche Verschlechterung beeinflusst zunehmend den Alltag und die Lebensqualität der SPMS-Patienten (14).

Die Erhebung des EDSS-Werts in der Studie EXPAND erfolgte während der Screening-Visite, zu Baseline (Studienbeginn) und danach alle drei Monate während der Studie sowie bei einem Krankheitsschub. Die Progression der Behinderung war definiert als Erhöhung des EDSS-Werts um einen Punkt für Patienten mit einem EDSS-Wert bei Studienbeginn von 3,0 bis 5,0 oder um 0,5 Punkte für Patienten mit einem Baseline-EDSS-Wert von 5,5 bis 6,5 Punkten. Eine bestätigte Progression der Behinderung (3- oder 6-month-confirmed disability progression, 3- oder 6-CDP) war definiert als eine drei bzw. sechs Monate nach dem ersten Anstieg fortdauernde Erhöhung des EDSS-Werts im Vergleich zu Studienbeginn. Außerdem musste jeder EDSS-Punktwert, der während der drei bzw. sechs Monate ermittelt wurde, ebenfalls die Kriterien für eine Progression erfüllen. EDSS-Werte, die während eines Schubes erhoben wurden, fanden keine Berücksichtigung. Zu diesem Zwecke wurde die maximale Schubdauer auf 90 Tage festgelegt.

Eine Behinderungsprogression ist für die Patienten unmittelbar spürbar und relevant. Sie wurde seitens des G-BA im Beratungsgespräch als patientenrelevanter Endpunkt definiert (4).

Behinderungsprogression (MSFC)

Der Schweregrad der Behinderungsprogression wurde zusätzlich anhand des MSFC (Multiple Sclerosis Functional Composite) erhoben. Der MSFC ist eine validierte neurologische Leistungsskala zur objektiven quantitativen Beurteilung des Schweregrads der Behinderungen bei MS-Patienten (15, 16). Die Leistung wird in drei klinischen Dimensionen gemessen:

- Die Beinfunktion mittels des Timed 25-foot Walking Test (T25W).
  Beim T25W wird die Zeit in Sekunden zur Bewältigung einer Gehstrecke von 25 Fuß gemessen (entspricht 7,62 Meter).
- Die Armfunktion mittels des Nine Hole Peg Test (9-HPT).
  Der 9-HPT ist ein Stecktest, bei dem neun Stifte in ein Steckbrett gesteckt und anschließend wieder entfernt werden müssen. Der Test ist zuerst zweimal mit dem dominanten Arm und danach zweimal mit dem nicht dominanten Arm durchzuführen. Es wird die Zeit in Sekunden pro Durchlauf gemessen und daraus der Durchschnittswert ermittelt.
- Die kognitive Funktion mittels des Paced Auditory Serial Addition Test (PASAT-3).
  Der PASAT-3 erhebt die Geschwindigkeit und Flexibilität bei der Verarbeitung von Lautinformationen sowie das Rechenvermögen. Dem Patienten wird alle drei Sekunden eine neue Zahl vorgegeben, die auf die unmittelbar vorhergehende addiert werden muss. Das Ergebnis ist die Anzahl richtiger Antworten innerhalb von drei Minuten (mögliche Spanne von 0 bis 60). Bei den Visiten kamen abwechselnd zwei verschiedene Aufgabensätze (Aufgabensatz A und B) zum Einsatz. Der PASAT-3 wurde in der jeweiligen Landessprache implementiert.

Beim T25W und 9-HPT zeigt eine Abnahme der Werte eine Verbesserung in der betreffenden Komponente an. Beim PASAT-3 bedeuten zunehmende Werte eine Verbesserung. Die Ergebnisse des T25W, 9-HPT und PASAT-3 werden in Form des MSFC-z zusammengefasst, um die durchschnittliche Änderung in allen drei Tests abzubilden. Beim MSFC-z zeigen höhere Werte ebenfalls eine Verbesserung an.

Der MSFC wurde zu Baseline und dann alle drei Monate (T25W und 9-HPT) bzw. alle sechs Monate (PASAT-3) erhoben. Der MSFC-z konnte dementsprechend zur Baseline und dann alle sechs Monate ermittelt werden.

Der MSFC bzw. dessen einzelne Komponenten erfassen typische Symptome der MS-bedingten Behinderung, nämlich motorische und kognitive Defizite. Diese Einschränkungen sind für den Patienten direkt spürbar. Der Endpunkt *Behinderungsprogression* gemessen anhand des MSFC wurde seitens des G-BA im Beratungsgespräch außerdem als patientenrelevanter Endpunkt definiert (4).

Schweregrad der kognitiven Dysfunktion (SDMT)

Beim SDMT (Symbol Digit Modalities Test) handelt es sich um einen sensitiven und spezifischen Test zur Erfassung von Aufmerksamkeit und Konzentration, genauer gesagt der Informationsverarbeitungsgeschwindigkeit. Diese ist die bei MS-Patienten am häufigsten

eingeschränkte kognitive Funktion (17). Einschränkungen der kognitiven Leistungsfähigkeit haben bedeutende Auswirkungen auf das Leben der Patienten und beeinträchtigen die Alltagskompetenz und die Arbeitsfähigkeit mitunter sehr stark (18).

Der SDMT bietet ausreichende Validität, Reliabilität, Ansprechempfindlichkeit und ist einfach in der Anwendung (19). Er hat sich in den letzten Jahren zum meist eingesetzten Test der Informationsverarbeitungsgeschwindigkeit bei MS-Patienten entwickelt (20). Der SDMT wird mittlerweile auch von der US-amerikanischen National MS Society zum regelmäßigen Screening von MS-Patienten auf kognitive Dysfunktionen empfohlen (21).

Beim SDMT wird den Patienten ein Testbogen vorgelegt, auf dem sich oben eine Reihe mit neun Zahlen und dazu zugeordneten geometrischen Figuren befindet. Darunter sind ungepaarte Zahlen bzw. Symbole angeordnet. Die Patienten müssen jeweils einer Zahl das dazugehörige Symbol bzw. einem Symbol die dazugehörige Zahl zuordnen. Gemessen wird die Anzahl der korrekten Antworten, die in 90 Sekunden gegeben werden. Beim SDMT zeigt eine Zunahme der Werte eine Verbesserung an. Eine Veränderung um vier oder mehr Punkte stellt eine klinisch relevante Veränderung der kognitiven Leistungsfähigkeit dar, da ab diesem Wert eine Korrelation mit der Arbeitsfähigkeit besteht (19).

Der SDMT wurde zu Baseline und dann alle sechs Monate erhoben. Der SDMT erfasst kognitive Einschränkungen, die als typische Symptome der MS-bedingten Behinderung gelten. Diese kognitiven Einschränkungen sind für den Patienten direkt spürbar. Der G-BA hat außerdem die kognitive Dysfunktion als patientenrelevant definiert (4). Der Endpunkt *Schweregrad der kognitiven Dysfunktion* gemessen anhand des SDMT, ist daher als patientenrelevant einzustufen.

Schweregrad der kognitiven Dysfunktion (BVMT-R)

Der BVMT-R (Brief Visuospatial Memory Test – Revised) erfasst das visuell-räumliche Erinnerungsvermögen (22). In drei aufeinanderfolgenden Lerntests wird den Patienten jeweils zehn Sekunden lang ein Testbogen mit geometrischen Figuren gezeigt. Im Anschluss sollen sie die Figuren so genau wie möglich nachzeichnen, und zwar an der Stelle, wo sie zu sehen waren. Nach 25 Minuten werden die Patienten in einem weiteren Erinnerungstest nochmals aufgefordert, die Figuren nachzuzeichnen (Delayed Recall Trial). Gemessen wird die Anzahl der richtigen Antworten im Erinnerungstest. Beim BVMT-R zeigt eine Zunahme der Werte eine Verbesserung an. Der sogenannte Recognition Trial, ein Wiedererkennungstest, bei dem der Proband aus zwölf Figuren diejenigen benennen soll, die unter den ursprünglich gezeigten geometrischen Figuren waren, kam in der EXPAND-Studie nicht zum Einsatz.

Der BVMT-R wurde zu Baseline und dann alle sechs Monate erhoben. Der BVMT-R erfasst kognitive Einschränkungen, die als typische Symptome der MS-bedingten Behinderung gelten. Die kognitive Dysfunktion ist für den Patienten direkt spürbar und wurde vom G-BA als patientenrelevant eingestuft (4). Der Endpunkt *Schweregrad der kognitiven Dysfunktion* gemessen anhand des BVMT-R ist daher als patientenrelevant einzustufen.

Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)

Bei MS-Patienten sind Sehstörungen häufig, jedoch fehlt diese Komponente in Tests wie dem MSFC. Um auch Einschränkungen im Sehvermögen zu erfassen, wurde in die EXPAND-Studie der LCVA-Test (Low Contrast Visual Acuity) aufgenommen. Die MS-typischen Beeinträchtigungen sind bei Sehtests mit hohem Kontrast oft nicht gut messbar. Der Einsatz von LCVA-Tafeln hat sich hingegen zu einem validen Instrument zur Bestimmung der Sehschärfe bei MS-Patienten entwickelt (23). Mit den sogenannten Sloan-Buchstaben-Sehprobentafeln mit geringem Kontrast ist es möglich, die Sehleistung quantitativ und standardisiert zu ermitteln. In der EXPAND-Studie kam die Tafel mit einem Kontrast von 2,5 % zum Einsatz. Auf jeder Tafel befinden sich auf weißem Hintergrund Reihen mit grauen Buchstaben, die von oben nach unten kleiner werden. Gemessen wird die Anzahl der richtig erkannten Buchstaben. Beim LCVA zeigt eine Zunahme der Werte eine Verbesserung an.

Der LCVA wurde zu Baseline und dann alle sechs Monate erhoben. Der LCVA erfasst Einschränkungen der Sehfunktion. Diese sind ein typisches Symptom der MS und für den Patienten direkt spürbar. Der Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens*, gemessen anhand des LCVA, ist daher als patientenrelevant einzustufen.

Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)

Die Multiple Sclerosis Walking Scale (MSWS-12) ist ein validiertes Messinstrument mit zwölf Fragen zur Beantwortung durch den Patienten, mit dem die Auswirkung der MS-Erkrankung auf die Gehfähigkeit erhoben wird (24, 25). In der Version 2 des Fragebogens können pro Frage 1 bis 3 bzw. 1 bis 5 Punkte erzielt werden (26). Dabei stehen höhere Werte für eine stärkere Beeinträchtigung. Die in den verschiedenen Fragen erzielten Punkte werden summiert und in Werte von 0 bis 100 umgerechnet (27).

Die MSWS-12 war vom Patienten selbständig zu Baseline und danach alle sechs Monate jeweils vor den anderen zu dieser Visite stattfindenden Untersuchungen auszufüllen. Die durch die MS erworbene Behinderung äußert sich oftmals in Form einer eingeschränkten Gehfähigkeit. Mit der MSWS-12 werden die durch den Patienten wahrgenommenen Beeinträchtigungen erfasst, der Endpunkt wird daher als patientenrelevant eingestuft.

<u>Gesundheitsbezogene Lebensqualität</u>

MSIS-29

Die Multiple Sclerosis Impact Scale (MSIS-29) ist ein validiertes, krankheitsspezifisches Instrument zur patientenberichteten Lebensqualität (28, 29). In der verwendeten Version 2 handelt es sich um einen Fragebogen mit 29 Fragen zur Beantwortung durch den Patienten. Sie umfasst jeweils eine Skala zur physischen bzw. psychischen Beeinträchtigung des täglichen Lebens durch die MS-Erkrankung. Die Antworten werden auf einer Vierpunkte-Skala erfasst, sie reicht von ‚gar nicht' beeinträchtigt (codiert als 1) bis ‚sehr' beeinträchtigt. Höhere Werte stehen also für eine stärkere Beeinträchtigung des Alltagslebens.

EQ-5D

Der EQ-5D ist ein krankheitsübergreifender, validierter Fragebogen zur präferenzbasierten Erfassung der gesundheitsbezogenen Lebensqualität (30). Der EQ-5D besteht aus zwei Teilen: EQ-5D Health Classification State und EQ-5D Visual Analogue Scale (VAS).

Die EQ-5D Health Classification State (Descriptive System) umfasst fünf Dimensionen: Beweglichkeit/Mobilität; Für-sich-selbst-sorgen; allgemeine Tätigkeiten; Schmerzen / körperliche Beschwerden; Angst/Niedergeschlagenheit. Die Dimensionen weisen jeweils drei Antwortmöglichkeiten (‚keine', ‚mäßige' und ‚extreme Probleme') auf. Insgesamt lassen sich 243 Gesundheitszustände durch Kombination der Antwortmöglichkeiten über die fünf Dimensionen abbilden. Die Gesundheitszustände wurden unter Verwendung des EQ-5D-Value-Sets für das Vereinigte Königreich in einen Lebensqualitätsindex zwischen -0,594 und 1,0 transformiert (Tod = 0; bestmöglicher Gesundheitszustand = 1). Einige Gesundheitszustände können demnach auch Werte von kleiner als 0 annehmen, d.h. sie werden aus Sicht des Patienten als schlimmer als der Tod bewertet (31). Nach Walters (2005) liegt eine klinisch bedeutsame Verbesserung bei einer Steigerung um durchschnittlich 0,074 Punkte im Indexwert vor (31).

Als zweiter Teil in der Erfassung der gesundheitsbezogenen Lebensqualität mittels EQ-5D wird die visuelle Analogskala (VAS) verwendet. In dieser schätzen die Patienten ihren momentanen Gesundheitszustand zwischen 0 (schlechtestmöglicher momentaner Gesundheitszustand) und 100 (bestmöglicher momentaner Gesundheitszustand) ein. Die VAS erlaubt eine allgemeine Bewertung des eigenen momentanen Gesundheitszustands, unabhängig von den Antworten auf den fünf Dimensionen des Fragebogens. Eine klinische Relevanzschwelle ist nicht etabliert.

Die gesundheitsbezogene Lebensqualität gilt als patientenrelevanter Endpunkt. Die Fragebögen hierzu waren vom Patienten selbständig zu Baseline und danach alle sechs Monate jeweils vor den anderen zu dieser Visite stattfindenden Untersuchungen auszufüllen.

Sicherheit und Verträglichkeit

Im vorliegenden Dossier werden jeweils die Häufigkeit sowie die Zeit bis zum ersten Ereignis berichtet für: Unerwünschte Ereignisse (UE), Schwerwiegende Unerwünschte Ereignisse (SUE), Schwere UE und UE von besonderem Interesse (gemäß Studienbericht) sowie Studien- bzw. Behandlungsabbrüche aufgrund von UE.

Zusätzlich werden die Ergebnisse der Columbia-Beurteilungsskala zur Suizidalität (Columbia Suicide Severity Rating Scale; C-SSRS) berichtet. Hierbei handelt es sich um einen validierten Fragebogen zur Erfassung suizidaler Gedanken und suizidalen Verhaltens (32). Der Fragebogen umfasst die Kategorien Suizidgedanken, suizidales Verhalten und Selbstverletzung ohne suizidale Absicht. Alle lebensbedrohenden Ereignisse, also z. B. ‚Ja'-Antworten im Fragenteil zum suizidalen Verhalten, wurden in der Studie auch als SUE berichtet. Berichtete Selbstverletzungen ohne suizidale Absicht, die die UE-Kriterien erfüllten, wurden in der Studie außerdem als UE erfasst.

Bildgebung

Läsionslast und Hirnvolumen

Aufgrund einer im Verlaufe der Erkrankung aufgetretenen Schädigung der Blut-Hirn-Schranke ist die akute entzündliche Aktivität im MRT als Gd-Anreicherung darstellbar. Aus dieser inflammatorischen Aktivität, bildgebend gezeigt als Gd-anreichernde T1-Läsion, gehen weiter differenzierbare Läsionstypen hervor, z. B. neue T2-Läsionen. Praktisch alle T2-Läsionen treten zusammen mit Gd-anreichernden Läsionen auf oder entwickeln sich aus diesen (33). Das Auftreten von Gd-anreichernden Läsionen oder neuen oder sich vergrößernden T2-Läsionen gilt als Kriterium einer aktiven Erkrankung. Zudem ist das Ausmaß einer MS-bedingten Hirnatrophie ein klinisch relevanter Prädiktionsfaktor für die Progression der kognitiven und physischen Behinderung (34-39). In der EXPAND-Studie wurde zu Studienbeginn sowie danach jährlich jeweils eine MRT-Untersuchung durchgeführt.

Die Läsionslast ist für die Definition der Aktivität der SPMS-Erkrankung und damit auch für die Definition der Zielpopulation von Bedeutuiung. Läsionslast und Hirnatrophie gelten außerdem als Surrogat für das Fortschreiten der Erkrankung und sind zentrale Parameter in der Überwachung des Therapieerfolgs in der klinischen Praxis. Daher werden sowohl die Läsionslast als auch die Veränderung des Hirnvolumens ergänzend im Dossier dargestellt. Eine Surrogatvalidierung ist nicht erfolgt.

Auswertungszeitpunkte

Die Kernphase der EXPAND wurde ursprünglich mit einer festgelegten Dauer von 24 Monaten geplant. Per Amendment wurde diese auf eine flexible Dauer von bis zu 24 Monaten geändert.

Die Wirksamkeitsendpunkte wurden daher wie folgt ausgewertet:

- Studienende (End of study, EOS) für alle Endpunkte, sofern nicht anders angegeben. Es gehen alle Patienten mit ihrer individuellen Behandlungsdauer ein.
- 24 Monate zusätzlich für alle Ereigniszeitanalysen (Anteil der Patienten ohne Ereignis zu Monat 24).
- 12 bzw. 24 Monate für alle Auswertungen zu Veränderungen seit Studienbeginn wie angegeben.
- 12 bzw. 24 Monate für patientenberichtete Endpunkte. Zusätzlich werden hier gemäß der neuen Dossier-Vorlage adjustierte Mittelwerte für alle erhobenen Zeitpunkte deskriptiv dargestellt.

Statistik

Die für das Dossier angefertigten Analysen wurden für die ITT-Population (FAS) durchgeführt. Die Ergebnisse der Studie werden durch adäquate Effektschätzer, korrespondierende Konfidenzintervalle (KI) und p-Werte dargestellt. Wenn nicht anders angegeben, werden das

95%-KI und p-Werte geeigneter (zweiseitiger) statistischer Tests dargestellt. Obwohl sich durch die Vielzahl an post-hoc durchgeführten statistischen Tests ein multiples Testproblem ergibt, wurde keine Adjustierung des Signifikanzniveaus vorgenommen. Bei einem p-Wert < 0,05 wird eine statistische Signifikanz angenommen.

Für kategorielle Variablen werden zur Beschreibung des Auftretens eines Ereignisses das Relative Risiko (RR), das Odds Ratio (OR) und die Risikodifferenz (RD) als Risikoschätzer angegeben. Das OR wurde dabei basierend auf logistischen Regressionsmodellen berechnet. Für das RR und die RD ergaben sich Intervallschätzer und p-Wert basierend auf einer Approximation durch die Normalverteilung.

Für die Analyse von Ereigniszeiten werden Kaplan-Meier-Schätzer für die Anteile der Patienten mit bzw. ohne Ereignis und die über ein Cox-Regressionsmodell berechnete Hazard-Ratio (HR) angegeben. Zudem wird die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko (zu mehreren Zeitpunkten) abgebildet.

Der Vergleich von (jährlichen) Raten erfolgt über Rate Ratios, die inklusive 95%-KI und p-Wert basierend auf einer Negativ-Binomialregression geschätzt wurden.

Der Vergleich der Anzahl neuer oder sich vergrößernder T2-Läsionen erfolgt relativ zum vorherigen Zeitpunkt basierend auf einer Negativ-Binomialregression, inklusive 95%-KI und p-Wert.

Für kontinuierliche Variablen werden die mittleren Werte zu Studienbeginn, die mittlere adjustierte Veränderung mit Standardfehler sowie die adjustierte Mittelwertdifferenz mit dazugehörigem Konfidenzintervall und p-Wert angegeben. Es wurde ein Mixed Model for Repeated Measures (MMRM) zugrunde gelegt, in das für jeden Patienten die Werte jeweils aller vorhandener Visiten eingingen

Kovariaten in Modellen

In den Regressionsmodellen wurden analog zu den im CSR (Clinical Study Report) dargestellten Analysen jeweils geeignete Kovariablen berücksichtigt. Die genaue Beschreibung der jeweils eingeschlossenen Variablen findet sich in den Fußnoten zu den jeweiligen Quelltabellen im Anhang zu diesem Modul 4.

Sofern nicht anders angegeben, entstammen alle dargestellten Ergebnisse den Zusatzanalysen zur EXPAND-Studie (siehe Anhang 1 und Anhang 2 zu Modul 4A).

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 nach der Knapp-Hartung-Methode mit der Paule-Mandel-Methode zur Heterogenitätsschätzung[5] erfolgen. Im Fall von sehr wenigen Studien ist die Heterogenität nicht verlässlich schätzbar. Liegen daher weniger als 5 Studien vor, ist auch die Anwendung eines Modells mit festem Effekt oder eine qualitative Zusammenfassung in Betracht zu ziehen. Kontextabhängig können auch alternative Verfahren wie z. B. Bayes'sche Verfahren oder Methoden aus dem Bereich der generalisierten linearen Modelle in Erwägung gezogen 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[6] 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 geeigneter statistische Maße auf Vorliegen von Heterogenität[7, 5] erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik $\geq 0{,}05$), 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).

[5] Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. Cochrane Database Syst Rev 2015: 25-27.
[6] 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.
[7] Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-560.

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

Es waren keine Meta-Analysen erforderlich, da nur eine relevante RCT identifiziert wurde (EXPAND-Studie), in der Siponimod gegen BSC geprüft wurde.

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.

Die im Dossier dargestellten Ergebnisse zur Gesamtpopulation der Studie können zur Einschätzung der Robustheit der Ergebnisse der Nutzenbewertung herangezogen werden.

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 (zu ergänzenden Kriterien zur Darstellung siehe Abschnitt 4.3.1.3.13).

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 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. 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.

Für das vorliegende Dossier wurden alle Endpunkte aus dem direkten Vergleich in den folgenden Subgruppen analysiert:

Präspezifizierte Subgruppen:
- Geschlecht (männlich; weiblich),
- Vorherige krankheitsmodifizierende Therapie (ja; nein),
- Vorherige Behandlung mit IFN-beta 1b (ja; nein),
- Schnell fortschreitende Erkrankung, d. h. EDSS-Anstieg um mindestens 1,5 Punkte in den zwei Jahren vor oder zu Studienbeginn (ja; nein),
- Moderater oder schwerer Erkrankungsverlauf, d. h. Wert von ≥ 4 auf der MS Status Scale (ja; nein).

Nachträglich definierte Subgruppen:
- Alter (18 – 30 Jahre; 31 – 40 Jahre bzw. 18 – 40 Jahre; 41 – 55 Jahre; > 55 Jahre),
- Region (Europa; andere).

Im Studienbericht definierte, aber im Dossier nicht aufgenommene Subgruppen:
- Vorhandensein aufgesetzter Schübe in den zwei Jahren vor Studienbeginn (ja; nein),
- Vorhandensein Gd-anreichernder Läsionen (ja; nein),
- Aufgesetzte Schübe während der Studie (ja; nein).

Es wurden die für den Studienbericht präspezifizierten Subgruppenanalysen herangezogen. Die Trennpunkte für die präspezifizierten Subgruppen wurden analog zum Studienbericht gewählt.

Zusätzlich wurden Auswertungen nach Alter und Region aufgenommen, um die Anforderungen der Verfahrensordnung zu erfüllen. Als Trennpunkte für die Subgruppenauswertung nach Alter wurde die Aufteilung der Altersgruppen entsprechend der Studienberichtsauswertung zu den Patientencharakteristika übernommen. Die Alterskategorien 18 – 30 und 31 – 40 Jahre wurden jedoch für die Auswertung der Subpopulation B zusammengefasst, da die unterste Altersklasse zu gering besetzt war. Für die Subgruppenauswertung nach Region wurde unterteilt in Europa und andere Regionen. Dafür wurden andere Regionen als Europa als eine Kategorie zusammengefasst, da die einzelnen Regionen zu gering besetzt waren und eine getrennte Auswertung nicht möglich war. Die Zuordnung erfolgte nach Definition der Region Europa durch die Vereinten Nationen (40).

Tabelle 4-13: Zuordnung der an der EXPAND-Studie teilnehmenden Ländern nach Regionen

Land	**UN-Region**	**Subgruppe**
Argentinien	Amerika	Andere
Australien	Ozeanien	Andere
Belgien	Europa	Europa
Bulgarien	Europa	Europa
China	Asien	Andere
Deutschland	Europa	Europa
Estland	Europa	Europa
Frankreich	Europa	Europa
Griechenland	Europa	Europa
Großbritannien	Europa	Europa
Irland	Europa	Europa
Israel	Asien	Andere
Italien	Europa	Europa
Japan	Asien	Andere
Kanada	Amerika	Andere
Lettland	Europa	Europa
Litauen	Europa	Europa
Niederlande	Europa	Europa
Österreich	Europa	Europa
Polen	Europa	Europa
Portugal	Europa	Europa
Rumänien	Europa	Europa
Russland	Europa	Europa
Schweden	Europa	Europa
Schweiz	Europa	Europa
Slowakei	Europa	Europa
Spanien	Europa	Europa
Tschechien	Europa	Europa
Türkei	Asien	Andere
Ungarn	Europa	Europa
USA	Amerika	Andere

Die Subgruppenanalysen zum Vorhandensein aufgesetzter Schübe in den vergangenen zwei Jahren und zum Vorhandensein Gd-anreichernder Läsionen wurden im Dossier nicht aufgegriffen. Dies begründet sich wie folgt: Die Subpopulation B umfasst die Patienten, die

keine aufgesetzten Schübe in den vergangenen zwei Jahren gezeigt hatten und die mindestens eine Gd-anreichernde Läsion aufwiesen. Die Kriterien für diese Analyse sind also über die Auswertung der Subpopulation B bereits abgedeckt. Eine Auswertung entsprechend dieser Subgruppen in der Subpopulation B ist nicht möglich, da in der Kategorie ‚Schübe: ja' bzw. ‚Gd-anreichernde Läsionen: nein' definitionsgemäß jeweils keine Patienten sind.

Die Subgruppenanalyse nach dem Kriterium des Auftretens von Schüben während der Studie wurde nicht aufgenommen, da eine Subgruppendefinition über ein Ereignis während der Studie durch die Behandlung beeinflusst und damit nicht sinnvoll interpretierbar ist.

Subgruppenanalysen für kategorielle Variablen wurden nur durchgeführt, wenn alle resultierenden Subgruppen jeweils mindestens zehn Patienten umfassten. Außerdem erfolgten Subgruppenanalysen für binäre Ereignisse nur, wenn in einer der Subgruppen mindestens zehn Ereignisse aufgetreten waren. Für Unerwünschte Ereignisse nach SOC und PT erfolgten Subgruppenanalysen generell nur, wenn das Ergebnis in der Gesamtpopulation bzw. in der gesamten Subpopulation B statistisch signifikant war.

Mögliche Effektmodifikationen wurden anhand von Interaktionstermen untersucht. Dazu wurden die den Analysen zugrundeliegenden Modelle im Rahmen der Subgruppenanalyse um die entsprechende Subgruppenvariable und den Interaktionsterm Behandlung × Subgruppenvariable erweitert. Ein p-Wert von < 0,05 zeigt einen Beleg für eine Interaktion an.

Ergebnisse wurden nur für diejenigen Subgruppenanalysen dargestellt, bei denen der Interaktionstest einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Subgruppenmerkmal ergab. Die Ergebnisse aller Subgruppenanalysen sind den separaten Anhängen zu Modul 4 zu entnehmen.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Es besteht dabei internationaler Konsens, dass Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator (häufig als nicht adjustierte indirekte Vergleiche bezeichnet) regelhaft keine valide Analysemethode darstellen[8]. Eine Ausnahme kann das Vorliegen von dramatischen Effekten sein. An Stelle von nicht adjustierten indirekten Vergleichen sollen je nach Datenlage einfache adjustierte indirekte Vergleiche[9] oder komplexere Netzwerk-Meta-Analysen (auch als „Mixed Treatment Comparison [MTC] Meta-Analysen" oder „Multiple Treatment Meta-Analysen"

[8] Bender R, Schwenke C, Schmoor C, Hauschke D. Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR [online]. [Zugriff: 31.10.2016]. URL: http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf.
[9] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997; 50(6): 683-691.

bezeichnet) für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche berechnet werden. Aktuelle Verfahren wurden beispielsweise von Lu und Ades (2004)[10] und Rücker (2012)[11] vorgestellt.

Alle Verfahren für indirekte Vergleiche gehen im Prinzip von den gleichen zentralen Annahmen aus. Hierbei handelt es sich um die Annahmen der Ähnlichkeit der eingeschlossenen Studien, der Homogenität der paarweisen Vergleiche und der Konsistenz zwischen direkter und indirekter Evidenz 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[12].

Das Ergebnis eines indirekten Vergleichs kann maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen. Als Brückenkomparatoren sind dabei insbesondere Interventionen zu berücksichtigen, für die sowohl zum bewertenden Arzneimittel als auch zur zweckmäßigen Vergleichstherapie mindestens eine direkt vergleichende Studie vorliegt (Brückenkomparatoren ersten Grades).

Insgesamt ist es notwendig, die zugrunde liegende Methodik für alle relevanten Endpunkte genau und reproduzierbar zu beschreiben und die zentralen Annahmen zu untersuchen[13, 14, 15]

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

- *Benennung aller potentiellen Brückenkomparatoren ersten Grades und ggf. Begründung für die Auswahl.*

- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayes'schen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten, die Art der Untersuchung der Konvergenz der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*

[10] Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004; 23(20): 3105-3124.
[11] Rücker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods 2012; 3(4): 312-324.
[12] Schöttker B, Lühmann D, Boulkhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.
[13] 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.
[14] Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study BMJ 2011; 343 :d4909
[15] Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. Res Synth Methods 2013; 4(4): 291-323.

- *Art der Prüfung der Ähnlichkeit der eingeschlossenen Studien.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenzannahme im Netzwerk.*
- *Bilden Sie den Code des Computerprogramms inklusive der einzulesenden Daten 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.*

Für die Subpopulationen war es aus folgenden Gründen nicht möglich, einen indirekten Vergleich durchzuführen:

Subpopulation A:

Die Subpopulation A umfasst Patienten mit SPMS mit einer aktiven Erkrankung definiert durch Schübe. Zur Identifikation potenziell passender Studien wurde eine unsystematische Literaturrecherche durchgeführt. Die dabei identifizierten Studien zu IFN-beta 1a s.c., IFN-beta 1b s.c., bzw. Ocrelizumab, unterscheiden sich in der untersuchten Patientenpopulation, sodass für die in Frage kommende Teilpopulation der jeweiligen zweckmäßigen Vergleichstherapien keine Daten für einen indirekten Vergleich identifiziert werden konnten.

Subpopulation B:

Die Subpopulation B umfasst alle SPMS-Patienten mit aktiver Erkrankung definiert durch bildgebende Verfahren. Für diese Subpopulation steht mit der EXPAND-Studie eine randomisierte kontrollierte Studie gegen die zweckmäßige Vergleichstherapie Best Supportive Care zur Verfügung, weshalb ein indirekter Vergleich nicht geprüft wurde.

Es wurden daher keine indirekten Vergleiche zur Bewertung des Zusatznutzens in den relevanten Subpopulationen vorgenommen.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

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

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

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

Falls für die Bewertung des Zusatznutzens mehrere Komparatoren (z.B. Wirkstoffe) herangezogen werden, sind die Aussagen zum Zusatznutzen primär gegenüber der Gesamtheit der gewählten Komparatoren durchzuführen (z. B. basierend auf Meta-Analysen unter gemeinsamer Betrachtung aller direkt vergleichender Studien). Spezifische methodische Argumente, die gegen eine gemeinsame Analyse sprechen (z. B. statistische oder inhaltliche Heterogenität), sind davon unbenommen. Eine zusammenfassende Aussage zum Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie ist in jedem Fall erforderlich.

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, Angabe zu geplanten und durchgeführten Datenschnitten und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-14: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	**Zulassungsstudie (ja/nein)**	**Sponsor (ja/nein)**	**Status (abgeschlossen / abgebrochen / laufend)**	**Studiendauer, ggf. Datenschnitt**	**Therapiearme**
EXPAND	ja	ja	abgeschlossen[a]	bis zu 24 Monate[a]	Siponimod; BSC
a: Die doppelblinde Behandlungsphase (Kernphase) wurde abgeschlossen; diese ist Gegenstand der Nutzenbewertung; die einarmige, offene Extensionsstudie läuft zum Zeitpunkt der Einreichung des Dossiers noch; diese ist aber nicht relevant für die Nutzenbewertung.					

Geben Sie an, welchen Stand die Information in Tabelle 4-14 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 obige Tabelle bildet den Studienstatus zum 1. Dezember 2019 ab.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-14 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-15: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	**Begründung für die Nichtberücksichtigung der Studie**
Subpopulation A	
EXPAND	andere Vergleichstherapie
Subpopulation B	
keine	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.

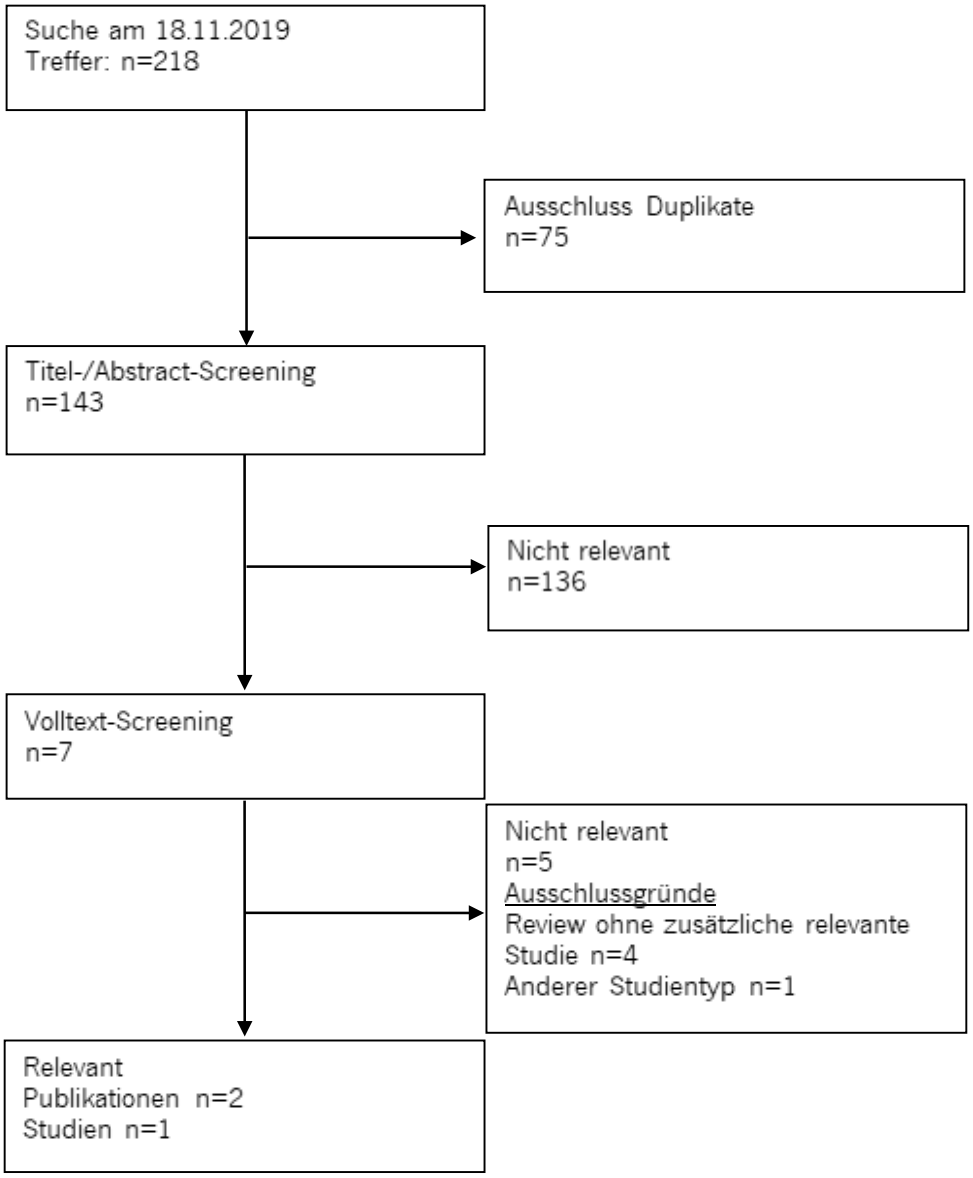


Abbildung 1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in welchem Studienregister / Studienergebnisdatenbank 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 4-14) 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 4-16: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank und Angabe der Zitate[a])	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Status (abgeschlossen/ abgebrochen/ laufend)
EXPAND	**clinicaltrials.gov** NCT01665144 (41)	ja	ja	abgeschlossen
	ICTRP NCT01665144 (42) JPRN-JapicCTI-142447 (43) EUCTR2012-003056-36-HU (44)[b] EUCTR2012-003056-36-SK (45)[b]			
	EU-CTR EudraCT 2012-003056-36 (46)			
	Clinical Data (EMA) (Keine Einträge)			
	AMIS (Keine Einträge)			

a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister / in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder –ergebnisse.

b: Bei der Recherche wurden zu diesem EU-CTR-Code die Ländereinträge 2012-003056-36-SK und 2012-003056-36-HU als separate Studien identifiziert.

Geben Sie an, welchen Stand die Information in Tabelle 4-16 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 obige Tabelle bildet den Studienstatus zum 25. November 2019 ab.

4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-14) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/ Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

[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 4-17: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	**Relevante Quellen**[a]	**Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)**	**Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)**	**Studie durch Suche in Studienregistern / Studienergebnis datenbanken identifiziert (ja/nein)**
keine				
a: Quellen aus der Suche auf der Internetseite des G-BA				

Geben Sie an, welchen Stand die Information in Tabelle 4-17 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 obige Tabelle bildet den Studienstatus zum 25. November 2019 ab.

4.3.1.1.5 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, 4.3.1.1.3 und 4.3.1.1.4) 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 aller verfügbaren Quellen in Abschnitt 4.3.1.4 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 4-18: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen[a]		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	**Gesponserte Studie[b] (ja/nein)**	**Studie Dritter (ja/nein)**	**Studienberichte (ja/nein [Zitat])**	**Registereinträge[c] (ja/nein [Zitat])**	**Publikation und sonstige Quellen[d] (ja/nein [Zitat])**
Subpopulation A						
keine						
Subpopulation B						
EXPAND	ja	ja	nein	ja (9)	ja (41-46)	ja (47, 48)

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 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (Liste der eingeschlossenen Studien) aufgeführt werden.

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.

d: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.

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. Falls Teilpopulationen berücksichtigt werden, ist die Charakterisierung der Studienpopulation auch für diese Teilpopulation durchzuführen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Geben Sie bei den Datenschnitten auch den Anlass des Datenschnittes an. Fügen Sie für jede Studie eine neue Zeile ein.

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

Tabelle 4-19: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	**Studiendesign**	**Population**	**Interventionen (Zahl der randomisierten Patienten)**	**Studiendauer/ Datenschnitte**	**Ort und Zeitraum der Durchführung**	**Primärer Endpunkt; patientenrelevante sekundäre Endpunkte**
EXPAND	RCT, doppelblind, parallel	Erwachsene Patienten mit SPMS (18 bis 60 Jahre)	Siponimod: N = 1.099 BSC: N = 546	Bis zu 24 Monate[a]	Weltweit in 31 Ländern mit 294 Studienzentren	Primärer Endpunkt: Progression der Behinderung (EDSS) Weitere Endpunkte: Bestätigte Krankheitsschübe, MSFC, SDMT, BVMT-R, LCVA, MSWS-12, Gesundheitsbezogene Lebensqualität (MSIS-29/EQ-5D), Unerwünschte Ereignisse, Suizidalität, Bildgebung (Läsionslast, Hirnvolumen)
a: Die Kernphase der Studie wurde nach drei Jahren Gesamtlaufzeit beendet, weil dann die erforderliche Anzahl Ereignisse für die Auswertung des primären Endpunkts erreicht war.						

Tabelle 4-20: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Siponimod	BSC	*ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in-Phase etc.*
EXPAND	Siponimod 2 mg[a]	Placebo	keine
a: Die Behandlung mit Siponimod bzw. entsprechendem Placebo begann mit einer Auftitrationsphase von fünf Tagen (Tag 1 und 2: je 0,25 mg; Tag 3: 0,50 mg; Tag 4: 0,75 mg; Tag 5: 1,25 mg; ab Tag 6: volle Dosis von 2 mg). Patienten mit einer Lymphozytenzahl < $0{,}2 \times 10^9$/l bei zwei aufeinanderfolgenden Visiten mussten die Dosis auf 1 mg reduzieren.			

Gesamtpopulation

Tabelle 4-21: Patientendisposition – RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND (RAN)	Siponimod (N = 1.105)	BSC (N = 546)	Gesamt (N = 1.651)
Full Analysis Set, n (%)			
In FAS eingeschlossen	1.099 (99,5)	546 (100,0)	1.645 (99,6)
Grund für Ausschluss aus dem FAS			
Keine Medikation erhalten	5 (0,5)	0 (0,0)	5 (0,3)
Keine Einwilligungserklärung	1 (0,1)	0 (0,0)	1 (0,1)
Studienabbrüche, n (%)			
Studie abgebrochen	202 (18,3)	122 (22,3)	324 (19,6)
Abbruchgründe			
Entscheidung des Patienten/Vormunds	96 (8,7)	77 (14,1)	173 (10,5)
Unerwünschtes Ereignis	45 (4,1)	18 (3,3)	63 (3,8)
Fehlende Wirksamkeit	16 (1,4)	11 (2,0)	27 (1,6)
Entscheidung des Arztes	13 (1,2)	1 (0,2)	14 (0,8)
Keine Nachbeobachtung möglich	9 (0,8)	8 (1,5)	17 (1,0)
Fortschreiten der Erkrankung	8 (0,7)	4 (0,7)	12 (0,7)
Non-Compliance hinsichtlich Medikation	5 (0,5)	0 (0,0)	5 (0,3)
Tod	3 (0,3)	1 (0,2)	4 (0,2)
Protokollabweichung	3 (0,3)	1 (0,2)	4 (0,2)
Neue Behandlung	2 (0,2)	1 (0,2)	3 (0,2)
Technische Probleme	2 (0,2)	0 (0,0)	2 (0,1)

EXPAND (RAN)	Siponimod (N = 1.105)	BSC (N = 546)	Gesamt (N = 1.651)
Medikation, n (%)			
Medikation erhalten	1.100 (99,5)	546 (100,0)	1.646 (99,7)
Keine Medikation erhalten	5 (0,5)	0 (0,0)	5 (0,3)
Doppelblinde Behandlung abgeschlossen	737 (66,7)	322 (59,0)	1.059 (64,1)
Doppelblinde Behandlung vorzeitig abgebrochen	368 (33,3)	224 (41,0)	592 (35,9)
Mit unverblindetem Siponimod weiterbehandelt	116 (10,5)	94 (17,2)	210 (12,7)
Mit verkürztem Visitenschema fortgefahren	64 (5,8)	25 (4,6)	89 (5,4)
Vollständiger Abbruch der Behandlungsphase	188 (17,0)	105 (19,2)	293 (17,7)
Abbruchgründe			
Entscheidung des Patienten/Vormunds	115 (10,4)	71 (13,0)	186 (11,3)
Fortschreiten der Erkrankung	101 (9,1)	81 (14,8)	182 (11,0)
Unerwünschtes Ereignis	94 (8,5)	28 (5,1)	122 (7,4)
Fehlende Wirksamkeit	36 (3,3)	27 (4,9)	63 (3,8)
Entscheidung des Arztes	14 (1,3)	13 (2,4)	27 (1,6)
Abbruch entspricht Protokollvorgaben	3 (0,3)	2 (0,4)	5 (0,3)
Protokollabweichung	3 (0,3)	0 (0,0)	3 (0,2)
Technische Probleme	2 (0,2)	1 (0,2)	3 (0,2)
Dosierungsfehler	0 (0,0)	1 (0,2)	1 (0,1)
Protokollabweichungen, n (%)			
Mindestens eine Protokollabweichung	682 (61,7)	328 (60,1)	1.010 (61,2)
Wichtige Prozeduren nicht protokollgemäß	415 (37,6)	198 (36,3)	613 (37,1)
Andere GCP-Abweichung	296 (26,8)	143 (26,2)	439 (26,6)
Unerlaubte Begleitmedikation	74 (6,7)	38 (7,0)	112 (6,8)
Einschlusskriterien verletzt	99 (9,0)	35 (6,4)	134 (8,1)
Nicht protokollgemäß aus der Studie genommen	3 (0,3)	2 (0,4)	5 (0,3)
Behandlungsabweichung	85 (7,7)	38 (7,0)	123 (7,5)

FAS: Full Analysis Set; GCP: Good Clinical Practice; N: Patienten gesamt; n: Anzahl Patienten in der Kategorie; RAN: Randomisierte Patienten.

Quelle: Anhang 1 zu Modul 4, Tabelle 1.1.

Tabelle 4-22: Beobachtungsdauer – RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND (FAS)	Siponimod (N = 1.099)	BSC (N = 546)	Gesamt (N = 1.645)
Beobachtungsdauer			
MW ± SD (Tage)	629,6 ± 237,3	628,7 ± 227,3	629,3 ± 233,9
Median (Tage)	638,0	640,5	639,0
Minimum; Maximum (Tage)	7; 1.109	16; 1.087	7; 1.109
Patienten ≥ 365 Tage in der Studie, n (%)	969 (88,2)	480 (87,9)	1.449 (88,1)
Patienten < 365 Tage in der Studie, n (%)	130 (11,8)	66 (12,1)	196 (11,9)
Patienten ≥ 730 Tage in der Studie, n (%)	399 (36,3)	195 (35,7)	594 (36,1)
Patienten < 730 Tage in der Studie, n (%)	700 (63,7)	351 (64,3)	1.051 (63,9)

FAS: Full Analysis Set; MW: Mittelwert; N: Patienten gesamt; n: Anzahl Patienten in der Kategorie; SD: Standard Deviation.
Quelle: Anhang 1 zu Modul 4, Tabelle 1.2.

Tabelle 4-23: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Demographie, Gesamtpopulation)

EXPAND (FAS)	Siponimod (N = 1.099)	BSC (N = 546)	Gesamt (N = 1.645)
Alter (Jahre)			
MW ± SD	48,0 ± 7,8	48,1 ± 7,9	48,0 ± 7,9
Median	49,0	49,0	49,0
Spanne	22 – 61	21 – 61	21 – 61
Altersgruppen (Jahre), n (%)			
18 – 30	26 (2,4)	12 (2,2)	38 (2,3)
31 – 40	161 (14,6)	91 (16,7)	252 (15,3)
41 – 55	712 (64,8)	331 (60,6)	1.043 (63,4)
> 55	200 (18,2)	112 (20,5)	312 (19,0)
Geschlecht, n (%)			
Weiblich	664 (60,4)	323 (59,2)	987 (60,0)
Männlich	435 (39,6)	223 (40,8)	658 (40,0)
Gewicht (kg)			
MW ± SD	71,6 ± 15,7	71,5 ± 16,0	71,5 ± 15,8
Median	70,0	70,0	70,0
Spanne	40,0 – 142,1	38,5 – 145,3	38,5 – 145,3

EXPAND (FAS)	Siponimod (N = 1.099)	BSC (N = 546)	Gesamt (N = 1.645)
Größe (cm)			
MW ± SD	169,3 ± 9,4	169,4 ± 9,9	169,3 ± 9,5
Median	169,0	170,0	169,0
Spanne	142,2 – 196,0	145,0 – 198,0	142,2 – 198,0
BMI (kg/m²)			
MW ± SD	24,9 ± 4,8	24,8 ± 4,8	24,9 ± 4,8
Median	24,1	24,2	24,1
Spanne	15,1 – 52,2	15,5 – 53,3	15,1 – 53,3
Rasse, n (%)			
Asiatisch	30 (2,7)	18 (3,3)	48 (2,9)
Schwarz oder Afroamerikanisch	7 (0,6)	3 (0,5)	10 (0,6)
Sonstige	11 (1,0)	7 (1,3)	18 (1,1)
Unbekannt	5 (0,5)	5 (0,9)	10 (0,6)
Weiß	1.046 (95,2)	513 (94,0)	1.559 (94,8)
Ethnie, n (%)			
Hispanisch oder Latino	74 (6,7)	32 (5,9)	106 (6,4)
Nicht-Hispanisch oder Latino	823 (74,9)	410 (75,1)	1.233 (75,0)
Nicht berichtet	95 (8,6)	58 (10,6)	153 (9,3)
Unbekannt	107 (9,7)	46 (8,4)	153 (9,3)

BMI: Body-Mass-Index; FAS: Full Analysis Set; MW: Mittelwert; N: Patienten gesamt; n: Patienten in der Kategorie; SD: Standard Deviation.
Quelle: Anhang 1 zu Modul 4, Tabelle 1.3.

Tabelle 4-24: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Erkrankungshistorie, Gesamtpopulation)

EXPAND (FAS)	Siponimod (N = 1.099)	BSC (N = 546)	Gesamt (N = 1.645)
Dauer der MS-Erkrankung seit Erstdiagnose (Jahre)			
MW ± SD	12,9 ± 7,9	12,1 ± 7,5	12,6 ± 7,8
Median	11,9	11,2	11,7
Spanne	0,1 – 44,4	0,4 – 39,4	0,1 – 44,4
Dauer der MS-Erkrankung seit den ersten Symptomen (Jahre)			
MW ± SD	17,1 ± 8,4	16,2 ± 8,2	16,8 ± 8,3
Median	16,3	15,4	16,0
Spanne	1,4 – 45,0	1,3 – 43,0	1,3 – 45,0

EXPAND (FAS)	Siponimod (N = 1.099)	BSC (N = 546)	Gesamt (N = 1.645)
Zeit seit Konversion zur SPMS (Jahre)			
MW ± SD	3,9 ± 3,6	3,6 ± 3,3	3,8 ± 3,5
Median	2,6	2,5	2,6
Spanne	0,1 – 24,2	0,1 – 21,7	0,1 – 24,2
Anzahl der Schübe in den letzten 24 Monaten			
MW ± SD	0,7 ± 1,2	0,7 ± 1,2	0,7 ± 1,2
Median	0,0	0,0	0,0
Spanne	0 – 12	0 – 8	0 – 12
Anzahl der Schübe in den letzten 24 Monaten, n (%)			
0	708 (64,4)	343 (62,8)	1.051 (63,9)
1	198 (18,0)	104 (19,0)	302 (18,4)
2 – 3	157 (14,3)	81 (14,8)	238 (14,5)
4 – 5	26 (2,4)	13 (2,4)	39 (2,4)
> 5	7 (0,6)	4 (0,7)	11 (0,7)
Keine Angabe	3 (0,3)	1 (0,2)	4 (0,2)
Anzahl der Schübe in den letzten 12 Monaten			
MW ± SD	0,3 ± 0,5	0,3 ± 0,6	0,3 ± 0,6
Median	0,0	0,0	0,0
Spanne	0 – 4	0 – 4	0 – 4
Anzahl der Schübe in den letzten 12 Monaten, n (%)			
0	874 (79,5)	416 (76,2)	1.290 (78,4)
1	187 (17,0)	111 (20,3)	298 (18,1)
2 – 3	35 (3,2)	16 (2,9)	51 (3,1)
4 – 5	2 (0,2)	2 (0,4)	4 (0,2)
Keine Angabe	1 (0,1)	1 (0,2)	2 (0,1)

FAS: Full Analysis Set; MW: Mittelwert; N: Patienten gesamt; n: Patienten in der Kategorie; SD: Standard Deviation.

Quelle: Anhang 1 zu Modul 4, Tabelle 1.4.

Tabelle 4-25: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (MS-Charakteristika, Gesamtpopulation)

EXPAND (FAS)	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
EDSS			
MW ± SD	5,4 ± 1,1	5,4 ± 1,0	5,4 ± 1,1
Median	6,0	6,0	6,0
Spanne	2,0 – 7,0	2,5 – 7,0	2,0 – 7,0
EDSS, n (%)			
< 3,0	6 (0,5)	2 (0,4)	8 (0,5)
3,0 – 4,5	310 (28,2)	148 (27,1)	458 (27,8)
5,0 – 5,5	165 (15,0)	100 (18,3)	265 (16,1)
6,0 – 6,5	616 (56,1)	295 (54,0)	911 (55,4)
> 6,5	2 (0,2)	1 (0,2)	3 (0,2)
Anzahl Gd-anreichernder T1-Läsionen			
MW ± SD	0,9 ± 3,6	0,7 ± 3,6	0,9 ± 3, 6
Median	0,0	0,0	0,0
Spanne	0 – 55	0 – 65	0 – 65
Anzahl Gd-anreichernder T1-Läsionen, n (%)			
0	828 (75,3)	415 (76,0)	1243 (75,6)
≥ 1	236 (21,5)	114 (20,9)	350 (21,3)
Keine Angabe	35 (3,2)	17 (3,1)	52 (3,2)
T2-Läsionsvolumen (mm^3)			
MW ± SD	15.598,8 ± 16.199,7	14.694,0 ± 15.619,8	15.298,3 ± 16.010,2
Median	10.286,0	9.994,0	10.083,0
Spanne	23,0 – 116.664,0	0,0 – 103.560,0	0,0 – 116.664,0
Normalisiertes Hirnvolumen (cm^3)			
MW ± SD	1.422,0 ± 86,4	1.424,5 ± 87,6	1.422,8 ± 86,8
Median	1.420,5	1.425,2	1.422,3
Spanne	1.136,1 – 1.723,2	1.198,8 – 1.691,2	1.136,1 – 1.723,2

EDSS: Expanded Disability Status Scale; FAS: Full Analysis Set; Gd: Gadolinium; MW: Mittelwert; N: Patienten gesamt; n: Patienten in der Kategorie; SD: Standard Deviation.

Quelle: Anhang 1 zu Modul 4, Tabelle 1.5.

Tabelle 4-26: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Kriterien der SPMS mit aktiver Erkrankung, Gesamtpopulation)

EXPAND (FAS)		Siponimod (N = 1.099)	BSC (N = 546)	Gesamt (N = 1.645)
Schübe in den letzten 24 Monaten	**Gd-anreichernde Läsionen**			
≥ 1	0	269 (24,5)	144 (26,4)	413 (25,1)
≥ 1	≥ 1	108 (9,8)	53 (9,7)	161 (9,8)
0	0	557 (50,7)	270 (49,5)	827 (50,3)
0	≥ 1	127 (11,6)	61 (11,2)	188 (11,4)

FAS: Full Analysis Set; N: Patienten gesamt.
Quelle: Anhang 1 zu Modul 4, Tabelle 1.6.

Subpopulation B

Tabelle 4-27: Patientendisposition – RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND (RAN)	Siponimod (N = 128)	BSC (N = 61)	Gesamt (N = 189)
Full Analysis Set, n (%)			
In FAS eingeschlossen	127 (99,2)	61 (100,0)	188 (99,5)
Grund für Ausschluss aus dem FAS			
Keine Medikation erhalten	1 (0,8)	0 (0,0)	1 (0,5)
Studienabbrüche, n (%)			
Studie abgebrochen	24 (18,8)	16 (26,2)	40 (21,2)
Abbruchgründe			
Entscheidung des Patienten/Vormunds	13 (10,2)	10 (16,4)	23 (12,2)
Unerwünschtes Ereignis	5 (3,9)	3 (4,9)	8 (4,2)
Entscheidung des Arztes	3 (2,3)	0 (0,0)	3 (1,6)
Keine Nachbeobachtung möglich	1 (0,8)	1 (1,6)	2 (1,1)
Tod	1 (0,8)	0 (0,0)	1 (0,5)
Fortschreiten der Erkrankung	1 (0,8)	0 (0,0)	1 (0,5)
Fehlende Wirksamkeit	0 (0,0)	1 (1,6)	1 (0,5)
Protokollabweichung	0 (0,0)	1 (1,6)	1 (0,5)

EXPAND (RAN)	Siponimod (N = 128)	BSC (N = 61)	Gesamt (N = 189)
Medikation, n (%)			
Medikation erhalten	127 (99,2)	61 (100,0)	188 (99,5)
Keine Medikation erhalten	1 (0,8)	0 (0,0)	1 (0,5)
Doppelblinde Behandlung abgeschlossen	83 (64,8)	36 (59,0)	119 (63,0)
Doppelblinde Behandlung vorzeitig abgebrochen	45 (35,2)	25 (41,0)	70 (37,0)
Mit unverblindetem Siponimod weiterbehandelt	14 (10,9)	9 (14,8)	23 (12,2)
Mit verkürztem Visitenschema fortgefahren	7 (5,5)	2 (3,3)	9 (4,8)
Vollständiger Abbruch der Behandlungsphase	24 (18,8)	14 (23,0)	38 (20,1)
Abbruchgründe			
Fortschreiten der Erkrankung	16 (12,5)	8 (13,1)	24 (12,7)
Entscheidung des Patienten/Vormunds	16 (12,5)	8 (13,1)	24 (12,7)
Unerwünschtes Ereignis	9 (7,0)	4 (6,6)	13 (6,9)
Entscheidung des Arztes	2 (1,6)	0 (0,0)	2 (1,1)
Fehlende Wirksamkeit	1 (0,8)	4 (6,6)	5 (2,6)
Abbruch entspricht Protokollvorgaben	1 (0,8)	0 (0,0)	1 (0,5)
Dosierungsfehler	0 (0,0)	1 (1,6)	1 (0,5)
Protokollabweichungen, n (%)			
Mindestens eine Protokollabweichung	74 (57,8)	38 (62,3)	112 (59,3)
Wichtige Prozeduren nicht protokollgemäß	45 (35,2)	23 (37,7)	68 (36,0)
Andere GCP-Abweichung	35 (27,3)	17 (27,9)	52 (27,5)
Unerlaubte Begleitmedikation	10 (7,8)	3 (4,9)	13 (6,9)
Einschlusskriterien verletzt	11 (8,6)	3 (4,9)	14 (7,4)
Behandlungsabweichung	4 (3,1)	5 (8,2)	9 (4,8)

FAS: Full Analysis Set; GCP: Good Clinical Practice; N: Patienten gesamt; n: Anzahl Patienten in der Kategorie; RAN: Randomisierte Patienten.

Quelle: Anhang 2 zu Modul 4, Tabelle 1.1.

Tabelle 4-28: Beobachtungsdauer – RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND (FAS)	Siponimod (N = 127)	BSC (N = 61)	Gesamt (N = 188)
Beobachtungsdauer			
MW ± SD (Tage)	650,0 ± 249,6	612,4 ± 245,9	637,8 ± 248,4
Median (Tage) [a]	654,0	605,0	642,0
Minimum; Maximum (Tage)	7; 1.101	48; 1.078	7; 1.101
Patienten ≥ 365 Tage in der Studie, n (%)	111 (87,4)	53 (86,9)	164 (87,2)
Patienten < 365 Tage in der Studie, n (%)	16 (12,6)	8 (13,1)	24 (12,8)
Patienten ≥ 730 Tage in der Studie, n (%)	54 (42,5)	22 (36,1)	76 (40,4)
Patienten < 730 Tage in der Studie, n (%)	73 (57,5)	39 (63,9)	112 (59,6)

FAS: Full Analysis Set; MW: Mittelwert; N: Patienten gesamt; n: Anzahl Patienten in der Kategorie; SD: Standard Deviation.

a: Der Unterschied von etwa 50 Tagen zwischen den Behandlungsgruppen beträgt somit weniger als 10 % der gesamten Beobachtungsdauer und wird insgesamt als nicht relevant eingestuft.

Quelle: Anhang 2 zu Modul 4, Tabelle 1.2.

Tabelle 4-29: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Demographie, Subpopulation B)

EXPAND (FAS)	Siponimod (N = 127)	BSC (N = 61)	Gesamt (N = 188)
Alter (Jahre)			
MW ± SD	47,1 ± 7,0	48,4 ± 8,7	47,6 ± 7,6
Median	47,0	50,0	48,0
Spanne	25 – 61	24 – 60	24 – 61
Altersgruppen (Jahre), n (%)			
18 – 40	18 (14,2)	13 (21,3)	31 (16,5)
41 – 55	95 (74,8)	33 (54,1)	128 (68,1)
> 55	14 (11,0)	15 (24,6)	29 (15,4)
Geschlecht, n (%)			
Weiblich	76 (59,8)	37 (60,7)	113 (60,1)
Männlich	51 (40,2)	24 (39,3)	75 (39,9)
Gewicht (kg)			
MW ± SD	70,5 ± 15,8	68,0 ± 12,4	69,7 ± 14,8
Median	69,0	67,4	68,0
Spanne	42,6 – 142,1	47,2 – 97,2	42,6 – 142,1

EXPAND (FAS)	Siponimod (N = 127)	BSC (N = 61)	Gesamt (N = 188)
Größe (cm)			
MW ± SD	169,0 ± 9,7	168,7 ± 9,6	168,9 ± 9,6
Median	168,0	168,0	168,0
Spanne	142,2 – 193,0	152,0 – 195,0	142,2 – 195,0
BMI (kg/m²)			
MW ± SD	24,6 ± 4,8	23,9 ± 3,3	24,4 ± 4,4
Median	24,2	23,9	24,0
Spanne	16,4 – 52,2	17,3 – 31,6	16,4 – 52,2
Rasse, n (%)			
Asiatisch	1 (0,8)	0 (0,0)	1 (0,5)
Schwarz oder Afroamerikanisch	0 (0,0)	1 (1,6)	1 (0,5)
Sonstige	0 (0,0)	2 (3,3)	2 (1,1)
Unbekannt	0 (0,0)	1 (1,6)	1 (0,5)
Weiß	126 (99,2)	57 (93,4)	183 (97,3)
Ethnie, n (%)			
Hispanisch oder Latino	8 (6,3)	1 (1,6)	9 (4,8)
Nicht-Hispanisch oder Latino	90 (70,9)	47 (77,0)	137 (72,9)
Nicht berichtet	18 (14,2)	7 (11,5)	25 (13,3)
Unbekannt	11 (8,7)	6 (9,8)	17 (9,0)

BMI: Body-Mass-Index; FAS: Full Analysis Set; MW: Mittelwert; N: Patienten gesamt; n: Patienten in der Kategorie; SD: Standard Deviation.
Quelle: Anhang 2 zu Modul 4, Tabelle 1.3.

Tabelle 4-30: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (Erkrankungshistorie, Subpopulation B)

EXPAND (FAS)	Siponimod (N = 127)	BSC (N = 61)	Gesamt (N = 188)
Dauer der MS-Erkrankung seit Erstdiagnose (Jahre)			
MW ± SD	13,2 ± 6,8	12,0 ± 6,7	12,8 ± 6,8
Median	12,4	10,5	11,9
Spanne	0,5 – 32,9	0,4 – 33,2	0,4 – 33,2
Dauer der MS-Erkrankung seit den ersten Symptomen (Jahre)			
MW ± SD	17,1 ± 7,5	15,8 ± 7,9	16,7 ± 7,6
Median	16,2	15,3	16,0
Spanne	3,2 – 41,3	2,3 – 33,5	2,3 – 41,3

EXPAND (FAS)	Siponimod (N = 127)	BSC (N = 61)	Gesamt (N = 188)
Zeit seit Konversion zur SPMS (Jahre)			
MW ± SD	4,4 ± 3,8	4,0 ± 3,0	4,3 ± 3,5
Median	3,3	3,2	3,2
Spanne	0,6 – 20,0	0,5 – 12,8	0,5 – 20,0
Anzahl der Schübe in den letzten 24 Monaten, n (%)			
0	127 (100,0)	61 (100,0)	188 (100,0)
Anzahl der Schübe in den letzten 12 Monaten, n (%)			
0	127 (100,0)	61 (100,0)	188 (100,0)

FAS: Full Analysis Set; MW: Mittelwert; N: Patienten gesamt; n: Patienten in der Kategorie; SD: Standard Deviation.
Quelle: Anhang 2 zu Modul 4, Tabelle 1.4.

Tabelle 4-31: Charakterisierung der Studienpopulationen – RCT mit dem zu bewertenden Arzneimittel (MS-Charakteristika, Subpopulation B)

EXPAND (FAS)	Siponimod (N = 127)	BSC (N = 61)	Gesamt (N = 188)
EDSS			
MW ± SD	5,6 ± 1,1	5,5 ± 1,1	5,6 ± 1,1
Median	6,0	6,0	6,0
Spanne	2,5 – 6,5	3,0 – 6,5	2,5 – 6,5
EDSS, n (%)			
< 3,0	1 (0,8)	0 (0,0)	1 (0,5)
3,0 – 4,5	29 (22,8)	16 (26,2)	45 (23,9)
5,0 – 5,5	15 (11,8)	9 (14,8)	24 (12,8)
6,0 – 6,5	82 (64,6)	36 (59,0)	118 (62,8)
> 6,5	0 (0,0)	0 (0,0)	0 (0,0)
Anzahl Gd-anreichernder T1-Läsionen			
MW ± SD	3,0 ± 3,2	3,1 ± 8,4	3,0 ± 5,4
Median	2,0	1,0	2,0
Spanne	1 – 18	1 – 65	1 – 65
Anzahl Gd-anreichernder T1-Läsionen, n (%)			
0	0 (0,0)	0 (0,0)	0 (0,0)
≥ 1	127 (100,0)	61 (100,0)	188 (100,0)

EXPAND (FAS)	Siponimod (N = 127)	BSC (N = 61)	Gesamt (N = 188)
T2-Läsionsvolumen (mm^3)			
MW ± SD	21.552,1 ± 16.818,9	22.260,6 ± 19.500,5	21.782,0 ± 17.684,0
Median	17.826,0	15.771,0	17.319,0
Spanne	918,0 – 81.591,0	918,0 – 103.560,0	918,0 – 103.560,0
Normalisiertes Hirnvolumen (cm^3)			
MW ± SD	1.414,8 ± 91,9	1.423,2 ± 92,9	1.417,5 ± 92,1
Median	1.411,5	1.417,5	1.411,6
Spanne	1.171,4 – 1644,1	1.230,2 – 1678,7	1.171,4 – 1.678,7

EDSS: Expanded Disability Status Scale; FAS: Full Analysis Set; MW: Mittelwert; N: Patienten gesamt; n: Patienten in der Kategorie; SD: Standard Deviation.
Quelle: Anhang 2 zu Modul 4, Tabelle 1.5.

Beschreiben Sie die Studien zusammenfassend. In der Beschreibung der Studien sollten Informationen zur Behandlungsdauer sowie zu geplanter und tatsächlicher Beobachtungsdauer enthalten sein. Sofern sich die Beobachtungsdauer zwischen den relevanten Endpunkten unterscheidet, sind diese unterschiedlichen Beobachtungsdauern endpunktbezogen anzugeben. Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

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

Die Studie EXPAND umfasst eine randomisierte, placebokontrollierte Kernphase und eine unverblindete, unkontrollierte Extensionsphase. Die Extensionsphase ist nicht Gegenstand der Nutzenbewertung. Die Kernphase wurde etwa drei Jahren nach Beginn der Randomisierung beendet, da angenommen wurde, dass zu diesem Zeitpunkt die für die Auswertung der Behinderungsprogression notwendigen 374 Ereignisse eingetreten waren und die Randomisierung für mehr als 95 % der Patienten mindestens ein Jahr zurücklag. Dieses ereignisgetriebene Design hat den Vorteil, dass Patienten schneller in die unverblindete Extensionsphase aufgenommen werden können und damit weniger Risiko ausgesetzt sind.

Eingeschlossen wurden erwachsene Patienten im Alter von 18 bis 60 Jahren mit sekundär progredient verlaufender MS, gekennzeichnet durch ein schubunabhängiges Fortschreiten der Behinderung über mindestens sechs Monate vor Studieneinschluss. Ein schubförmig remittierender Verlauf gemäß der 2010 überarbeiteten McDonald-Kriterien musste vorausgegangen sein (49). Der EDSS-Wert zu Studienbeginn musste bei 3,0 bis einschließlich

6,5 liegen und die Patienten sollten eines der folgende Kriterien der Krankheitsaktivität erfüllen:

- dokumentierte EDSS-Progression ≥ 1 innerhalb von zwei Jahren vor Einschluss bei Patienten mit einem EDSS < 6,0 beim Screening;

- dokumentierte EDSS-Progression ≥ 0,5 innerhalb von zwei Jahren vor Einschluss bei Patienten mit einem EDSS ≥ 6,0 beim Screening.

Innerhalb von drei Monaten vor Randomisierung durften keine Schübe aufgetreten sein und keine Kortikosteroid-Behandlung erfolgt sein. Eine weitere Einschränkung hinsichtlich der Vorbehandlung wurde nicht vorgenommen. Patienten, die während der Studie einen Schub hatten, mussten erneut zur Teilnahme einwilligen.

Die Studie bestand aus zwei Behandlungsarmen. In einem erhielten die Patienten 1 × täglich Siponimod 2 mg zur oralen Einnahme, im anderen Arm das entsprechende Placebo. Die Behandlung mit Siponimod bzw. Placebo begann mit einer Auftitrationsphase von fünf Tagen (Tag 1 und 2: je 0,25 mg; Tag 3: 0,50 mg; Tag 4: 0,75 mg; Tag 5: 1,25 mg; ab Tag 6: volle Dosis von 2 mg). Patienten mit einer Lymphozytenzahl $< 0{,}2 \times 10^9$/l bei zwei aufeinanderfolgenden Visiten mussten die Dosis auf 1 mg reduzieren.

Laut Fachinformation beträgt die Erhaltungsdosis 1 mg für Patienten mit einm Polymorphismus des CYP2C9-Gens (Genotyp CYP2C9*2*3 oder1*3). In der EXPAND-Studie erhielten diese Patienten Siponimod auch in einer Dosierung von 2 mg. In der Gesamtpopulation trifft das auf 237 von 1.645 Patienten zu (*2/*3 (= CYP2C9*2/*3): n = 43; 3/WT (= CYP2C9*1/*3): n = 194). Für vier Patienten fehlt die Angabe zum Genotyp. Unter der konservativen Annahme, dass diese vier Patienten auch einen relevanten Polymorphismus gehabt hätten, liegt der Anteil der Patienten, die nicht mit der Dosierung gemäß Fachinformation behandelt wurden, in der Gesamtpopulation bei 14,7 %. In der Subpopulation B betrifft dies 25 von 188 Patienten (*2/*3 (= CYP2C9*2/*3): n = 1; 3WT (= CYP2C9*1/*3): n = 24), für einen Patienten fehlt die Angabe zum Genotyp. Unter der konservativen Annahme, dass der genannte Patient auch einen relevanten Polymorphismus gehabt hätte, liegt der Anteil der Patienten, die nicht mit der Dosierung gemäß Fachinformation behandelt wurden in der Subpopulation B bei 13,8 %. Der Anteil der nicht gemäß Fachinformation behandelten Patienten liegt also deutlich unter 20 %. Die Subpopulation B ist daher für die Nutzenbewertung im relevanten Anwendungsgebiet geeignet.

In den Behandlungsarmen bestand die zusätzliche Möglichkeit der Anwendung pharmakologischer sowie nicht-pharmakologischer Interventionen zur Behandlung der Symptome. Ausgeschlossen waren Behandlungen, die immunsuppressiv oder immunmodulierend wirken. Die Operationalisierung von BSC war in der EXPAND-Studie daher korrekt umgesetzt.

Die individuelle Zeit in der Studie war aufgrund des ereignisgetriebenen Designs von Patient zu Patient sehr unterschiedlich. Die mediane Beobachtungsdauer lag bei 639 Tagen in der Gesamtpopulation ohne relevanten Unterschied zwischen den Gruppen (638,0 Tage unter

Siponimod; 640,5 Tage unter BSC). 88,1 % der Patienten wiesen eine Beobachtungsdauer von mindestens 365 Tagen auf. Der Anteil war in den Behandlungsgruppen mit 88,2 % (Siponimod) und 87,9 % (BSC) ähnlich. Eine Beobachtungsdauer von 730 Tagen erreichten insgesamt 36,1 % der Patienten (Tabelle 4-22). In der Subpopulation B lag die mediane Beobachtungsdauer bei 654,0 Tagen in der Siponimod-Gruppe und bei 605,0 Tagen in der BSC-Gruppe. Der Unterschied von etwa 50 Tagen zwischen den Behandlungsgruppen beträgt somit weniger als 10 % der gesamten Beobachtungsdauer und wird insgesamt als nicht relevant eingestuft. Außerdem war der Anteil der Patienten mit einer Beobachtungsdauer von mindestens 365 Tagen in den Behandlungsgruppen mit 87,4 % (Siponimod) und 86,9 % (BSC) ähnlich. Eine Beobachtungsdauer von 730 Tagen erreichten insgesamt 40,4 % der Patienten der Subpopulation B (Tabelle 4-28). Die Auswertung der Endpunkte zu Monat 24 ist daher durch hohe Nichtberücksichtigungsanteile gekennzeichnet.

Der primäre Endpunkt der Studie war die Zeit bis zur Behinderungsprogression, gemessen anhand des EDSS mit Bestätigung nach drei Monaten. Patientenrelevante sekundäre Endpunkte waren bestätigte Krankheitsschübe, die Behinderungsprogression (gemessen anhand von EDSS und MSFC), die kognitive Beeinträchtigung (gemessen anhand von SDMT und BVMT-R), die visuelle Beeinträchtigung (gemessen anhand von LCVA), die Beeinträchtigung der Gehfähigkeit (gemessen anhand von MSWS-12) sowie T2-Läsionslast, Hirnvolumen, Unerwünschte Ereignisse, Suizidalität und die gesundheitsbezogene Lebensqualität (gemessen anhand von MSIS-29 und EQ-5D).

In der Studienpopulation waren sich die beiden Behandlungsarme hinsichtlich der Patientenmerkmale bei Studienbeginn sehr ähnlich.

Die Studienpopulation sowie die Subpopulation B sind in ihren demographischen und anderen Charakteristika zu Studienbeginn, in der Diagnose und in der Vorbehandlung weitestgehend strukturgleich zur Population mit SPMS bzw. SPMS mit aktiver Erkrankung in Deutschland (50). Die Studienergebnisse sowie insbesondere die Ergebnisse zur Subpopulation B sind somit auf den deutschen Versorgungskontext übertragbar.

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 4-32: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungs-sequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
EXPAND	ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Da die Randomisierungssequenz adäquat erzeugt wurde, die Gruppenzuteilung verdeckt erfolgte, Patienten und Behandler verblindet waren und keine sonstigen das Verzerrungspotenzial beeinflussenden verzerrenden Punkte vorliegen, wird das Verzerrungspotenzial auf Studienebene insgesamt als niedrig bewertet. Die post-hoc durchgeführten Subgruppenanalysen führen hierbei zu keiner relevant erhöhten Verzerrung.

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 4-33: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität	Bestätigte Krankheits-schübe	Behinderungs-progression	Kognition/ Sehvermögen/ Gehfähigkeit	Gesundheits-bezogene Lebens-qualität	Sicherheit/ Verträglich-keit
EXPAND	ja	ja	ja	ja	ja	ja

Für die Analyse der Endpunkte wurden folgende Populationen herangezogen (die jeweils verwendete Analysepopulation ist in den Ergebnistabellen zu jedem Endpunkt spezifiziert):

ITT/FAS Das Intention-to-treat Set (ITT) / Full Analysis Set (FAS) beinhaltet alle randomisierten Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben. Gemäß dem Intention-to-treat-Prinzip wurden die Patienten

in der Behandlungsgruppe analysiert, in die sie randomisiert worden waren. Alle Analysen zu Endpunkten der Wirksamkeit basieren auf dem ITT/FAS.

SAF — Das Safety Set (SAF) beinhaltet alle Patienten, die mindestens eine Dosis der Studienmedikation erhalten haben. Die Patienten wurden in der Behandlungsgruppe analysiert, deren Behandlung sie erhalten haben. Alle Analysen zu Endpunkten der Sicherheit basieren auf dem SAF.

4.3.1.3.1 Mortalität – 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 inkl. Angaben zur Häufigkeit von und zum Umgang mit nicht oder nicht vollständig beobachteten Patienten (bei Verlaufsbeobachtungen pro Messzeitpunkt)
– 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.

Unterschiedliche Beobachtungszeiten zwischen den Behandlungsgruppen sollen durch adäquate Analysen (z.B. Überlebenszeitanalysen) adressiert werden, und zwar für alle Endpunkte (einschließlich UE nach den nachfolgend genannten Kriterien), für die eine solche Analyse aufgrund deutlich unterschiedlicher Beobachtungszeiten erforderlich ist.

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden. Dabei ist für jeden Endpunkt, für den eine solche Analyse durchgeführt wird, eine separate Kaplan-Meier-Kurve darzustellen.

Zu mit Skalen erhobenen patientenberichteten Endpunkten (z.B. zur gesundheitsbezogenen Lebensqualität oder zu Symptomen) sind immer auch die Werte im Studienverlauf anzugeben, auch als grafische Darstellung, sowie eine Auswertung, die die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt (z.B. als Symptomlast über die Zeit, geschätzt mittels MMRM-Analyse [falls aufgrund der Datenlage geeignet]).

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen:

1. Gesamtrate UE,

2. Gesamtrate schwerwiegender UE (SUE),

3. Gesamtrate der Abbrüche wegen UE,

4. Gesamtraten von UE differenziert nach Schweregrad, sofern dies in der/den relevante/n Studie/n erhoben wurde (z.B. gemäß CTCAE und/oder einer anderen etablierten bzw. validierten indikationsspezifischen Klassifikation) einschließlich einer Abgrenzung schwerer und nicht schwerer UE,

5. zu den unter 1, 2 und 4 genannten Kategorien (UE ohne weitere Differenzierung, SUE, UE differenziert nach Schweregrad) soll zusätzlich zu den Gesamtraten die Darstellung nach Organsystemen und Einzelereignissen (als System Organ Class [SOCs] und Preferred Terms [PT] nach MedDRA) jeweils nach folgenden Kriterien erfolgen:

- UE (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10% der Patienten in einem Studienarm aufgetreten sind

- Schwere UE (z.B. CTCAE-Grad ≥ 3) und SUE: Ereignisse, die bei mindestens 5% der Patienten in einem Studienarm aufgetreten sind

- zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patienten UND bei mindestens 1 % der Patienten in einem Studienarm aufgetreten sind.

6. A priori definierte UE von besonderem Interesse [AESI]) sowie prädefinierte SOC-übergreifende UE-Auswertungen (z.B. als Standardised MedDRA Queries, SMQs) sollen unabhängig von der Ereignisrate dargestellt werden und zwar differenziert nach Schweregrad (dargestellt als Gesamtrate und differenziert nach Schweregrad, nicht schwer, schwer, schwerwiegend).

7. zu Kategorie 3: Die Abbruchgründe auf SOC/PT-Ebene müssen vollständig, jedoch nur deskriptiv dargestellt werden.

Sofern bei der Erhebung unerwünschter Ereignisse erkrankungsbezogenen Ereignisse (z. B. Progression, Exazerbation) berücksichtigt werden (diese Ereignisse also in die UE-Erhebung eingehen), sollen für die Gesamtraten (UE, schwere UE und SUE) zusätzliche UE-Analysen durchgeführt werden, bei denen diese Ereignisse unberücksichtigt bleiben. Alle Auswertungen

zu UE können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine Darstellung ausschließlich in Modul 5 ist nicht ausreichend. Davon unbenommen sind die Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE), sowie die für die Gesamtaussage zum Zusatznutzen herangezogenen Ergebnisse im vorliegenden Abschnitt darzustellen.

Auswertungen zu den im Abschnitt 4.3.1.2.1 aufgeführten Datenschnitten sollen vollständig, d.h. für alle erhobenen relevanten Endpunkte, durchgeführt und vorgelegt werden. Das gilt auch dann wenn ein Datenschnitt ursprünglich nur zur Auswertung einzelner Endpunkte geplant war. Auf die Darstellung der Ergebnisse einzelner Endpunkte eines Datenschnitts bzw. eines gesamten Datenschnitts kann verzichtet werden, wenn hierdurch kein wesentlicher Informationsgewinn gegenüber einem anderen Datenschnitt zu erwarten ist (z. B. wenn die Nachbeobachtung zu einem Endpunkt bereits zum vorhergehenden Datenschnitt nahezu vollständig war oder ein Datenschnitt in unmittelbarer zeitlicher Nähe zu einem anderen Datenschnitt liegt).

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

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

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

Tabelle 4-34: Operationalisierung des Endpunkts *Mortalität*

Studie	Operationalisierung
EXPAND	Informationen zu Todesfällen wurden im Rahmen der (S)UE-Dokumentation erhoben. Berichtet wird die: • Inzidenz von Todesfällen jeglicher Ursache (binäre Analyse) mit den Risikoschätzern OR, RR und RD, jeweils mit den dazugehörigen Konfidenzintervallen und p-Werten.

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 4-35: Bewertung des Verzerrungspotenzials für den Endpunkt *Mortalität (Todesfälle)* 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**
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Mortalität (Todesfälle)* als niedrig bewertet, da das Verzerrungspotenzial auf Studienebene niedrig war, der Endpunkterheber (Prüfarzt) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und sonstige das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden. Daher kann hier von einem objektivierbaren Endpunkt ausgegangen werden und es ist nur mit einem niedrigen daraus resultierenden Verzerrungspotenzial zu rechnen.

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

Tabelle 4-36: Ergebnisse für den Endpunkt *Mortalität* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 1.099)	**BSC (N = 546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Mortalität					
N'	1.099	546			
Todesfälle – n (%)	4 (0,4)	4 (0,7)	0,49 [0,12; 1,99] 0,321	0,50 [0,12; 1,98] 0,321	0,00 [-0,01; 0,00] 0,366

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 1 zu Modul 4, Tabelle 2.1.

Tabelle 4-37: Ergebnisse für den Endpunkt *Mortalität* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N = 61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Mortalität					
N'	127	61			
Todesfälle – n (%)	1 (0,8)	1 (1,6)	0,48 [0,03; 7,74] 0,602	0,48 [0,03; 7,55] 0,602	-0,01 [-0,04; 0,03] 0,637

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 2 zu Modul 4, Tabelle 2.1.

In der EXPAND-Studie ereigneten sich insgesamt acht Todesfälle (jeweils vier pro Behandlungsarm), zwei davon in der Subpopulation B (jeweils einer pro Behandlungsarm). Es ergab sich kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen.

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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 Bestätigte Krankheitsschübe – RCT

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

Tabelle 4-38: Operationalisierung des Endpunkts *Bestätigte Krankheitsschübe*

Studie	Operationalisierung
EXPAND	Ein Krankheitsschub war im Prüfplan definiert als das Auftreten neuer neurologische Symptome oder die Verschlechterung von zuvor stabilen oder sich verbessernden vorbestehenden Symptomen, die mindestens 30 Tage nach einem vorhergehenden demyelinisierenden Ereignis auftraten. Die Symptome mussten für mindestens 24 Stunden andauern und ohne Fieber (< 37,5° C) oder eine Infektion aufgetreten sein. Ein bestätigter Krankheitsschub lag vor, wenn ein wie oben definierter Krankheitsschub mit einem klinisch relevanten Anstieg im EDSS-Wert einherging (durch einen unabhängigen EDSS-Bewerter festgestellt). Die EDSS-Bewertung sollte binnen sieben Tagen nach Einsetzen der Symptome vorgenommen werden. Ein klinisch relevanter Anstieg war definiert als ein Anstieg um 0,5 Punkte im EDSS-Wert oder als ein Anstieg um einen Punkt auf zwei verschiedenen „Funktionellen Systemen“ (FS) der EDSS oder als ein Anstieg um zwei Punkte auf einem FS (ausschließlich Darm/Blase oder zerebrales FS). Das Auftreten von Symptomen eines Krankheitsschubs wurde bei jeder Visite durch den Arzt erfragt. Zudem konnten Symptome auch jederzeit durch den Patienten selbst spontan gemeldet werden. Berichtet werden folgende Endpunkte: • die jährliche Rate bestätigter Krankheitsschübe (Annualized Relapse Rate; ARR) als Rate Ratio, basierend auf einem negativ-binomialen Regressionsmodell. Die Zeit, die ein Patient in der Studie war (logarithmierte Zeit in Jahren) wird als Offset-Variable modelliert, um die patientenindividuelle Studiendauer zu berücksichtigen und eine Schätzung der ARR zu erhalten; • die Zeit bis zum ersten bestätigten Schub bzw. das Risiko für einen bestätigten Schub aus einer Kaplan-Meier-Schätzung (Hazard Ratio basierend auf Cox-Regression), inklusive Kaplan-Meier-Schätzer für den Anteil der Patienten ohne bestätigten Krankheitsschub.

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 4-39: Bewertung des Verzerrungspotenzials für den Endpunkt *Bestätigte Krankheitsschübe* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Bestätigte Krankheitsschübe* als niedrig bewertet, da das Verzerrungspotenzial auf Studienebene niedrig war, die Endpunkterheber (Prüfarzt und unabhängiger EDSS-Bewerter) verblindet waren, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und sonstige, das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden. Die Kriterien für einen bestätigten Krankheitsschub waren a priori im Prüfplan definiert. Die Herleitung der jährlichen Schubrate (ARR) aus den dokumentierten (bestätigten) Krankheitsschüben war ebenfalls a priori im Prüfplan definiert. Daher kann hier von einem objektivierbaren Endpunkt ausgegangen werden und es ist nur mit einem niedrigen daraus resultierenden Verzerrungspotenzial zu rechnen.

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

Jährliche Rate bestätigter Krankheitsschübe

Tabelle 4-40: Ergebnisse für den Endpunkt *Jährliche Rate bestätigter Krankheitsschübe* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Rate Ratio [95%-KI]**	**p-Wert**
Jährliche Rate bestätigter Krankheitsschübe				
N'	1.061	528		
Anzahl der Patienten mit Schüben	113	100		
Anzahl der Schübe	134	141		
Jährliche Schubrate [95%-KI]	0,07 [0,06; 0,09]	0,16 [0,12; 0,21]	0,45 [0,34; 0,59]	< 0,001

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
Quelle: Anhang 1 zu Modul 4, Tabelle 3.1.1.

Tabelle 4-41: Ergebnisse für den Endpunkt *Jährliche Rate bestätigter Krankheitsschübe* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Rate Ratio [95%-KI]**	**p-Wert**
Jährliche Rate bestätigter Krankheitsschübe				
N'	127	61		
Anzahl der Patienten mit Schüben	13	12		
Anzahl der Schübe	14	15		
Jährliche Schubrate [95%-KI]	0,06 [0,03; 0,10]	0,14 [0,08; 0,26]	0,41 [0,18; 0,92]	0,031

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
Quelle: Anhang 2 zu Modul 4, Tabelle 3.1.1.

In den Auswertungen der Gesamtpopulation sowie der Subpopulation B der EXPAND-Studie war die jährliche Rate bestätigter Krankheitsschübe unter Siponimod statistisch signifikant niedriger als unter der Behandlung mit BSC (Tabelle 4-40 und Tabelle 4-41).

Zeit bis zum ersten bestätigten Krankheitsschub

Tabelle 4-42: Ergebnisse für den Endpunkt *Zeit bis zum ersten bestätigten Krankheitsschub* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**HR [95%-KI][b]**	**p-Wert**
Zeit bis zum ersten bestätigten Krankheitsschub				
N'	1.061	528		
Anzahl der Patienten mit Schub	113	100		
Median in Tagen [95%-KI][a]	n. b.	n. b.		
Anteil schubfreier Patienten nach 24 Monaten in % – KM-Schätzer [95%-KI][a]	89,00 [86,99; 91,01]	79,59 [75,84; 83,35]		
Anteil schubfreier Patienten zum Studienende in % – KM-Schätzer [95%-KI][a]	82,19 [74,71; 89,66]	76,92 [72,20; 81,65]		
			0,54 [0,41; 0,70]	< 0,001

HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar (Überlebenskurve erreicht das 50%-Level nicht).

a: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

b: Cox proportional hazards model.

Quelle: Anhang 1 zu Modul 4, Tabelle 3.2.1.

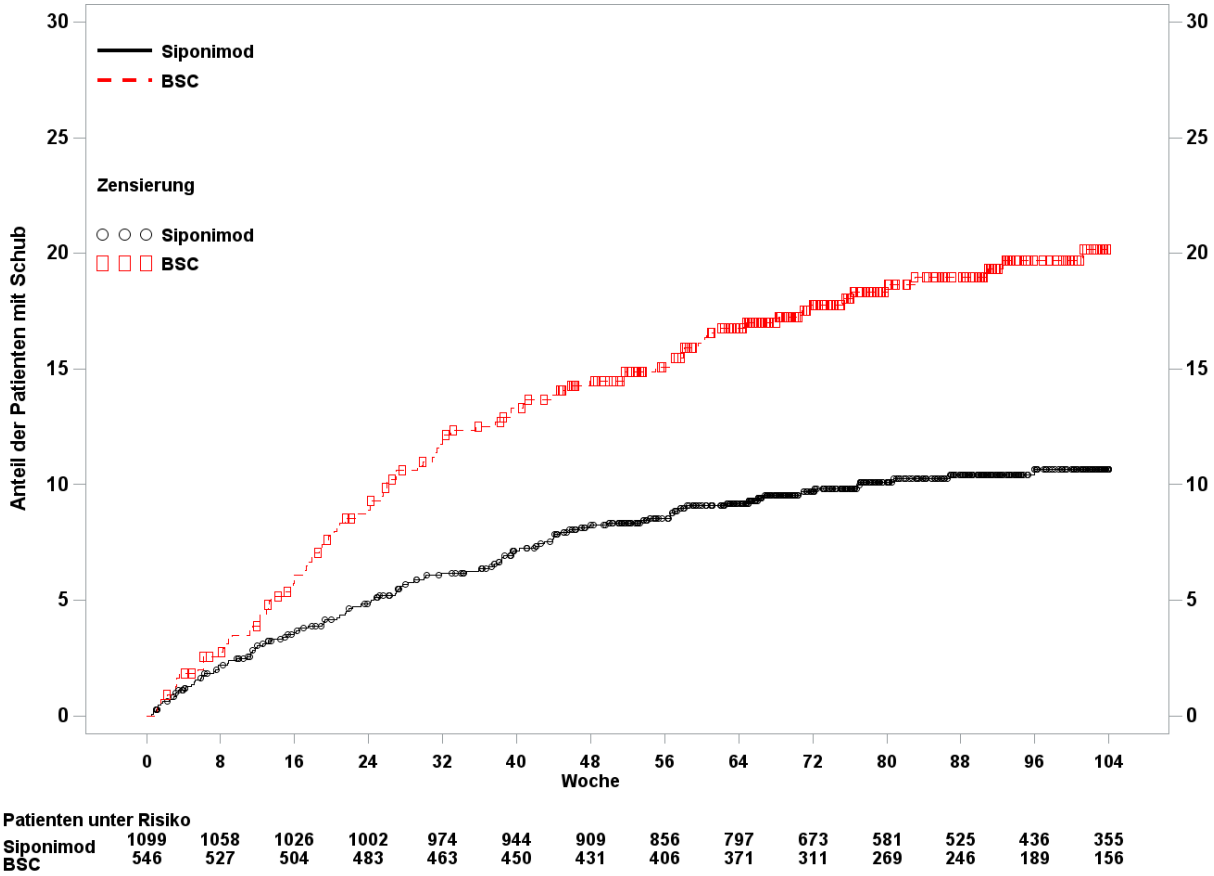


Abbildung 2: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten bestätigten Krankheitsschub* (Studienende) aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Tabelle 4-43: Ergebnisse für den Endpunkt *Zeit bis zum ersten bestätigten Krankheitsschub* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**HR [95%-KI][b]**	**p-Wert**
Zeit bis zum ersten bestätigten Krankheitsschub				
N'	127	61		
Anzahl der Patienten mit Schub	13	12		
Median in Tagen [95%-KI][a]	n. b.	n. b.		
Anteil schubfreier Patienten nach 24 Monaten in % – KM-Schätzer [95%-KI][a]	90,81 [85,63; 95,99]	79,47 [69,08; 89,86]		
Anteil schubfreier Patienten zum Studienende in % – KM-Schätzer [95%-KI][a]	80,32 [64,05; 96,60]	79,47 [69,08; 89,86]		
			0,38 [0,16; 0,89]	0,026

HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b. nicht bestimmbar (Überlebenskurve erreicht das 50%-Level nicht).

a: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

b: Cox proportional hazards model.

Quelle: Anhang 2 zu Modul 4, Tabelle 3.2.1.

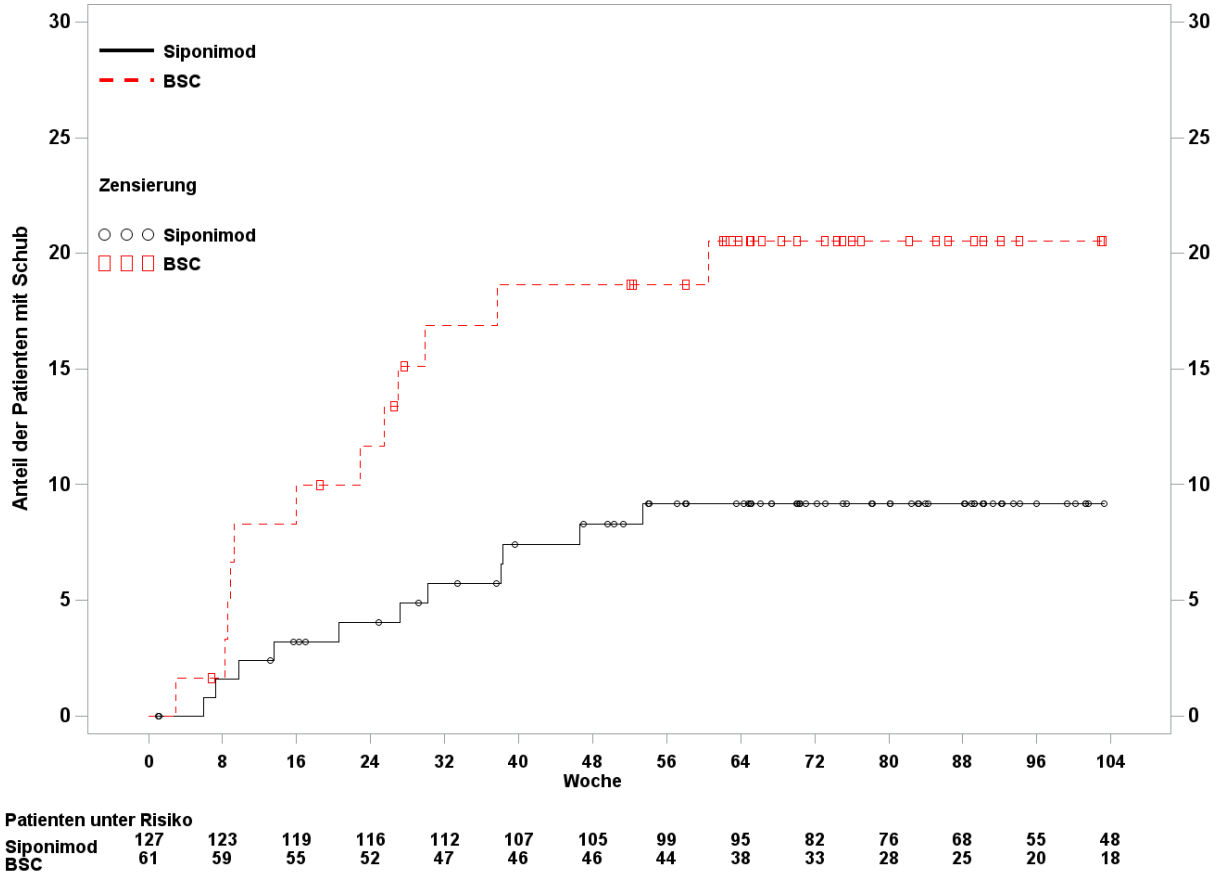


Abbildung 3: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten bestätigten Krankheitsschub* (Studienende) aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

In der Auswertung der Gesamtpopulation war das Schubrisiko unter Siponimod statistisch signifikant um 46 % reduziert im Vergleich zu BSC (HR 0,54; 95%-KI [0,41; 0,70]; Tabelle 4-42). In der Subpopulation B war das Schubrisiko unter Siponimod statistisch signifikant um 62 % reduziert (HR 0,38; 95%-KI [0,16; 0,89]; Tabelle 4-43).

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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 Behinderungsprogression (EDSS) – RCT

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

Tabelle 4-44: Operationalisierung des Endpunkts *Behinderungsprogression*

Studie	Operationalisierung
EXPAND	Die Behinderungsprogression war definiert als Erhöhung des EDSS-Werts um einen Punkt für Patienten mit einem EDSS-Wert von 3,0 bis 5,0 bei Studienbeginn oder um 0,5 Punkte für Patienten mit einem EDSS-Wert von 5,5 bis 6,5 Punkten ebenfalls bei Studienbeginn. Dies entspricht der CHMP-Empfehlung zur Auswertung der Behinderungsprogression bei MS (8). Die Erhebung des EDSS-Wertes erfolgte während der Screening-Visite, zu Baseline und danach alle drei Monate während der Studie sowie bei einem Krankheitsschub. Eine bestätigte Progression der Behinderung (3- or 6-month-confirmed disability progression, 3- or 6-CDP) war definiert als eine auch drei bzw. sechs Monate nach der ersten Erhöhung fortdauernde Erhöhung des EDSS-Werts im Vergleich zum Studienbeginn. Außerdem musste jeder EDSS-Punktwert, der während der drei bzw. sechs Monate ermittelt wurde, ebenfalls die Kriterien für eine Progression erfüllen. EDSS-Werte, die während eines Schubs erhoben wurden, fanden keine Berücksichtigung. Zu diesem Zwecke wurde die maximale Schubdauer auf 90 Tage festgelegt. Im Dossier berichtet werden: • die Inzidenz einer nach drei bzw. sechs Monaten bestätigten Behinderungsprogression (binäre Analyse) mit den Risikoschätzern OR, RR und RD, jeweils mit den dazugehörigen Konfidenzintervallen und p-Werten; • die Zeit bis zur ersten nach drei bzw. sechs Monaten bestätigten Behinderungsprogression bzw. das Risiko für eine nach drei bzw. sechs Monaten bestätigte Behinderungsprogression aus einer Kaplan-Meier-Schätzung (Hazard Ratio basierend auf Cox-Regression), inklusive Kaplan-Meier-Schätzer für den Anteil der Patienten ohne bestätigte Behinderungsprogression.

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 4-45: Bewertung des Verzerrungspotenzials für den Endpunkt *Behinderungsprogression* 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**
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Behinderungsprogression* als niedrig bewertet, da es auf Studienebene niedrig war, die Endpunkterheber (Prüfarzt und unabhängiger EDSS-Bewerter) verblindet waren, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und sonstige, das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden. Daher kann hier von einem objektivierbaren Endpunkt ausgegangen werden und es ist nur mit einem niedrigen daraus resultierenden Verzerrungspotenzial zu rechnen.

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

Inzidenz der bestätigten Behinderungsprogression

Tabelle 4-46: Ergebnisse für den Endpunkt *Inzidenz der bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 1.099)	**BSC (N = 546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Inzidenz der nach 3 Monaten bestätigten Behinderungsprogression					
N'	1.099	546			
Patienten mit Behinderungs-progression – n (%)	289 (26,3)	173 (31,7)	0,77 [0,61; 0,96] 0,022	0,83 [0,71; 0,97] 0,021	-0,05 [-0,10; -0,01] 0,024
Inzidenz der nach 6 Monaten bestätigten Behinderungsprogression					
N'	1.099	546			
Patienten mit Behinderungs-progression – n (%)	219 (19,9)	139 (25,5)	0,73 [0,57; 0,93] 0,011	0,78 [0,65; 0,94] 0,010	-0,06 [-0,10; -0,01] 0,013

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 1 zu Modul 4, Tabellen 4.1.1 und 5.1.1.

Tabelle 4-47: Ergebnisse für den Endpunkt *Inzidenz der bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N = 61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Inzidenz der nach 3 Monaten bestätigten Behinderungsprogression					
N'	127	61			
Patienten mit Behinderungs-progression – n (%)	30 (23,6)	19 (31,1)	0,68 [0,35; 1,35] 0,272	0,76 [0,47; 1,23] 0,266	-0,08 [-0,21; 0,06] 0,284
Inzidenz der nach 6 Monaten bestätigten Behinderungsprogression					
N'	127	61			
Patienten mit Behinderungs-progression – n (%)	24 (18,9)	16 (26,2)	0,66 [0,32; 1,35] 0,252	0,72 [0,41; 1,25] 0,246	-0,07 [-0,20; 0,06] 0,268

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 2 zu Modul 4, Tabellen 4.1.1 und 5.1.1.

In der Gesamtpopulation war die *Inzidenz einer bestätigten Behinderungsprogression* unter Siponimod statistisch signifikant geringer als unter BSC. Dies gilt sowohl für eine nach drei als auch für eine nach sechs Monaten bestätigte Behinderungsprogression (Tabelle 4-46).

In der Auswertung der Subpopulation B wurde kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen festgestellt (Tabelle 4-47).

Zeit bis zur bestätigten Behinderungsprogression

Tabelle 4-48: Ergebnisse für den Endpunkt *Zeit bis zur bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**HR [95%-KI][b]**	**p-Wert**
Zeit bis zur nach 3 Monaten bestätigten Behinderungsprogression				
N'	1.096	545		
Anzahl der Patienten mit Ereignis	288	173		
Median (in Tagen) [95%-KI][a]	n. b.	n. b.		
Anteil progressionsfreier Patienten nach 24 Monaten in % – KM-Schätzer [95%-KI][a]	68,68 [65,45; 71,91]	64,39 [59,78; 68,99]		
Anteil progressionsfreier Patienten zu Studienende in % – KM-Schätzer [95%-KI][a]	64,21 [58,98; 69,44]	56,36 [48,50; 64,23]		
			0,79 [0,65; 0,95]	0,013
Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression				
N'	1.096	545		
Anzahl der Patienten mit Ereignis	218	139		
Median (in Tagen) [95%-KI][a]	n. b.	n. b.		
Anteil progressionsfreier Patienten nach 24 Monaten in % – KM-Schätzer [95%-KI][a]	76,47 [73,58; 79,35]	70,81 [66,47; 75,14]		
Anteil progressionsfreier Patienten zu Studienende in % – KM-Schätzer [95%-KI][a]	75,33 [72,20; 78,45]	70,02 [65,46; 74,57]		
			0,74 [0,60; 0,92]	0,006

HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar (Überlebenskurve erreicht das 50%-Level nicht).

a: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

b: Cox proportional hazards model.

Quelle: Anhang 1 zu Modul 4, Tabellen 4.2.1 und 5.2.1.

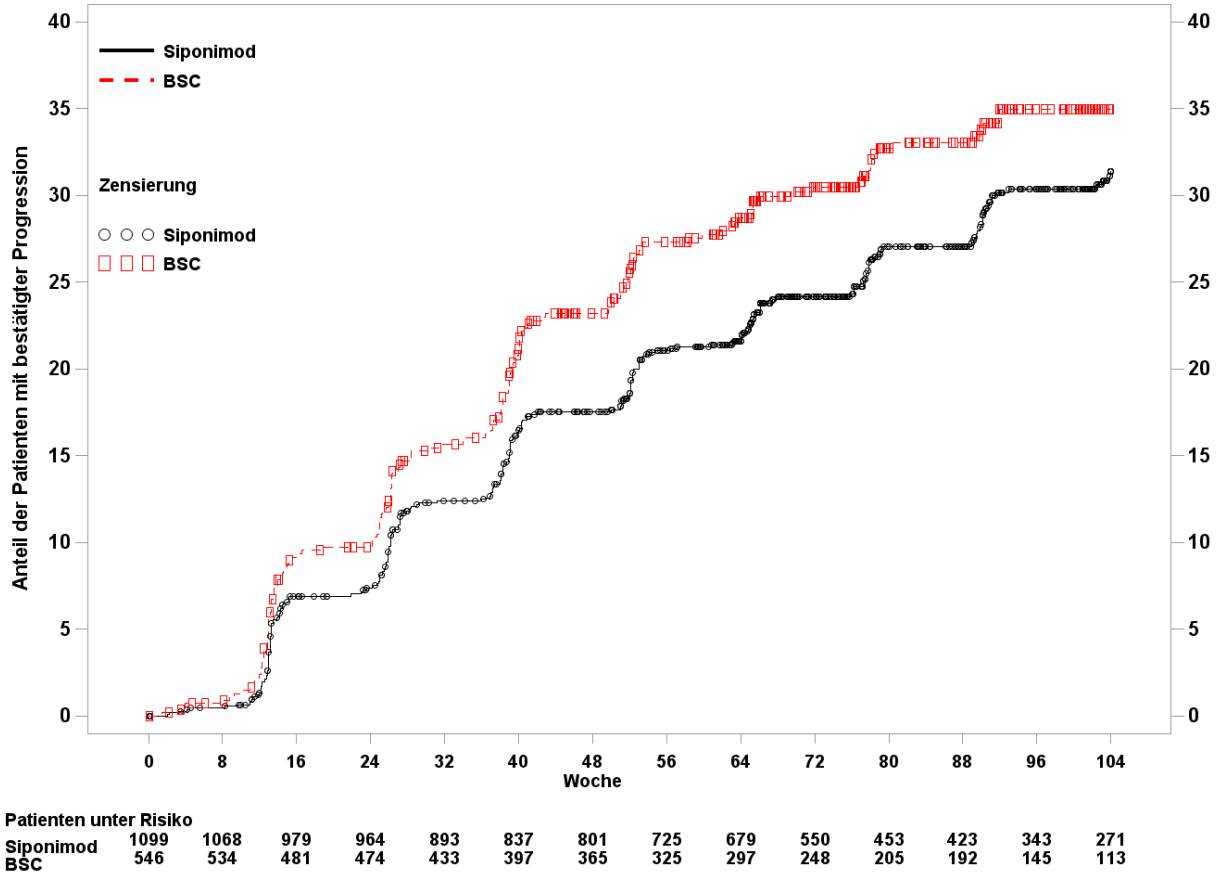


Abbildung 4: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur nach 3 Monaten bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

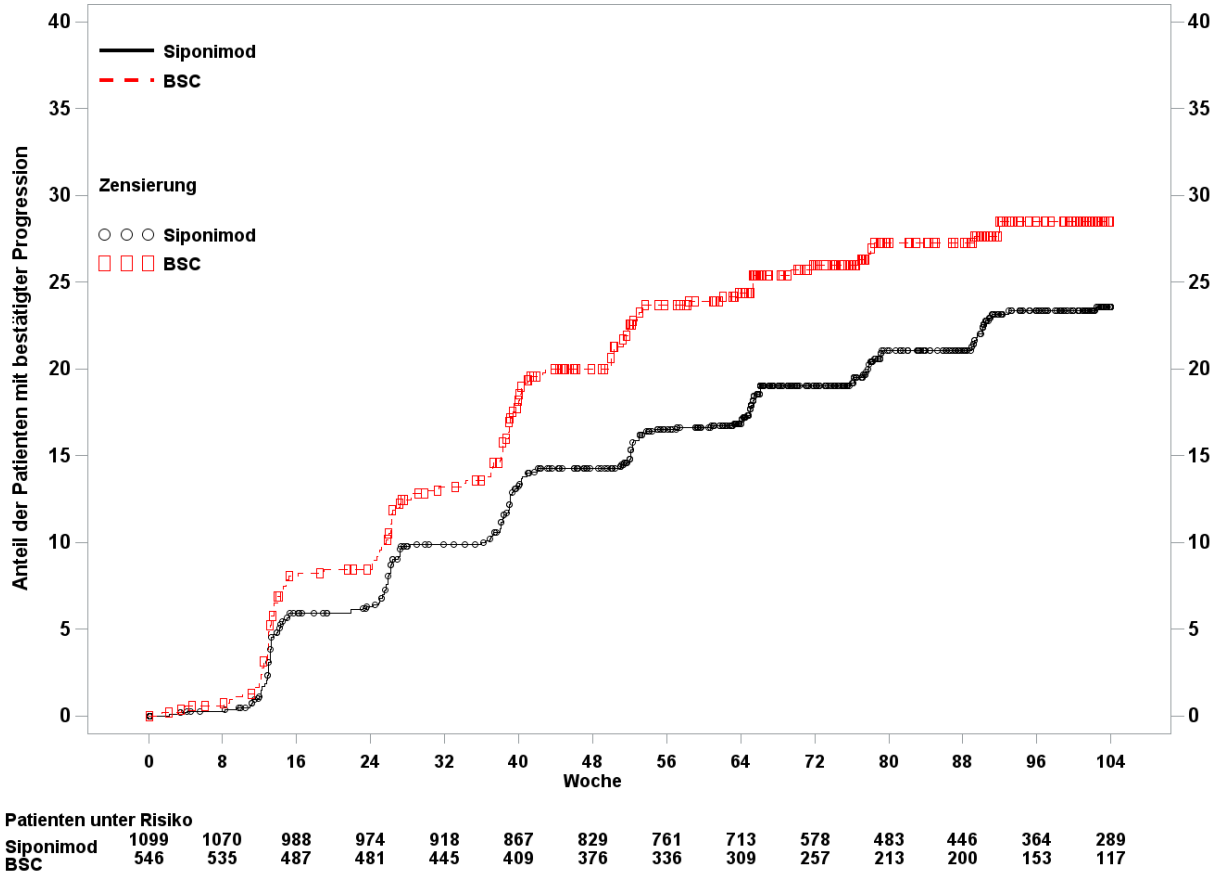


Abbildung 5: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Tabelle 4-49: Ergebnisse für den Endpunkt *Zeit bis zur bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**HR [95%-KI][b]**	**p-Wert**
Zeit bis zur nach 3 Monaten bestätigten Behinderungsprogression				
N'	127	61		
Anzahl der Patienten mit Ereignis	30	19		
Median (in Tagen) [95%-KI][a]	n. b.	n. b.		
Anteil progressionsfreier Patienten nach 24 Monaten in % – KM-Schätzer [95%-KI][a]	70,98 [61,77; 80,18]	63,25 [47,77; 78,73]		
Anteil progressionsfreier Patienten zu Studienende in % – KM-Schätzer [95%-KI][a]	70,98 [61,77; 80,18]	57,50 [39,79; 75,20]		
			0,65 [0,34; 1,23]	0,185
Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression				
N'	127	61		
Anzahl der Patienten mit Ereignis	24	16		
Median (in Tagen) [95%-KI][a]	n. b.	n. b.		
Anteil progressionsfreier Patienten nach 24 Monaten in % – KM-Schätzer [95%-KI][a]	76,79 [68,25; 85,33]	66,05 [50,38; 81,72]		
Anteil progressionsfreier Patienten zu Studienende in % – KM-Schätzer [95%-KI][a]	76,79 [68,25; 85,33]	66,05 [50,38; 81,72]		
			0,57 [0,28; 1,16]	0,121

HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar (Überlebenskurve erreicht das 50%-Level nicht).

a: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

b: Cox proportional hazards model.

Quelle: Anhang 2 zu Modul 4, Tabellen 4.2.1 und 5.2.1.

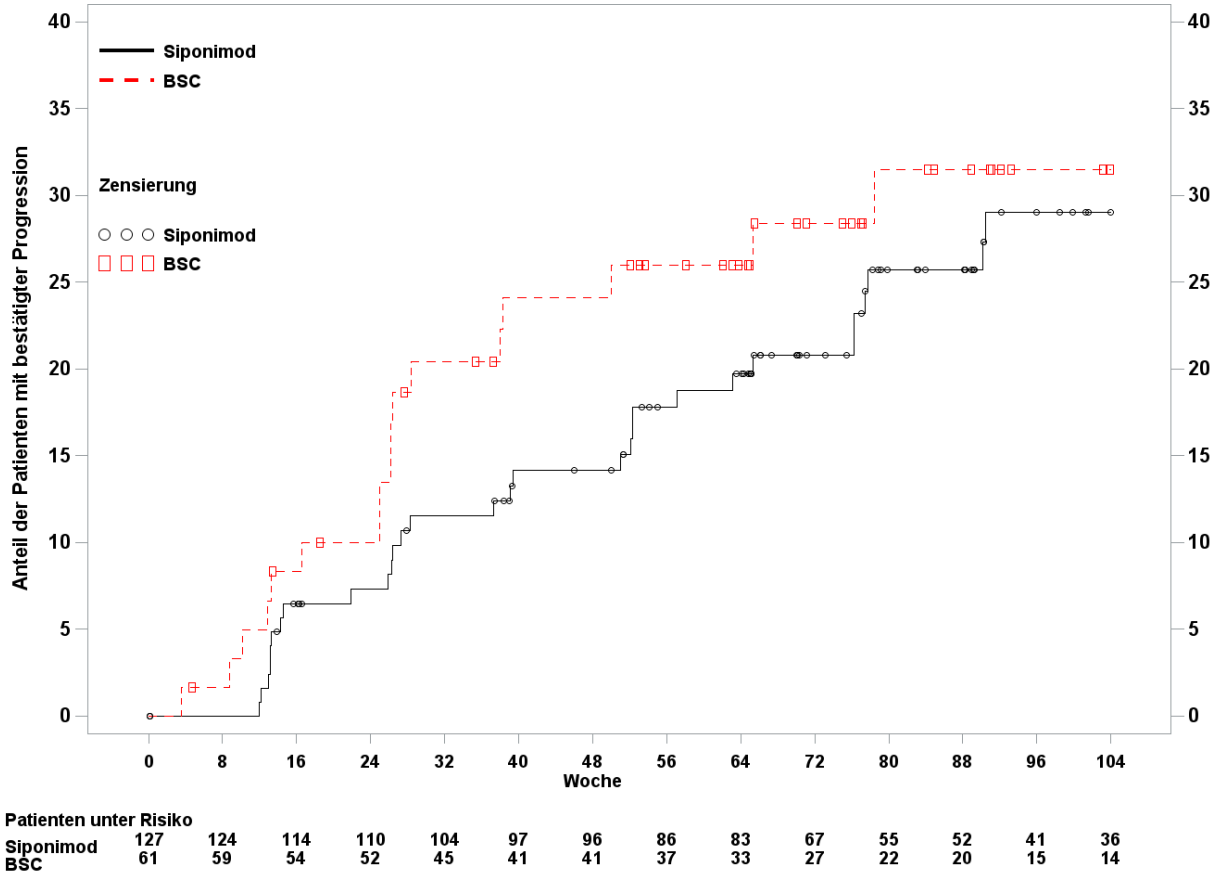


Abbildung 6: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur nach 3 Monaten bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

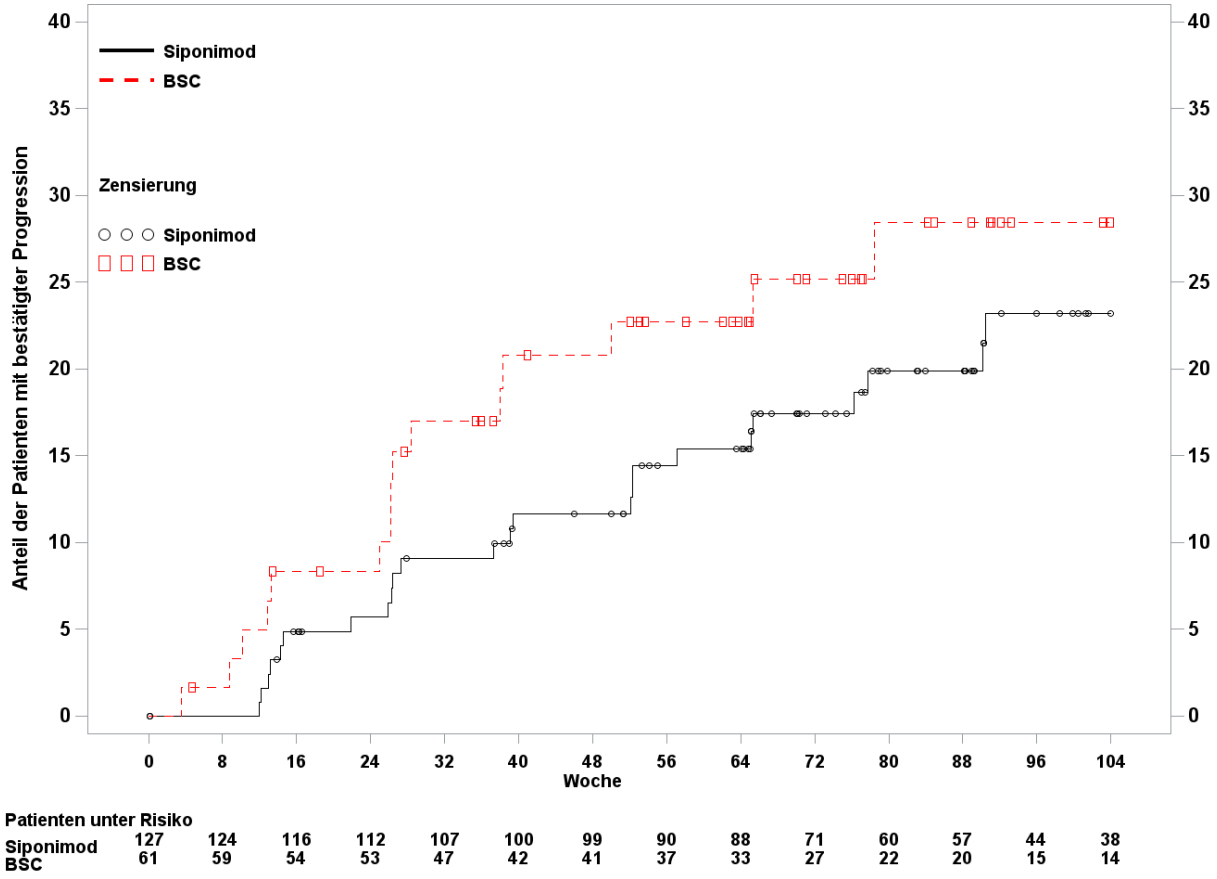


Abbildung 7: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur nach 6 Monaten bestätigten Behinderungsprogression* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

In der Auswertung der Gesamtpopulation zeigte sich ein statistisch signifikanter Vorteil im Endpunkt *Zeit bis zur nach drei bzw. sechs Monaten bestätigten Progression der Behinderung* zugunsten von Siponimod. Das Risiko einer Behinderungsprogression wurde hier um 21 % gesenkt (HR 0,79; 95%-KI [0,65; 0,95]; Tabelle 4-48).

In der Subpopulation B ergab die Auswertung des Endpunkts *Zeit bis zur nach drei bzw. sechs Monaten bestätigten Progression der Behinderung* keinen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen (Tabelle 4-49).

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

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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 Behinderungsprogression (MSFC) – RCT

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

Tabelle 4-50: Operationalisierung des Endpunkts *Behinderungsprogression (MSFC)*

Studie	Operationalisierung
EXPAND	Die Behinderungsprogression wurde anhand des MSFC und seiner Einzelkomponenten erhoben. Der MSFC wurde zu Baseline und dann alle drei Monate (T25W und 9-HPT) bzw. alle sechs Monate (PASAT-3) erhoben. Der MSFC-z kann dementsprechend für Baseline und dann alle sechs Monate ermittelt werden. Im MSFC-z und im PASAT-3 zeigt eine Erhöhung des Punktwerts eine Verbesserung an, im 9-HPT und im T25W zeigt eine Erhöhung des Punktwerts eine Verschlechterung an. Im Dossier berichtet werden: • die Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn für den MSFC-z und die Komponenten T25W, 9-HPT sowie PASAT-3 (jeweils dargestellt werden die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den zugehörigen Standardabweichungen sowie die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht); • die Inzidenz einer nach drei Monaten bestätigten Verschlechterung im T25W um mindestens 20 % (binäre Analyse) mit den Risikoschätzern OR, RR und RD, jeweils mit den dazugehörigen Konfidenzintervallen und p-Werten; • die Zeit bis zur ersten nach drei Monaten bestätigten Verschlechterung im T25W um mindestens 20 % bzw. das Risiko für eine nach drei Monaten bestätigte Verschlechterung im T25W um mindestens 20 % aus einer Kaplan-Meier-Schätzung (Hazard Ratio basierend auf Cox-Regression), inklusive Kaplan-Meier-Schätzer für den Anteil der Patienten ohne bestätigte Verschlechterung im T25W.

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 4-51: Bewertung des Verzerrungspotenzials für den Endpunkt *Behinderungsprogression (MSFC)* 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
EXPAND	niedrig	ja	ja[a]/nein[b]	ja	ja	niedrig[a]/hoch[b]

a: alle Auswertungen in der Gesamtpopulation; PASAT-3 und 9-HPT in der Subpopulation B.
b: MSFCz und T25W in der Subpopulation B.

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Behinderungsprogression (MSFC)* als niedrig bewertet für die Gesamtpopulation und die Auswertungen des PASAT-3 und 9-HPT in der Subpopulation B. Das Verzerrungspotenzial war auf Studienebene niedrig, der Endpunkterheber (unabhängiger Bewerter) war verblindet, das ITT-Prinzip bei der Analyse wurde adäquat umgesetzt, eine ergebnisgesteuerte Berichterstattung war unwahrscheinlich und sonstige das Verzerrungspotenzial beeinflussende Punkte wurden nicht festgestellt. Der Nichtberücksichtigungsanteil liegt zu Monat 12 insgesamt bei unter 30 % mit Ausnahme der Auswertung des MSFC-z in der BSC-Gruppe der Subpopulation B (32,8 %). Der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen liegt unter 10 % mit Ausnahme der Auswertung des MSFC-z und des T25W in der Subpopulation B (13,9 % bzw. 13,8 %). Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %.

Es kann zwar von einem objektivierbaren Endpunkt ausgegangen werden. Für die Auswertung zu Monat 12 ist für den MSFC-z und den T25W in der Subpopulation B aber mit einem hohen Verzerrungspotenzial zu rechnen. Für alle Auswertungen in der Gesamtpopulation und für den PASAT-3 sowie den 9-HPT in der Subpopulation B ist das Verzerrungspotenzial niedrig.

Tabelle 4-52: Rücklaufquoten MSFC-Komponenten (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
MSFC-z			
Rückläufer Studienbeginn, n (%)	1.067 (97,1)	536 (98,2)	1.603 (97,4)
Rückläufer Monat 6, n (%)	959 (87,3)	482 (88,3)	1.441 (87,6)
Rückläufer Monat 12, n (%)	862 (78,4)	411 (75,3)	1.273 (77,4)
Rückläufer Monat 18, n (%)	605 (55,1)	285 (52,2)	890 (54,1)
Rückläufer Monat 24, n (%)	372 (33,8)	172 (31,5)	544 (33,1)
PASAT-3			
Rückläufer Studienbeginn, n (%)	1.082 (98,5)	538 (98,5)	1.620 (98,5)
Rückläufer Monat 6, n (%)	995 (90,5)	501 (91,8)	1.496 (90,9)
Rückläufer Monat 12, n (%)	906 (82,4)	440 (80,6)	1.346 (81,8)
Rückläufer Monat 18, n (%)	634 (57,7)	309 (56,6)	943 (57,3)
Rückläufer Monat 24, n (%)	395 (35,9)	191 (35,0)	586 (35,6)
9-HPT			
Rückläufer Studienbeginn, n (%)	1.089 (99,1)	545 (99,8)	1.634 (99,3)
Rückläufer Monat 3, n (%)	1.050 (95,5)	528 (96,7)	1.578 (95,9)
Rückläufer Monat 6, n (%)	1.021 (92,9)	518 (94,9)	1.539 (93,6)
Rückläufer Monat 9, n (%)	985 (89,6)	492 (90,1)	1.477 (89,8)
Rückläufer Monat 12, n (%)	928 (84,4)	453 (83,0)	1.381 (84,0)
Rückläufer Monat 15, n (%)	818 (74,4)	408 (74,7)	1.226 (74,5)
Rückläufer Monat 18, n (%)	654 (59,5)	316 (57,9)	970 (59,0)
Rückläufer Monat 21, n (%)	531 (48,3)	259 (47,4)	790 (48,0)
Rückläufer Monat 24, n (%)	404 (36,8)	195 (35,7)	599 (36,4)
T25W			
Rückläufer Studienbeginn, n (%)	1.090 (99,2)	544 (99,6)	1.634 (99,3)
Rückläufer Monat 3, n (%)	1.033 (94,0)	529 (96,9)	1.562 (95,0)
Rückläufer Monat 6, n (%)	987 (89,8)	504 (92,3)	1.491 (90,6)
Rückläufer Monat 9, n (%)	952 (86,6)	480 (87,9)	1.432 (87,1)
Rückläufer Monat 12, n (%)	895 (81,4)	425 (77,8)	1.320 (80,2)
Rückläufer Monat 15, n (%)	787 (71,6)	385 (70,5)	1.172 (71,2)
Rückläufer Monat 18, n (%)	624 (56,8)	294 (53,8)	918 (55,8)
Rückläufer Monat 21, n (%)	498 (45,3)	239 (43,8)	737 (44,8)
Rückläufer Monat 24, n (%)	385 (35,0)	176 (32,2)	561 (34,1)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**

9-HPT: 9 Hole Peg Test; MSFC: Multiple Sclerosis Functional Composite; N: Patienten gesamt; n: Patienten mit Ereignis; PASAT-3: Paced Auditory Serial Addition Test; T25W: Timed 25 Foot Walking Test.
Quelle: Anhang 1 zu Modul 4, Tabellen 6.1.0, 6.2.0, 6.3.0 und 6.4.0.

Tabelle 4-53: Rücklaufquoten MSFC-Komponenten (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
MSFC-z			
Rückläufer Studienbeginn, n (%)	125 (98,4)	59 (96,7)	184 (97,9)
Rückläufer Monat 6, n (%)	113 (89,0)	54 (88,5)	167 (88,8)
Rückläufer Monat 12, n (%)	103 (81,1)	41 (67,2)	144 (76,6)
Rückläufer Monat 18, n (%)	75 (59,1)	27 (44,3)	102 (54,3)
Rückläufer Monat 24, n (%)	48 (37,8)	17 (27,9)	65 (34,6)
PASAT-3			
Rückläufer Studienbeginn, n (%)	126 (99,2)	59 (96,7)	185 (98,4)
Rückläufer Monat 6, n (%)	115 (90,6)	57 (93,4)	172 (91,5)
Rückläufer Monat 12, n (%)	105 (82,7)	47 (77,0)	152 (80,9)
Rückläufer Monat 18, n (%)	77 (60,6)	31 (50,8)	108 (57,4)
Rückläufer Monat 24, n (%)	52 (40,9)	21 (34,4)	73 (38,8)
9-HPT			
Rückläufer Studienbeginn, n (%)	127 (100,0)	61 (100,0)	188 (100,0)
Rückläufer Monat 3, n (%)	121 (95,3)	58 (95,1)	179 (95,2)
Rückläufer Monat 6, n (%)	117 (92,1)	56 (91,8)	173 (92,0)
Rückläufer Monat 9, n (%)	113 (89,0)	54 (88,5)	167 (88,8)
Rückläufer Monat 12, n (%)	108 (85,0)	49 (80,3)	157 (83,5)
Rückläufer Monat 15, n (%)	100 (78,7)	43 (70,5)	143 (76,1)
Rückläufer Monat 18, n (%)	82 (64,6)	29 (47,5)	111 (59,0)
Rückläufer Monat 21, n (%)	71 (55,9)	25 (41,0)	96 (51,1)
Rückläufer Monat 24, n (%)	52 (40,9)	20 (32,8)	72 (38,3)
T25W			
Rückläufer Studienbeginn, n (%)	126 (99,2)	61 (100,0)	187 (99,5)
Rückläufer Monat 3, n (%)	119 (93,7)	59 (96,7)	178 (94,7)
Rückläufer Monat 6, n (%)	114 (89,8)	55 (90,2)	169 (89,9)
Rückläufer Monat 9, n (%)	111 (87,4)	52 (85,2)	163 (86,7)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
Rückläufer Monat 12, n (%)	107 (84,3)	43 (70,5)	150 (79,8)
Rückläufer Monat 15, n (%)	100 (78,7)	39 (63,9)	139 (73,9)
Rückläufer Monat 18, n (%)	79 (62,2)	27 (44,3)	106 (56,4)
Rückläufer Monat 21, n (%)	66 (52,0)	23 (37,7)	89 (47,3)
Rückläufer Monat 24, n (%)	50 (39,4)	17 (27,9)	67 (35,6)

9-HPT: 9 Hole Peg Test; MSFC: Multiple Sclerosis Functional Composite; N: Patienten gesamt; n: Patienten mit Ereignis; PASAT-3: Paced Auditory Serial Addition Test; T25W: Timed 25 Foot Walking Test.
Quelle: Anhang 2 zu Modul 4, Tabellen 6.1.0, 6.2.0, 6.3.0 und 6.4.0.

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

Veränderung im MSFC und seinen Komponenten seit Studienbeginn

Tabelle 4-54: Ergebnisse für den Endpunkt *Behinderungsprogression (MSFC)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
MSFC-z – Veränderung seit Studienbeginn				
N'	972	496		
Mittelwert zu Studienbeginn (SD)	0,04 (0,67)	0,05 (0,65)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,00 (0,01)	-0,02 (0,02)	0,02 [-0,03; 0,07]	0,373
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,07 (0,02)	-0,07 (0,03)	-0,01 [-0,08; 0,07]	0,856
PASAT-3 – Veränderung seit Studienbeginn				
N'	1.011	509		
Mittelwert zu Studienbeginn (SD)	39,47 (14,39)	39.46 (13,86)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	2,25 (0,30)	1,80 (0,43)	0,45 [-0,57; 1,47]	0,387
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	2,98 (0,40)	2,38 (0,58)	0,59 [-0,79; 1,98]	0,401

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
9-HPT – Veränderung seit Studienbeginn				
N'	1.056	535		
Mittelwert zu Studienbeginn (SD)	34,07 (18,37)	34,48 (19,92)		
Monat 12: Adjustierte mittlere Veränderung (SE)[b]	1,01 (0,44)	1,01 (0,63)	0,00 [-1,50; 1.51]	0,996
Monat 24: Adjustierte mittlere Veränderung (SE)[b]	2,31 (0,58)	3,52 (0,82)	-1,20 [-3,18; 0,78]	0,233
T25W – Veränderung seit Studienbeginn				
N'	1.045	532		
Mittelwert zu Studienbeginn (SD)	16,40 (18,95)	15,24 (17,81)		
Monat 12: Adjustierte mittlere Veränderung (SE)[b]	4,28 (0,67)	4,24 (0,95)	0,03 [-2,24; 2,31]	0,976
Monat 24: Adjustierte mittlere Veränderung (SE)[b]	9,40 (1,05)	7,02 (1,52)	2,38 [-1,25; 6,01]	0,199

9-HPT: 9 Hole Peg Test; KI: Konfidenzintervall; MSFC: Multiple Sclerosis Functional Composite; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; PASAT-3: Paced Auditory Serial Addition Test; SD: Standard Deviation; SE: Standard Error; T25W: Timed 25 Foot Walking Test.

a: Positive Werte zeigen eine Verbesserung an.

b: Negative Werte zeigen eine Verbesserung an.

Quelle: Anhang 1 zu Modul 4, Tabellen 6.1.1, 6.2.1, 6.3.1 und 6.4.1.

Tabelle 4-55: Ergebnisse für den Endpunkt *Behinderungsprogression (MSFC)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
MSFC-z – Veränderung seit Studienbeginn				
N'	113	53		
Mittelwert zu Studienbeginn (SD)	-0,10 (0,65)	-0,11 (0,60)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,04 (0,04)	0,02 (0,06)	0,02 [-0,11; 0,15]	0,743
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,07 (0,05)	-0,10 (0,08)	0,03 [-0,17; 0,22]	0,789

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
PASAT-3 – Veränderung seit Studienbeginn				
N'	115	56		
Mittelwert zu Studienbeginn (SD)	37,80 (14,41)	34,57 (13,48)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	3,13 (0,87)	3,44 (1,29)	-0,31 [-3,39; 2,77]	0,842
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	2,27 (1,21)	0,96 (1,90)	1,31 [-3,16; 5,78]	0,563
9-HPT – Veränderung seit Studienbeginn				
N'	124	60		
Mittelwert zu Studienbeginn (SD)	39,67 (22,21)	34,81 (17,19)		
Monat 12: Adjustierte mittlere Veränderung (SE)[b]	-1,03 (1,16)	0,00 (1,70)	-1,03 [-5,13; 3,07]	0,620
Monat 24: Adjustierte mittlere Veränderung (SE)[b]	0,69 (2,30)	3,59 (3,53)	-2,89 [-11,28; 5,49]	0,495
T25W – Veränderung seit Studienbeginn				
N'	121	60		
Mittelwert zu Studienbeginn (SD)	16,45 (15,80)	14,44 (12,19)		
Monat 12: Adjustierte mittlere Veränderung (SE)[b]	6,14 (1,76)	7,90 (2,61)	-1,76 [-8,05; 4,54]	0,579
Monat 24: Adjustierte mittlere Veränderung (SE)[b]	12,31 (2,94)	13,05 (4,48)	-0,74 [-11,55; 10,08]	0,891

9-HPT: 9 Hole Peg Test; KI: Konfidenzintervall; MSFC: Multiple Sclerosis Functional Composite; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; PASAT-3: Paced Auditory Serial Addition Test. SD: Standard Deviation; SE: Standard Error; T25W: Timed 25 Foot Walking Test.

a: Positive Werte zeigen eine Verbesserung an.

b: Negative Werte zeigen eine Verbesserung an.

Quelle: Anhang 2 zu Modul 4, Tabellen 6.1.1, 6.2.1, 6.3.1 und 6.4.1.

In der Auswertung des MSFC bzw. seiner Komponenten zeigte sich weder in der Gesamtpopulation noch in der Subpopulation B ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen (Tabelle 4-54 und Tabelle 4-55).

Inzidenz einer bestätigten Verschlechterung im T25W

Tabelle 4-56: Ergebnisse für den Endpunkt *Inzidenz einer bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 1.099)	BSC (N = 546)	OR [95%-KI][a] p-Wert	RR [95%-KI][a] p-Wert	RD [95%-KI][a] p-Wert
Inzidenz einer nach 3 Monaten bestätigten Verschlechterung im T25W					
N'	1.099	546			
Patienten mit Verschlechterung ≥ 20 % – n (%)	436 (39,7)	226 (41,4)	0,93 [0,76; 1,15] 0,502	0,96 [0,85; 1,08] 0,501	-0,02 [-0,07; 0,03] 0,504

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; T25W: Timed 25 Foot Walking Test.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 1 zu Modul 4, Tabelle 6.5.1.

Tabelle 4-57: Ergebnisse für den Endpunkt *Inzidenz einer bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 127)	BSC (N = 61)	OR [95%-KI][a] p-Wert	RR [95%-KI][a] p-Wert	RD [95%-KI][a] p-Wert
Inzidenz einer nach 3 Monaten bestätigten Verschlechterung im T25W					
N'	127	61			
Patienten mit Verschlechterung ≥ 20 % – n (%)	62 (48,8)	28 (45,9)	1,12 [0,61; 2,07] 0,708	1,06 [0,77; 1,47] 0,711	0,03 [-0,12; 0,18] 0,707

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko: T25W: Timed 25 Foot Walking Test.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 2 zu Modul 4, Tabelle 6.5.1.

In beiden Populationen (Gesamtpopulation und Subpopulation B) war die Inzidenz der T25W-Verschlechterung zwischen den Behandlungsgruppen nicht statistisch signifikant unterschiedlich (Tabelle 4-56 und Tabelle 4-57).

Zeit bis zur ersten bestätigten Verschlechterung im T25W

Tabelle 4-58: Ergebnisse für den Endpunkt *Zeit bis zur ersten bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**HR [95%-KI][b]**	**p-Wert**
Zeit bis zur ersten nach 3 Monaten bestätigten Verschlechterung im T25W				
N'	1.087	543		
Anzahl der Patienten mit Ereignis	432	225		
Median (in Wochen) [95%-KI][a]	821 [722; n. b.]	804 [565; n. b.]		
Anteil Patienten ohne Verschlechterung zu Monat 24 in % – KM-Schätzer [95%-KI][a]	52,50 [48,87; 56,14]	50,37 [45,25; 55,48]		
Anteil Patienten ohne Verschlechterung zu Studienende in % – KM-Schätzer [95%-KI][a]	47,99 [43,69; 52,30]	44,16 [36,51; 51,81]		
			0,94 [0,80; 1,10]	0,440

HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar; T25W: Timed 25-Foot Walking Test.

a: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

b: Cox proportional hazards model.

Quelle: Anhang 1 zu Modul 4, Tabelle 6.6.1.

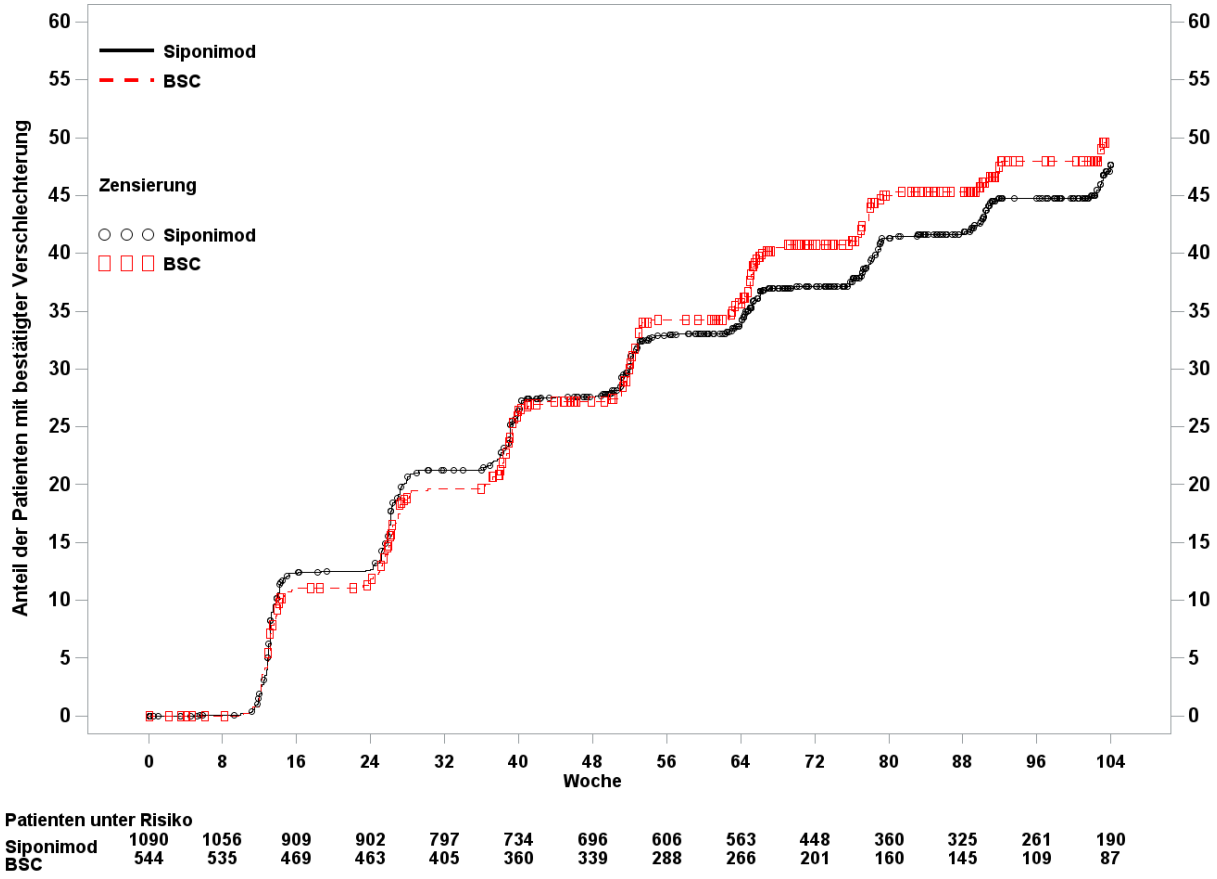


Abbildung 8: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur ersten nach 3 Monaten bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Tabelle 4-59: Ergebnisse für den Endpunkt *Zeit bis zur ersten bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**HR [95%-KI][b]**	**p-Wert**
Zeit bis zur ersten nach 3 Monaten bestätigten Verschlechterung im T25W				
N'	126	61		
Anzahl der Patienten mit Ereignis	61	28		
Median (in Wochen) [95%-KI][a]	633 [444; 819]	457 [371; n. b.]		
Anteil Patienten ohne Verschlechterung zu Monat 24 in % – KM-Schätzer [95%-KI][a]	43,65 [33,03; 54,28]	48,42 [34,39; 62,45]		
Anteil Patienten ohne Verschlechterung zu Studienende in % – KM-Schätzer [95%-KI][a]	32,91 [19,31; 46,51]	43,04 [27,09; 58,99]		
			1,27 [0,76; 2,11]	0,367

HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar; T25W: Timed 25-Foot Walking Test.

a: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

b: Cox proportional hazards model.

Quelle: Anhang 2 zu Modul 4, Tabelle 6.6.1.

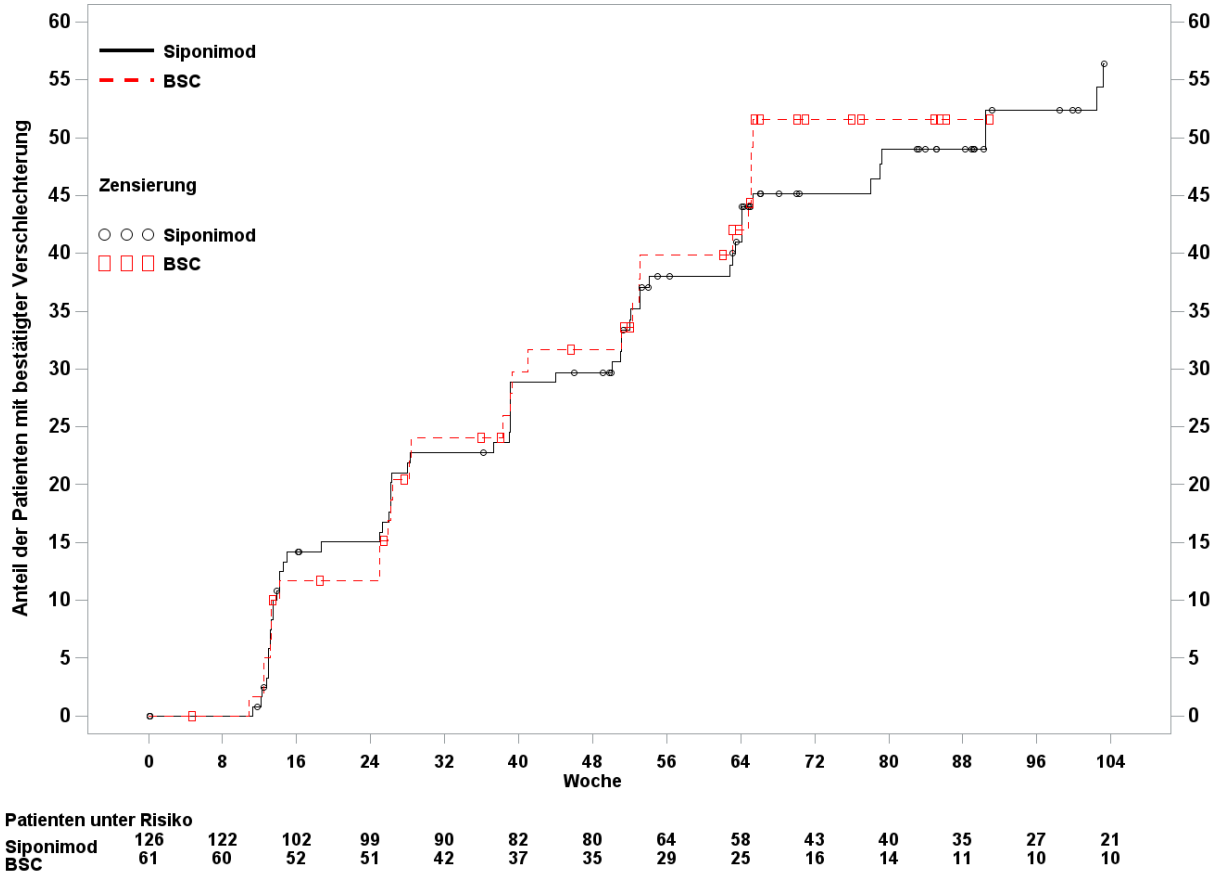


Abbildung 9: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zur ersten nach 3 Monaten bestätigten Verschlechterung im T25W* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

Das Risiko für eine T25W-Verschlechterung war in den Behandlungsgruppen nicht statistisch signifikant unterschiedlich, weder in der Gesamtpopulation noch in der Subpopulation B (Tabelle 4-58 und Tabelle 4-59).

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Auf Endpunktebene liegen in den Auswertungen des MSFCz und T25W in der Subpopulation B ein hohes Verzerrungspotenzial vor, das hier ggf. die Übertragbarkeit einschränkt.

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 Schweregrad der kognitiven Dysfunktion (SDMT) – RCT

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

Tabelle 4-60: Operationalisierung des Endpunkts *Schweregrad der kognitiven Dysfunktion (SDMT)*

Studie	Operationalisierung
EXPAND	Der Schweregrad der kognitiven Dysfunktion wurde anhand des SDMT (Symbol Digit Modalities Test) ermittelt. Der SDMT wurde zu Baseline und dann alle sechs Monate erhoben. Eine Erhöhung des Punktwerts zeigt eine Verbesserung an. Im Dossier berichtet werden: • die Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn für den SDMT (dargestellt werden die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den dazugehörigen Standardabweichungen sowie die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht); • die Inzidenz einer Verbesserung bzw. einer Verschlechterung im SDMT um mindestens 4 Punkte nach 12 Monaten gegenüber dem Baselinewert (binäre Analyse) mit den Risikoschätzern OR, RR und RD, jeweils mit den dazugehörigen Konfidenzintervallen und p-Werten. Fehlende Werte nach 12 Monaten wurden wie folgt imputiert: LOCF, falls der SDMT zu mindestens einem Zeitpunkt nach Baseline und vor der 12-Monats-Visite durchgeführt wurde, sonst Non-Responder-Imputation.

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 4-61: Bewertung des Verzerrungspotenzials für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkterheber (unabhängiger Bewerter) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und sonstige das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden. Der Nichtberücksichtigungsanteil liegt zu Monat 12 insgesamt unter 30 %, der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen unter 10 %. Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden und für die Auswertung zu Monat 12 ist nur mit einem niedrigen Verzerrungspotenzial zu rechnen.

Tabelle 4-62: Rücklaufquoten SDMT (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
SDMT			
Rückläufer Studienbeginn, n (%)	1.090 (99,2)	541 (99,1)	1.631 (99,1)
Rückläufer Monat 6, n (%)	1.004 (91,4)	512 (93,8)	1.516 (92,2)
Rückläufer Monat 12, n (%)	924 (84,1)	448 (82,1)	1.372 (83,4)
Rückläufer Monat 18, n (%)	644 (58,6)	313 (57,3)	957 (58,2)
Rückläufer Monat 24, n (%)	306 (27,8)	140 (25,6)	446 (27,1)
Rückläufer Monat 30, n (%)	87 (7,9)	35 (6,4)	122 (7,4)

N: Patienten gesamt; n: Patienten mit Ereignis; SDMT: Symbol Digital Modalities Test.
Quelle: Anhang 1 zu Modul 4, Tabelle 7.1.0.

Tabelle 4-63: Rücklaufquoten SDMT (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
SDMT			
Rückläufer Studienbeginn, n (%)	127 (100,0)	60 (98,4)	187 (99,5)
Rückläufer Monat 6, n (%)	116 (91,3)	56 (91,8)	172 (91,5)
Rückläufer Monat 12, n (%)	107 (84,3)	48 (78,7)	155 (82,4)
Rückläufer Monat 18, n (%)	79 (62,2)	30 (49,2)	109 (58,0)
Rückläufer Monat 24, n (%)	41 (32,3)	15 (24,6)	56 (29,8)
Rückläufer Monat 30, n (%)	14 (11,0)	3 (4,9)	17 (9,0)

N: Patienten gesamt; n: Patienten mit Ereignis; SDMT: Symbol Digital Modalities Test.
Quelle: Anhang 2 zu Modul 4, Tabelle 7.1.0.

Stellen Sie die Ergebnisse für den Endpunkt Schweregrad der kognitiven Dysfunktion (SDMT) für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Veränderung im SDMT seit Studienbeginn

Tabelle 4-64: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
SDMT – Veränderung seit Studienbeginn				
N'	1.024	517		
Mittelwert zu Studienbeginn (SD)	38,92 (14,00)	39,8 (13,20)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,15 (0,29)	-0,94 (0,38)	1,09 [0,23; 1,95]	0,013
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	1,42 (0,42)	-1,06 (0,59)	2,48 [1,11; 3,85]	< 0,001

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SDMT: Symbol Digital Modalities Test; SD: Standard Deviation; SE: Standard Error.
a: Positive Werte zeigen eine Verbesserung an.
Quelle: Anhang 1 zu Modul 4, Tabelle 7.1.1.

Tabelle 4-65: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
SDMT – Veränderung seit Studienbeginn				
N'	118	57		
Mittelwert zu Studienbeginn (SD)	36,54 (13,85)	37,07 (12,14)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,16 (0,93)	-3,04 (1,20)	2,89 [0,33; 5,44]	0,027
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,34 (1,31)	-4,11 (1,97)	3,77 [-0,70; 8,24]	0,097

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SDMT: Symbol Digital Modalities Test; SD: Standard Deviation; SE: Standard Error.

a: Positive Werte zeigen eine Verbesserung an.

Quelle: Anhang 2 zu Modul 4, Tabelle 7.1.1.

In der Auswertung der Veränderung des SDMT seit Studienbeginn zeigte sich ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten von Siponimod zu Monat 12. Dabei zeigte sich unter BSC eine Verschlechterung im SDMT, unter Siponimod hingegen blieb der Wert weitestgehend stabil. Dies gilt gleichermaßen für die Gesamtpopulation wie für die Subpopulation B. Die Auswertung zu Monat 24 ist wegen sehr geringer Rücklaufquoten nicht zuverlässig interpretierbar (Tabelle 4-64 und Tabelle 4-65).

Inzidenz der Verbesserung bzw. Verschlechterung im SDMT nach 12 Monaten

Tabelle 4-66: Ergebnisse für den Endpunkt *Inzidenz der Verbesserung bzw. Verschlechterung im SDMT nach 12 Monaten* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 1.099)	BSC (N = 546)	OR [95%-KI][a] p-Wert	RR [95%-KI][a] p-Wert	RD [95%-KI][a] p-Wert
Inzidenz der Verbesserung im SDMT nach 12 Monaten					
N'	1.099	546			
Patienten mit Verbesserung um ≥ 4 Punkte – n (%)	340 (30,9)	130 (23,8)	1,43 [1,13; 1,81] 0,003	1,30 [1,09; 1,55] 0,003	0,07 [0,03; 0,12] 0,002
Inzidenz der Verschlechterung im SDMT nach 12 Monaten					
N'	1.099	546			
Patienten mit Verschlechterung um ≥ 4 Punkte – n (%)	291 (26,5)	185 (33,9)	0,70 [0,56; 0,88] 0,002	0,78 [0,67; 0,91] 0,002	-0,07 [-0,12; -0,03] 0,002

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; SDMT: Symbol Digital Modalities Test.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 1 zu Modul 4, Tabellen 7.2.1 und 7.3.1.

Tabelle 4-67: Ergebnisse für den Endpunkt *Inzidenz der Verbesserung bzw. Verschlechterung im SDMT nach 12 Monaten* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 127)	BSC (N = 61)	OR [95%-KI][a] p-Wert	RR [95%-KI][a] p-Wert	RD [95%-KI][a] p-Wert
Inzidenz der Verbesserung im SDMT nach 12 Monaten					
N'	127	61			
Patienten mit Verbesserung um ≥ 4 Punkte – n (%)	38 (29,9)	16 (26,2)	1,20 [0,61; 2,38] 0,601	1,14 [0,69; 1,88] 0,604	0,04 [-0,10; 0,17] 0,595

Inzidenz der Verschlechterung im SDMT nach 12 Monaten					
N'	127	61			
Patienten mit Verschlechterung um ≥ 4 Punkte – n (%)	41 (32,3)	23 (37,7)	0,79 [0,42; 1,49] 0,463	0,86 [0,57; 1,29] 0,457	-0,05 [-0,20; 0,09] 0,468

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; SDMT: Symbol Digital Modalities Test.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 2 zu Modul 4, Tabellen 7.2.1 und 7.3.1.

In der Gesamtpopulation, nicht jedoch in der Subpopulation B, zeigte sich unter Siponimod ein statistisch signifikant höherer Anteil an Patienten mit einer relevanten Verbesserung im SDMT zu Monat 12 und ein statistisch signifikant niedrigerer Anteil an Patienten mit einer relevanten Verschlechterung zu Monat 12 (Tabelle 4-66 und Tabelle 4-67).

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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 Schweregrad der kognitiven Dysfunktion (BVMT-R) – RCT

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

Tabelle 4-68: Operationalisierung des Endpunkts *Schweregrad der kognitiven Dysfunktion (BVMT-R)*

Studie	Operationalisierung
EXPAND	Der Schweregrad der kognitiven Dysfunktion wurde zusätzlich anhand des BVMT-R (Brief Visuospatial Memory Test – Revised) ermittelt. Der BVMT-R wurde zu Baseline und dann alle sechs Monate erhoben. Beim BVMT-R zeigt eine Zunahme der Werte eine Verbesserung an. Im Dossier berichtet wird: • die Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn für den BVMT-R Delayed Recall Score und den BVMT-R Total Recall Score (dargestellt werden jeweils die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den zugehörigen Standardabweichungen sowie die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht).

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 4-69: Bewertung des Verzerrungspotenzials für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkterheber (unabhängiger Bewerter) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und sonstige das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden. Der Nichtberücksichtigungsanteil liegt zu Monat 12 insgesamt unter 30 %, der Unterschied in den

Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen unter 10 %. Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden und für die Auswertung zu Monat 12 ist nur mit einem niedrigen Verzerrungspotenzial zu rechnen.

Tabelle 4-70: Rücklaufquoten BVMT-R (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
BVMT-R Delayed Recall Score			
Rückläufer Studienbeginn, n (%)	1.069 (97,3)	525 (96,2)	1.594 (96,9)
Rückläufer Monat 6, n (%)	987 (89,8)	504 (92,3)	1.491 (90,6)
Rückläufer Monat 12, n (%)	917 (83,4)	445 (81,5)	1.362 (82,8)
Rückläufer Monat 18, n (%)	637 (58,0)	306 (56,0)	943 (57,3)
Rückläufer Monat 24, n (%)	301 (27,4)	138 (25,3)	439 (26,7)
Rückläufer Monat 30, n (%)	87 (7,9)	34 (6,2)	121 (7,4)
BVMT-R Total Recall Score			
Rückläufer Studienbeginn, n (%)	1.078 (98,1)	534 (97,8)	1612 (98,0)
Rückläufer Monat 6, n (%)	990 (90,1)	509 (93,2)	1.499 (91,1)
Rückläufer Monat 12, n (%)	920 (83,7)	445 (81,5)	1.365 (83,0)
Rückläufer Monat 18, n (%)	639 (58,1)	307 (56,2)	946 (57,5)
Rückläufer Monat 24, n (%)	302 (27,5)	138 (25,3)	440 (26,7)
Rückläufer Monat 30, n (%)	87 (7,9)	34 (6,2)	121 (7,4)

BVMT-R: Brief Visuospatial Memory Test – Revised; N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 1 zu Modul 4, Tabellen 8.1.0 und 8.2.0.

Tabelle 4-71: Rücklaufquoten BVMT-R (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
BVMT-R Delayed Recall Score			
Rückläufer Studienbeginn, n (%)	123 (96,9)	59 (96,7)	182 (96,8)
Rückläufer Monat 6, n (%)	115 (90,6)	56 (91,8)	171 (91,0)
Rückläufer Monat 12, n (%)	107 (84,3)	48 (78,7)	155 (82,4)
Rückläufer Monat 18, n (%)	76 (59,8)	29 (47,5)	105 (55,9)
Rückläufer Monat 24, n (%)	39 (30,7)	15 (24,6)	54 (28,7)
Rückläufer Monat 30, n (%)	14 (11,0)	2 (3,3)	16 (8,5)
BVMT-R Total Recall Score			
Rückläufer Studienbeginn, n (%)	123 (96,9)	59 (96,7)	182 (96,8)
Rückläufer Monat 6, n (%)	115 (90,6)	56 (91,8)	171 (91,0)
Rückläufer Monat 12, n (%)	108 (85,0)	48 (78,7)	156 (83,0)
Rückläufer Monat 18, n (%)	77 (60,6)	30 (49,2)	107 (56,9)
Rückläufer Monat 24, n (%)	40 (31,5)	15 (24,6)	55 (29,3)
Rückläufer Monat 30, n (%)	14 (11,0)	2 (3,3)	16 (8,5)

BVMT-R: Brief Visuospatial Memory Test – Revised; N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 2 zu Modul 4, Tabellen 8.1.0 und 8.2.0.

Stellen Sie die Ergebnisse für den Endpunkt Schweregrad der kognitiven Dysfunktion (BVMT-R) für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Veränderung im BVMT-R seit Studienbeginn

Tabelle 4-72: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
BVMT-R Delayed Recall Score – Veränderung seit Studienbeginn				
N'	1.003	498		
Mittelwert zu Studienbeginn (SD)	8,12 (3,40)	8,07 (3,36)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,11 (0,09)	0,04 (0,12)	-0,15 [-0,42; 0,12]	0,284

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,44 (0,13)	0,43 (0,19)	0,01 [-0,43; 0,45]	0,964
BVMT-R Total Recall Score – Veränderung seit Studienbeginn				
N'	1.011	508		
Mittelwert zu Studienbeginn (SD)	20,79 (8,77)	20,6 (8,73)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,21 (0,22)	0,56 (0,30)	-0,35 [-1,01; 0,32]	0,307
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	1,78 (0,32)	1,33 (0,45)	0,45 [-0,59; 1,49]	0,394

BVMT-R: Brief Visuospatial Memory Test – Revised; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation [*Standardabweichung*]; SE: Standard Error [*Standardfehler*].

a: Positive Werte zeigen eine Verbesserung an.

Quelle: Anhang 1 zu Modul 4, Tabellen 8.1.1 und 8.2.1.

Tabelle 4-73: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
BVMT-R Delayed Recall Score – Veränderung seit Studienbeginn				
N'	113	56		
Mittelwert zu Studienbeginn (SD)	7,9 (3,29)	7,2 (3,28)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,53 (0,32)	0,36 (0,42)	-0,89 [-1,79; 0,01]	0,052
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,47 (0,43)	0,88 (0,64)	-0,41 [-1,84; 1,03]	0,574
BVMT-R Total Recall Score – Veränderung seit Studienbeginn				
N'	113	56		
Mittelwert zu Studienbeginn (SD)	20,33 (8,91)	18,18 (7,87)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,72 (0,74)	-0,21 (0,96)	-0,52 [-2,55; 1,52]	0,616
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,93 (0,99)	1,06 (1,47)	-0,13 [-3,43; 3,16]	0,937

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
BVMT-R: Brief Visuospatial Memory Test – Revised; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error. a: Positive Werte zeigen eine Verbesserung an. Quelle: Anhang 2 zu Modul 4, Tabellen 8.1.1 und 8.2.1.				

Die Auswertung des BVMT-R ergab keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen (Tabelle 4-72 und Tabelle 4-73).

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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 Schweregrad der Beeinträchtigung des Sehvermögens (LCVA) – RCT

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

Tabelle 4-74: Operationalisierung des Endpunkts *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)*

Studie	Operationalisierung
EXPAND	Der Schweregrad der Beeinträchtigung des Sehvermögens wurde anhand des LCVA (Low Contrast Visual Acuity) ermittelt. Der LCVA wurde zu Baseline und dann alle sechs Monate erhoben. Beim LCVA zeigt eine Zunahme der Werte eine Verbesserung an. Im Dossier berichtet wird: • die Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn für den LCVA (dargestellt werden die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den zugehörigen Standardabweichungen sowie die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht).

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 4-75: Bewertung des Verzerrungspotenzials für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkterheber (unabhängiger Bewerter) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und sonstige das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden. Der Nichtberücksichtigungsanteil liegt zu Monat 12 insgesamt unter 30 %,

der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen unter 10 %. Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden und für die Auswertung zu Monat 12 ist nur mit einem niedrigen Verzerrungspotenzial zu rechnen.

Tabelle 4-76: Rücklaufquoten LCVA (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
LCVA			
Rückläufer Studienbeginn, n (%)	1.046 (95,2)	526 (96,3)	1.572 (95,6)
Rückläufer Monat 6, n (%)	930 (84,6)	476 (87,2)	1.406 (85,5)
Rückläufer Monat 12, n (%)	824 (75,0)	407 (74,5)	1.231 (74,8)
Rückläufer Monat 18, n (%)	540 (49,1)	249 (45,6)	789 (48,0)
Rückläufer Monat 24, n (%)	327 (29,8)	141 (25,8)	468 (28,4)
Rückläufer Monat 30, n (%)	118 (10,7)	38 (7,0)	156 (9,5)

LCVA: Low Contrast Visual Acuity; N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 1 zu Modul 4, Tabelle 8.3.0.

Tabelle 4-77: Rücklaufquoten LCVA (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
LCVA			
Rückläufer Studienbeginn, n (%)	121 (95,3)	60 (98,4)	181 (96,3)
Rückläufer Monat 6, n (%)	111 (87,4)	56 (91,8)	167 (88,8)
Rückläufer Monat 12, n (%)	103 (81,1)	46 (75,4)	149 (79,3)
Rückläufer Monat 18, n (%)	70 (55,1)	23 (37,7)	93 (49,5)
Rückläufer Monat 24, n (%)	41 (32,3)	16 (26,2)	57 (30,3)
Rückläufer Monat 30, n (%)	15 (11,8)	6 (9,8)	21 (11,2)

LCVA: Low Contrast Visual Acuity; N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 2 zu Modul 4, Tabelle 8.3.0.

Stellen Sie die Ergebnisse für den Endpunkt Schweregrad der Beeinträchtigung des Sehvermögens (LCVA) für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Veränderung im LCVA seit Studienbeginn

Tabelle 4-78: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
LCVA – Veränderung seit Studienbeginn				
N'	940	486		
Mittelwert zu Studienbeginn (SD)	0,37 (0,26)	0,36 (0,26)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,02 (0,01)	0,01 (0,01)	0,01 [-0,01; 0,04]	0,366
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,01 (0,01)	0,00 (0,02)	0,00 [-0,03; 0,04]	0,877

KI: Konfidenzintervall; LCVA: Low Contrast Visual Acuity; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation ; SE: Standard Error .

a: Positive Werte zeigen eine Verbesserung an.

Quelle: Anhang 1 zu Modul 4, Tabelle 8.3.1.

Tabelle 4-79: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
LCVA – Veränderung seit Studienbeginn				
N'	110	55		
Mittelwert zu Studienbeginn (SD)	0,39 (0,26)	0,39 (0,26)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,00 (0,03)	-0,01 (0,03)	0,02 [-0,06; 0,09]	0,632
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,00 (0,03)	0,04 (0,05)	-0,04 [-0,15; 0,06]	0,401

KI: Konfidenzintervall; LCVA: Low Contrast Visual Acuity; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error .

a: Positive Werte zeigen eine Verbesserung an.

Quelle: Anhang 2 zu Modul 4, Tabelle 8.3.1.

Die Auswertung des LCVA ergab keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen (Tabelle 4-78 und Tabelle 4-79).

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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 Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12) – RCT

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

Tabelle 4-80: Operationalisierung des Endpunkts *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)*

Studie	Operationalisierung
EXPAND	Der Schweregrad der Beeinträchtigung der Gehfähigkeit wurde anhand des MSWS-12 (Multiple Sclerosis Walking Scale) ermittelt. Der Fragebogen zu MSWS-12 war vom Patienten selbständig zu Baseline und danach alle sechs Monate jeweils vor den anderen zu dieser Visite stattfindenden Untersuchungen auszufüllen. Beim MSWS-12 zeigt eine Zunahme der Werte eine Verschlechterung an. Im Dossier berichtet wird: • die Veränderung gegenüber Studienbeginn zu jedem Erfassungszeitpunkt für den MSWS-12 Converted Score (jeweils dargestellt werden die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den zugehörigen Standardabweichungen sowie für die Monate 12 und 24 die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht).

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 4-81: Bewertung des Verzerrungspotenzials für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkterheber (Patient) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und

sonstige das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden. Der Nichtberücksichtigungsanteil liegt zu Monat 12 insgesamt unter 30 %, der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen unter 10 %. Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden und für die Auswertung zu Monat 12 ist nur mit einem niedrigen Verzerrungspotenzial zu rechnen.

Tabelle 4-82: Rücklaufquoten MSWS-12 (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
MSWS-12			
Rückläufer Studienbeginn, n (%)	1.089 (99,1)	539 (98,7)	1.628 (99,0)
Rückläufer Monat 6, n (%)	1.012 (92,1)	518 (94,9)	1.530 (93,0)
Rückläufer Monat 12, n (%)	925 (84,2)	452 (82,8)	1.377 (83,7)
Rückläufer Monat 18, n (%)	516 (47,0)	246 (45,1)	762 (46,3)
Rückläufer Monat 24, n (%)	403 (36,7)	195 (35,7)	598 (36,4)
Rückläufer Monat 30, n (%)	144 (13,1)	59 (10,8)	203 (12,3)

MSWS: Multiple Sclerosis Walking Scale; N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 1 zu Modul 4, Tabelle 10.5.0.

Tabelle 4-83: Rücklaufquoten MSWS-12 (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
MSWS-12			
Rückläufer Studienbeginn, n (%)	127 (100,0)	60 (98,4)	187 (99,5)
Rückläufer Monat 6, n (%)	118 (92,9)	58 (95,1)	176 (93,6)
Rückläufer Monat 12, n (%)	107 (84,3)	47 (77,0)	154 (81,9)
Rückläufer Monat 18, n (%)	66 (52,0)	26 (42,6)	92 (48,9)
Rückläufer Monat 24, n (%)	51 (40,2)	20 (32,8)	71 (37,8)
Rückläufer Monat 30, n (%)	19 (15,0)	9 (14,8)	28 (14,9)

MSWS: Multiple Sclerosis Walking Scale; N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 2 zu Modul 4, Tabelle 10.5.0.

Stellen Sie die Ergebnisse für den Endpunkt Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12) für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-84: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
MSWS-12 – Veränderung seit Studienbeginn				
N'	1.022	516		
Mittelwert zu Studienbeginn (SD)	68,34 (23,39)	67,39 (22,28)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	0,11 (0,64)	1,09 (0,84)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	1,53 (0,68)	3,36 (0,91)	-1,83 [-3,85; 0,19]	0,076
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	3,28 (0,74)	5,01 (1,01)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	4,16 (0,85)	5,38 (1,17)	-1,23 [-3,89; 1,44]	0,367
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	4,39 (1,10)	7,47 (1,63)		

KI: Konfidenzintervall; MSWS: Multiple Sclerosis Walking Scale; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Negative Werte zeigen eine Verbesserung an.

Quelle: Anhang 1 zu Modul 4, Tabelle 10.5.1.

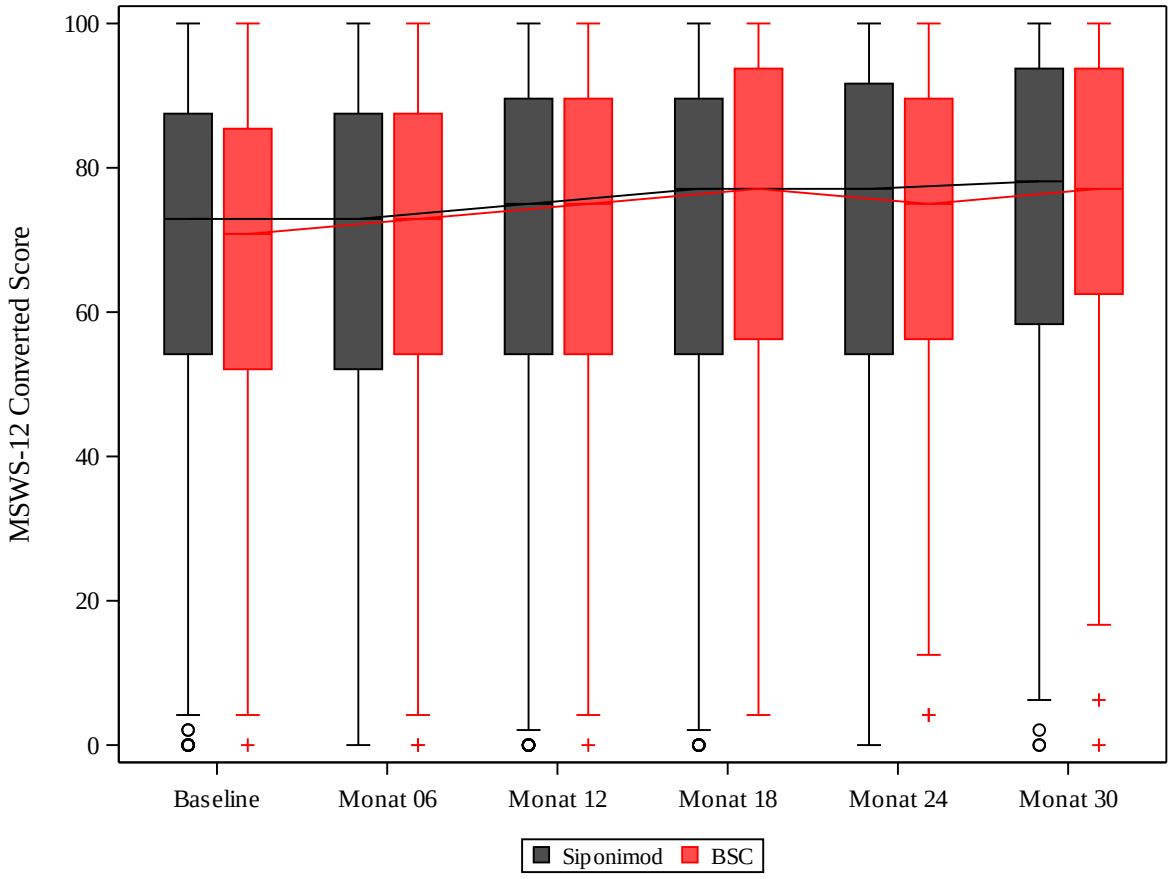


Abbildung 10: Zeitverlauf für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Tabelle 4-85: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
MSWS-12 – Veränderung seit Studienbeginn				
N'	118	57		
Mittelwert zu Studienbeginn (SD)	70,06 (24,52)	68,25 (23,57)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	1,98 (2,31)	1,02 (2,99)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	4,46 (2,32)	4,72 (3,08)	-0,26 [-6,89; 6,36]	0,938
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	8,73 (2,30)	6,71 (3,09)		

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	9,99 (2,64)	-0,14 (3,70)	10,13 [1,96; 18,29]	0,016
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	7,90 (3,06)	8,98 (4,04)		

KI: Konfidenzintervall; MSWS: Multiple Sclerosis Walking Scale; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation;SE: Standard Error.

a: Negative Werte zeigen eine Verbesserung an.

Quelle: Anhang 2 zu Modul 4, Tabelle 10.5.1.

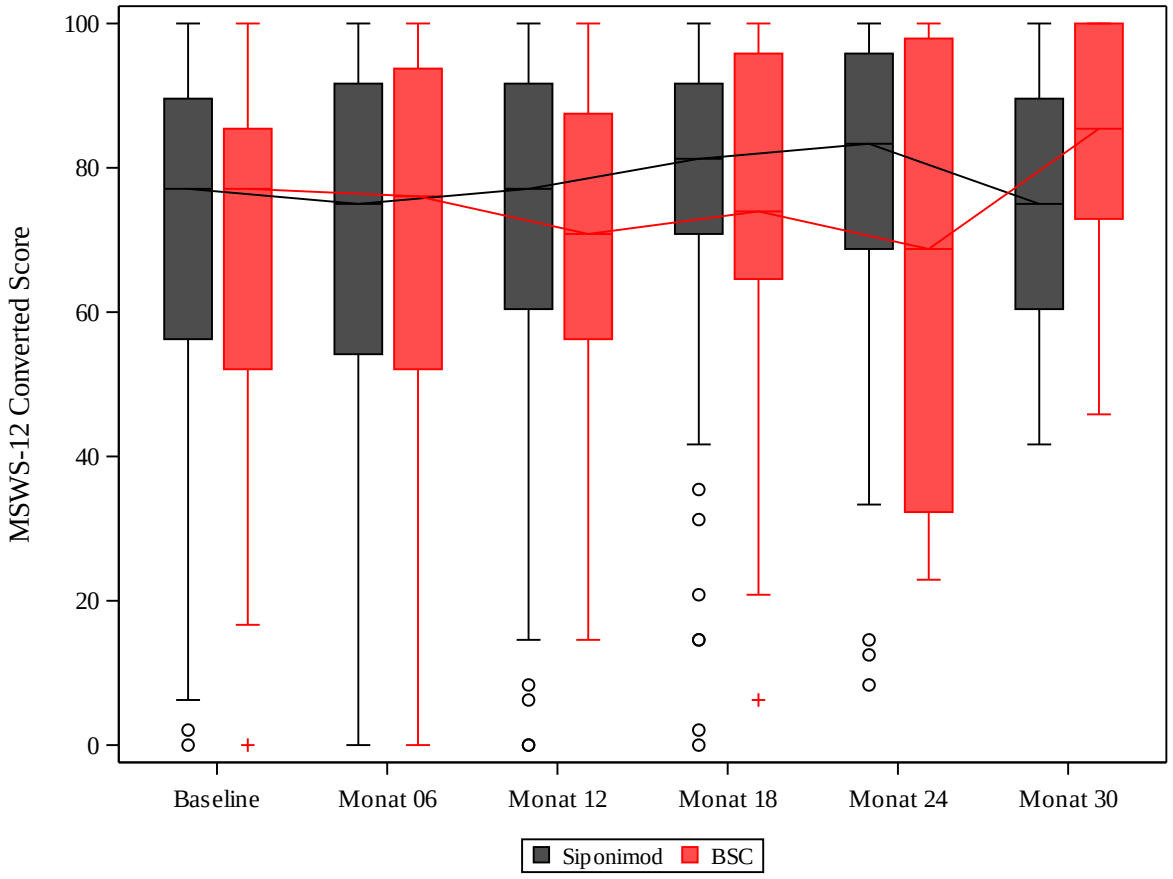


Abbildung 11: Zeitverlauf für den Endpunkt *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

Die Auswertungen des MSWS-12 zeigten lediglich in der Subpopulation B einen statistisch signifikanten Vorteil von BSC zu Monat 24. Die Auswertung zu Monat 24 ist wegen hoher Nichtberücksichtigungsanteile jedoch nicht zuverlässig interpretierbar (Tabelle 4-84 und Tabelle 4-85).

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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 Gesundheitsbezogene Lebensqualität (MSIS-29) – RCT

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

Tabelle 4-86: Operationalisierung des Endpunkts *Gesundheitsbezogene Lebensqualität (MSIS-29)*

Studie	Operationalisierung
EXPAND	Der Endpunkt Gesundheitsbezogene Lebensqualität wurde mittels der Multiple Sclerosis Impact Scale (MSIS-29) ermittelt. Die Fragebögen waren vom Patienten selbständig zu Baseline und danach alle sechs Monate auszufüllen, jeweils vor den anderen zu dieser Visite stattfindenden Untersuchungen. Beim MSIS-29 zeigt eine Zunahme der Werte eine Verschlechterung an. Im Dossier berichtet wird: • die Veränderung gegenüber Studienbeginn zu jedem Erhebungszeitpunkt für die MSIS-29 Physical Impact Scale und die MSIS-29 Psychological Impact Scale (jeweils dargestellt werden die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den zugehörigen Standardabweichungen sowie für die Monate 12 und 24 die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht).

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 4-87: Bewertung des Verzerrungspotenzials für *Gesundheitsbezogene Lebensqualität (MSIS-29)* 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
EXPAND	niedrig	ja	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* als hoch bewertet. Es war auf Studienebene niedrig, der Endpunkterheber (Patient) war verblindet, eine ergebnisgesteuerte Berichterstattung war unwahrscheinlich und sonstige das Verzerrungspotenzial beeinflussende Punkte wurden nicht festgestellt. Der Nichtberücksichtigungsanteil liegt zu Monat 12 in der Gesamtpopulation und

der Subpopulation B insgesamt jeweils unter 30 %, Der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen liegt in der Gesamtpopulation unter 10 % und in der Subpopulation B bei über 10 %. Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden. Für die Auswertung zu Monat 12 ist in der Gesamtpopulation nur mit einem niedrigen Verzerrungspotenzial und in der Subpopulation B mit einem hohen Verzerrungspotenzial zu rechnen.

Tabelle 4-88: Rücklaufquoten MSIS-29 (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
MSIS-29 Physical Impact Scale			
Rückläufer Studienbeginn, n (%)	1.066 (97,0)	530 (97,1)	1.596 (97,0)
Rückläufer Monat 6, n (%)	992 (90,3)	509 (93,2)	1.501 (91,2)
Rückläufer Monat 12, n (%)	909 (82,7)	441 (80,8)	1.350 (82,1)
Rückläufer Monat 18, n (%)	634 (57,7)	306 (56,0)	940 (57,1)
Rückläufer Monat 24, n (%)	388 (35,3)	188 (34,4)	576 (35,0)
Rückläufer Monat 30, n (%)	146 (13,3)	61 (11,2)	207 (12,6)
MSIS-29 Psychological Impact Scale			
Rückläufer Studienbeginn, n (%)	1.064 (96,8)	529 (96,9)	1.593 (96,8)
Rückläufer Monat 6, n (%)	986 (89,7)	507 (92,9)	1.493 (90,8)
Rückläufer Monat 12, n (%)	909 (82,7)	441 (80,8)	1.350 (82,1)
Rückläufer Monat 18, n (%)	632 (57,5)	306 (56,0)	938 (57,0)
Rückläufer Monat 24, n (%)	386 (35,1)	188 (34,4)	574 (34,9)
Rückläufer Monat 30, n (%)	146 (13,3)	61 (11,2)	207 (12,6)

MSIS: Multiple Sclerosis Impact Scale; N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 1 zu Modul 4, Tabellen 10.1.0 und 10.2.0.

Tabelle 4-89: Rücklaufquoten MSIS-29 (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
MSIS-29 Physical Impact Scale			
Rückläufer Studienbeginn, n (%)	126 (99,2)	59 (96,7)	185 (98,4)
Rückläufer Monat 6, n (%)	117 (92,1)	56 (91,8)	173 (92,0)
Rückläufer Monat 12, n (%)	107 (84,3)	44 (72,1)	151 (80,3)
Rückläufer Monat 18, n (%)	81 (63,8)	29 (47,5)	110 (58,5)
Rückläufer Monat 24, n (%)	50 (39,4)	19 (31,1)	69 (36,7)
Rückläufer Monat 30, n (%)	20 (15,7)	9 (14,8)	29 (15,4)
MSIS-29 Psychological Impact Scale			
Rückläufer Studienbeginn, n (%)	126 (99,2)	59 (96,7)	185 (98,4)
Rückläufer Monat 6, n (%)	116 (91,3)	55 (90,2)	171 (91,0)
Rückläufer Monat 12, n (%)	107 (84,3)	44 (72,1)	151 (80,3)
Rückläufer Monat 18, n (%)	81 (63,8)	29 (47,5)	110 (58,5)
Rückläufer Monat 24, n (%)	50 (39,4)	19 (31,1)	69 (36,7)
Rückläufer Monat 30, n (%)	20 (15,7)	9 (14,8)	29 (15,4)

MSIS: Multiple Sclerosis Impact Scale; N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 2 zu Modul 4, Tabellen 10.1.0 und 10.2.0.

Stellen Sie die Ergebnisse für den Endpunkt Gesundheitsbezogene Lebensqualität (MSIS-29) für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-90: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
MSIS-29 Physical Impact Scale – Veränderung seit Studienbeginn				
N'	1.001	508		
Mittelwert zu Studienbeginn (SD)	51,20 (21,54)	50,53 (21,35)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	0,24 (0,61)	1,55 (0,80)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	1,33 (0,65)	4,22 (0,87)	-2,89 [-4,83; -0,96]	0,003
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	2,79 (0,70)	5,09 (0,95)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	3,31 (0,82)	4,67 (1,12)	-1,36 [-3,93; 1,21]	0,300
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	3,79 (1,07)	6,37 (1,57)		
MSIS-29 Psychological Impact Scale – Veränderung seit Studienbeginn				
N'	998	506		
Mittelwert zu Studienbeginn (SD)	38,05 (23,32)	38,46 (23,87)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-0,73 (0,66)	0,95 (0,86)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,11 (0,71)	2,14 (0,95)	-2,03 [-4,15; 0,09]	0,060
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	0,88 (0,79)	2,48 (1,08)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,81 (0,96)	1,48 (1,33)	-0,66 [-3,73; 2,40]	0,670
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-1,05 (1,20)	2,82 (1,78)		

KI: Konfidenzintervall; MSIS: Multiple Sclerosis Impact Scale; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation ; SE: Standard Error .

a: Negative Werte zeigen eine Verbesserung an.

Quelle: Anhang 1 zu Modul 4, Tabellen 10.1.1 und 10.2.1.

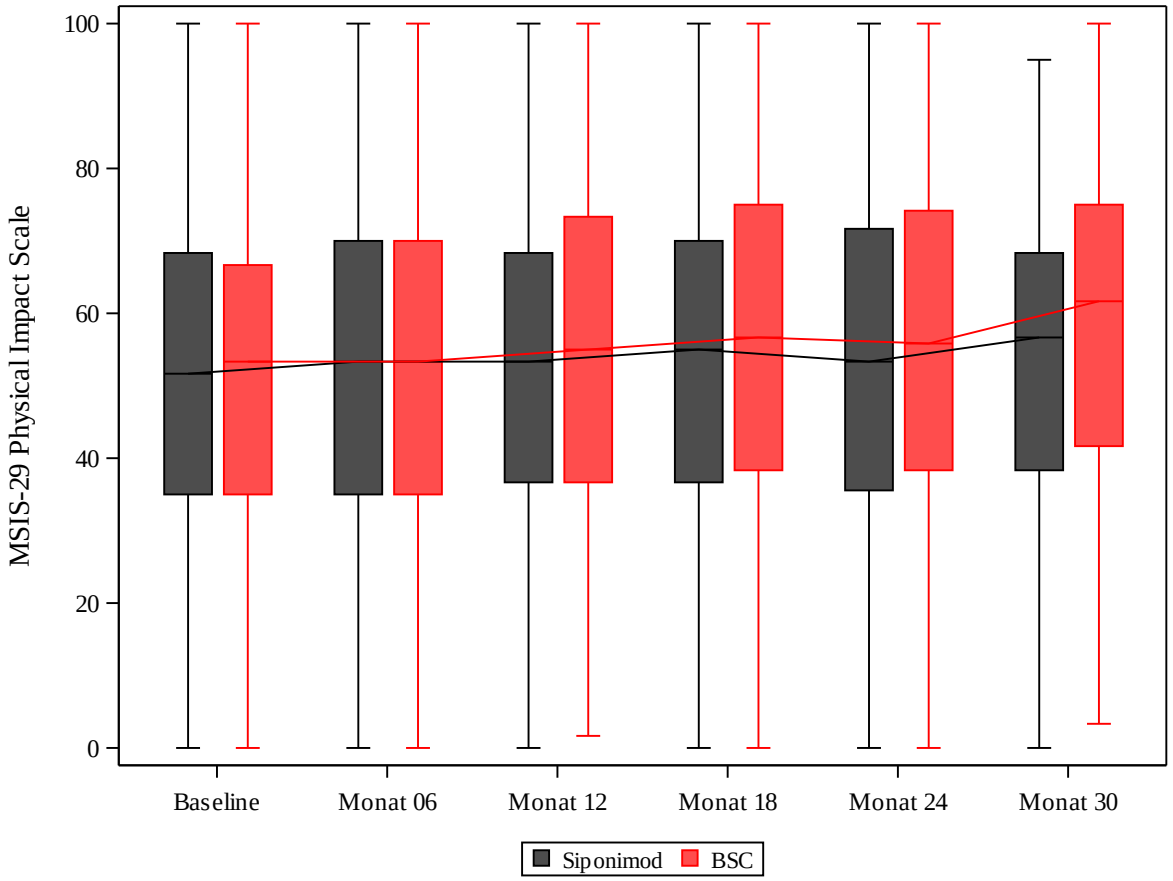


Abbildung 12: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

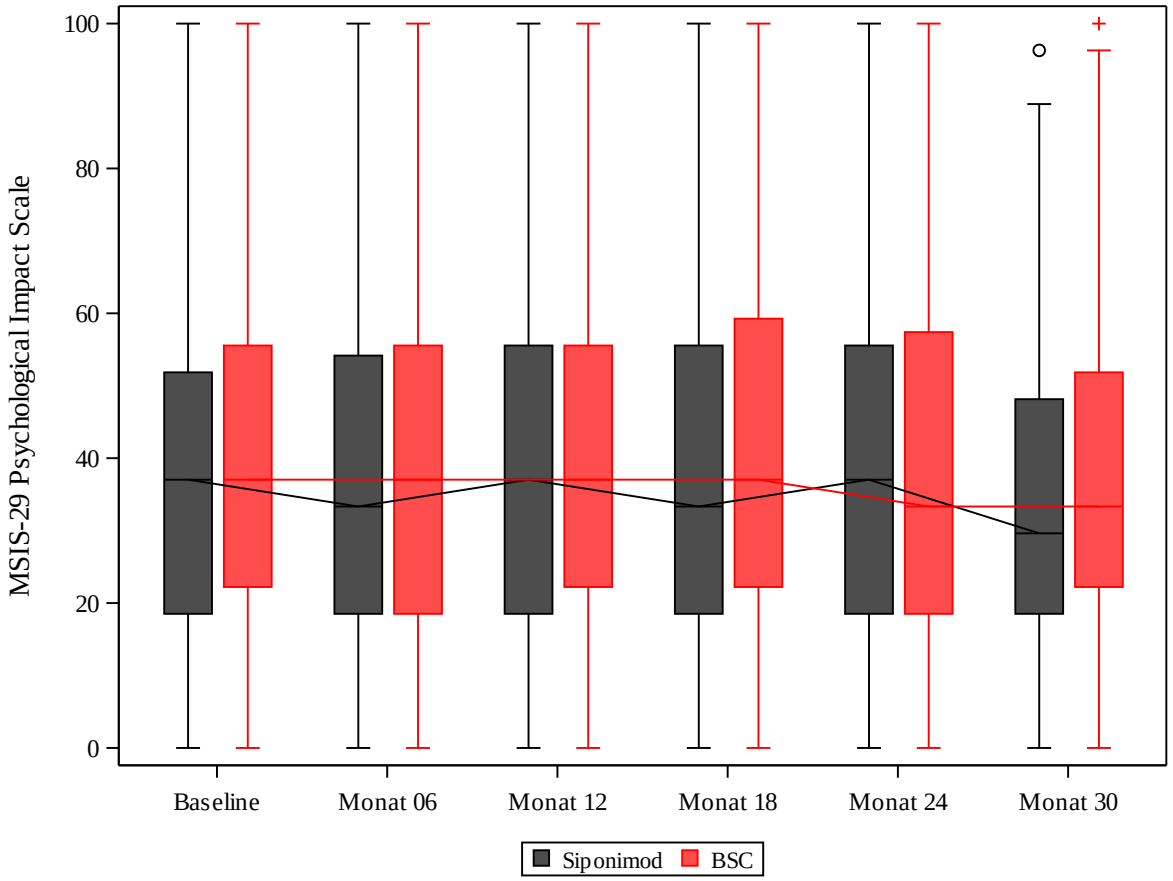


Abbildung 13: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Tabelle 4-91: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
MSIS-29 Physical Impact Scale – Veränderung seit Studienbeginn				
N'	117	56		
Mittelwert zu Studienbeginn (SD)	52,15 (21,74)	53,61 (23,26)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	3,23 (2,32)	0,77 (2,98)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	1,53 (2,28)	0,41 (3,00)	1,12 [-5,22; 7,46]	0,727
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	7,29 (2,32)	4,65 (3,17)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	7,25 (2,60)	-0,92 (3,65)	8,16 [0,17; 16,16]	0,045
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	8,16 (3,29)	2,62 (4,67)		
MSIS-29 Psychological Impact Scale – Veränderung seit Studienbeginn				
N'	117	55		
Mittelwert zu Studienbeginn (SD)	34,73 (21,90)	41,68 (24,26)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	2,22 (2,44)	-0,36 (3,20)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	3,33 (2,39)	1,35 (3,21)	1,98 [-4,77; 8,73]	0,563
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	5,85 (2,53)	4,69 (3,61)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	7,51 (2,93)	-3,24 (4,25)	10,75 [1,34; 20,16]	0,026
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	7,71 (3,77)	-2,13 (5,34)		

KI: Konfidenzintervall; MSIS: Multiple Sclerosis Impact Scale; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Negative Werte zeigen eine Verbesserung an.

Quelle: Anhang 2 zu Modul 4, Tabellen 10.1.1 und 10.2.1.

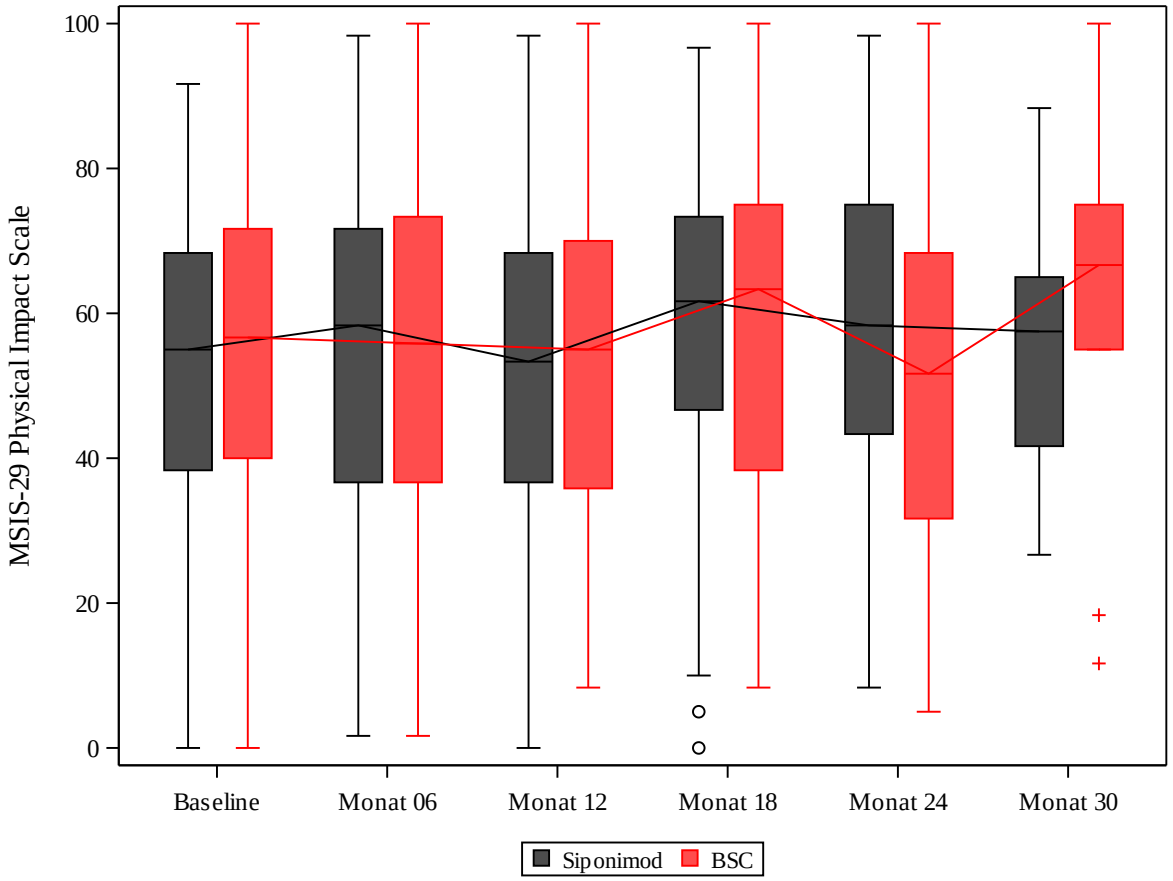


Abbildung 14: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

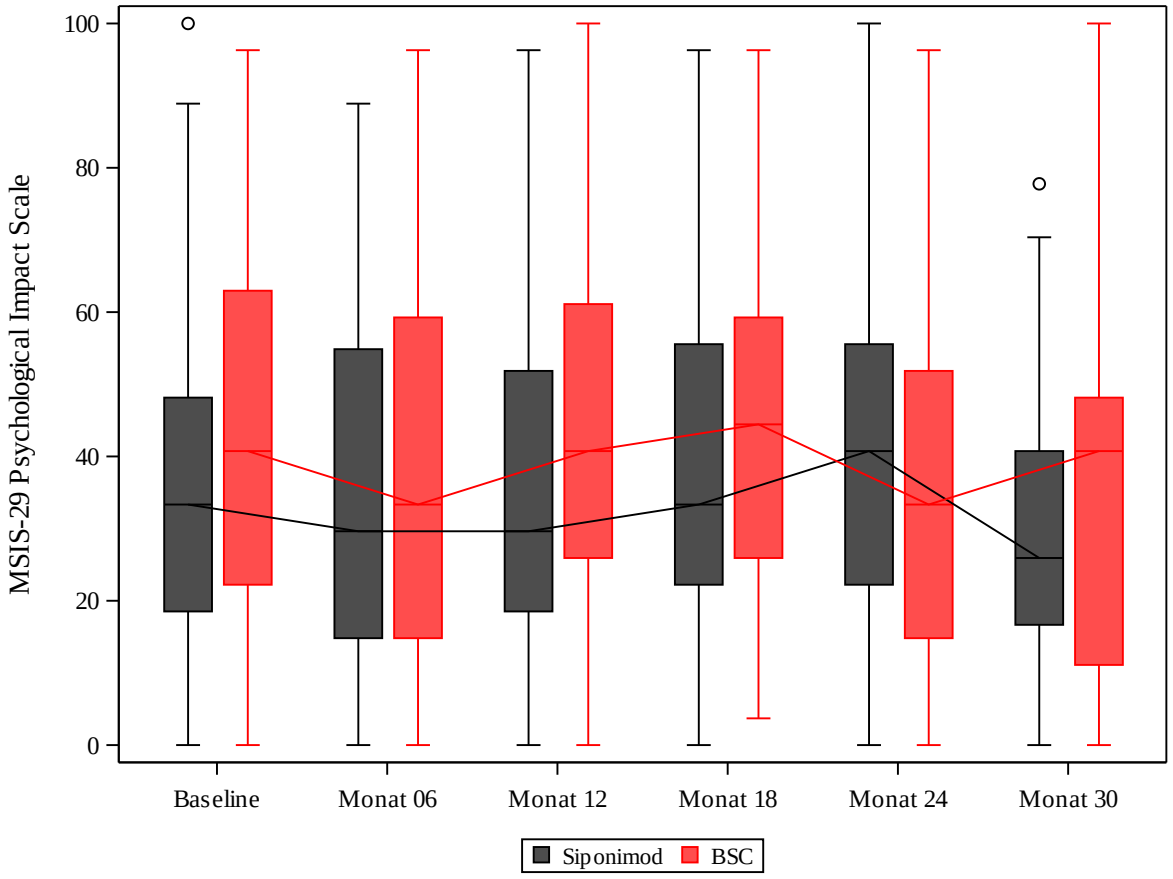


Abbildung 15: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

Die Auswertungen der MSIS-29 in der Gesamtpopulation zeigten einen statistisch signifikanten Vorteil von Siponimod in der Physical Impact Scale zu Monat 12. Unter Siponimod stieg dabei die Beeinträchtigung durch die Erkrankung weniger stark an als unter BSC (Tabelle 4-90).

Die Auswertungen der MSIS-29 in der Subpopulation B zeigten einen statistisch signifikanten Vorteil von BSC sowohl in der Physical Impact Scale als auch in der Psychological Impact Scale zu Monat 24. Unter BSC verringerte sich dabei die Beeinträchtigung durch die Erkrankung, wohingegen sie unter Siponimod stieg. Die Auswertung zu Monat 24 ist wegen hoher Nichtberücksichtigungsanteile jedoch nicht zuverlässig interpretierbar (Tabelle 4-91).

Es ergaben sich keine weiteren statistisch signifikanten Unterschiede in der Auswertung der MSIS-29.

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

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Auf Endpunktebene liegt ein hohes Verzerrungspotenzial vor, die die Übertragbarkeit ggf. einschränken.

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 Gesundheitsbezogene Lebensqualität (EQ-5D) – RCT

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

Tabelle 4-92: Operationalisierung des Endpunkts *Gesundheitsbezogene Lebensqualität (EQ-5D)*

Studie	Operationalisierung
EXPAND	Der Endpunkt Gesundheitsbezogene Lebensqualität wurde mittels des Fragebogen EQ-5D erhoben. Die Fragebögen waren vom Patienten selbständig zu Baseline und danach alle sechs Monate auszufüllen, jeweils vor den anderen zu dieser Visite stattfindenden Untersuchungen. Im Dossier berichtet wird: • die Veränderung gegenüber Studienbeginn zu jedem Erfassungszeitpunkt für den EQ-5D-Indexwert (basierend auf der Health State Classification) und die EQ-5D-VAS (jeweils dargestellt werden die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den zugehörigen Standardabweichungen sowie für die Monate 12 und 24 die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht).

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 4-93: Bewertung des Verzerrungspotenzials für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkterheber (Patient) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und sonstige das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden. Der Nichtberücksichtigungsanteil liegt zu Monat 12 insgesamt unter 30 %, der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen unter 10 %. Zu Monat 24 liegen die Nichtberücksichtigungsanteile jedoch bei über 30 %.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden und für die Auswertung zu Monat 12 ist prinzipiell nur mit einem niedrigen Verzerrungspotenzial zu rechnen.

Tabelle 4-94: Rücklaufquoten EQ-5D-Komponenten (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
EQ-5D Health State Classification			
Rückläufer Studienbeginn, n (%)	1.084 (98,6)	536 (98,2)	1.620 (98,5)
Rückläufer Monat 6, n (%)	1.012 (92,1)	517 (94,7)	1.529 (92,9)
Rückläufer Monat 12, n (%)	910 (82,8)	438 (80,2)	1.348 (81,9)
Rückläufer Monat 18, n (%)	646 (58,8)	308 (56,4)	954 (58,0)
Rückläufer Monat 24, n (%)	305 (27,8)	143 (26,2)	448 (27,2)
Rückläufer Monat 30, n (%)	87 (7,9)	35 (6,4)	122 (7,4)
EQ-5D-VAS			
Rückläufer Studienbeginn, n (%)	1.082 (98,5)	537 (98,4)	1.619 (98,4)
Rückläufer Monat 6, n (%)	1.004 (91,4)	512 (93,8)	1.516 (92,2)
Rückläufer Monat 12, n (%)	907 (82,5)	439 (80,4)	1.346 (81,8)
Rückläufer Monat 18, n (%)	643 (58,5)	308 (56,4)	951 (57,8)
Rückläufer Monat 24, n (%)	306 (27,8)	143 (26,2)	449 (27,3)
Rückläufer Monat 30, n (%)	87 (7,9)	36 (6,6)	123 (7,5)

EQ-5D: EuroQol – 5 Dimensionen, N: Patienten gesamt; n: Patienten mit Ereignis, VAS: Visuelle Analogskala.

Quelle: Anhang 1 zu Modul 4, Tabellen 10.3.0 und 10.4.0.

Tabelle 4-95: Rücklaufquoten EQ-5D-Komponenten (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
EQ-5D Health State Classification			
Rückläufer Studienbeginn, n (%)	127 (100,0)	61 (100,0)	188 (100,0)
Rückläufer Monat 6, n (%)	118 (92,9)	57 (93,4)	175 (93,1)
Rückläufer Monat 12, n (%)	107 (84,3)	47 (77,0)	154 (81,9)
Rückläufer Monat 18, n (%)	81 (63,8)	31 (50,8)	112 (59,6)
Rückläufer Monat 24, n (%)	40 (31,5)	16 (26,2)	56 (29,8)
Rückläufer Monat 30, n (%)	14 (11,0)	3 (4,9)	17 (9,0)
EQ-5D-VAS			
Rückläufer Studienbeginn, n (%)	125 (98,4)	61 (100,0)	186 (98,9)
Rückläufer Monat 6, n (%)	117 (92,1)	57 (93,4)	174 (92,6)
Rückläufer Monat 12, n (%)	106 (83,5)	47 (77,0)	153 (81,4)
Rückläufer Monat 18, n (%)	82 (64,6)	31 (50,8)	113 (60,1)
Rückläufer Monat 24, n (%)	40 (31,5)	16 (26,2)	56 (29,8)
Rückläufer Monat 30, n (%)	14 (11,0)	3 (4,9)	17 (9,0)

EQ-5D: EuroQol – 5 Dimensionen; N: Patienten gesamt; n: Patienten mit Ereignis; VAS: Visuelle Analogskala.

Quelle: Anhang 2 zu Modul 4, Tabellen 10.3.0 und 10.4.0.

Stellen Sie die Ergebnisse für den Endpunkt Gesundheitsbezogene Lebensqualität (MSIS-29) für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-96: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
EQ-5D-Indexwert – Veränderung seit Studienbeginn				
N'	1.017	514		
Mittelwert zu Studienbeginn (SD)	0,61 (0,23)	0,59 (0,23)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-0,02 (0,01)	-0,02 (0,01)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,02 (0,01)	-0,05 (0,01)	0,03 [0,00; 0,05]	0,037
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-0,04 (0,01)	-0,04 (0,01)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,04 (0,01)	-0,07 (0,02)	0,03 [0,00; 0,07]	0,088
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-0,01 (0,02)	-0,11 (0,03)		
EQ-5D-VAS – Veränderung seit Studienbeginn				
N'	1.016	514		
Mittelwert zu Studienbeginn (SD)	57,97 (18,93)	56,28 (19,25)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-1,03 (0,53)	-0,91 (0,74)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-2,08 (0,57)	-3,88 (0,82)	1,79 [-0,16; 3,75]	0,072
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-2,17 (0,67)	-3,27 (0,96)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-2,06 (0,90)	-3,00 (1,31)	0,94 [-2,18; 4,07]	0,555
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-3,70 (1,74)	-5,90 (2,67)		

EQ-5D: EuroQol – 5 Dimensionen; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error ; VAS: Visuelle Analogskala.

a: Positive Werte zeigen eine Verbesserung an.

Quelle: Anhang 1 zu Modul 4, Tabellen 10.3.1 und 10.4.1.

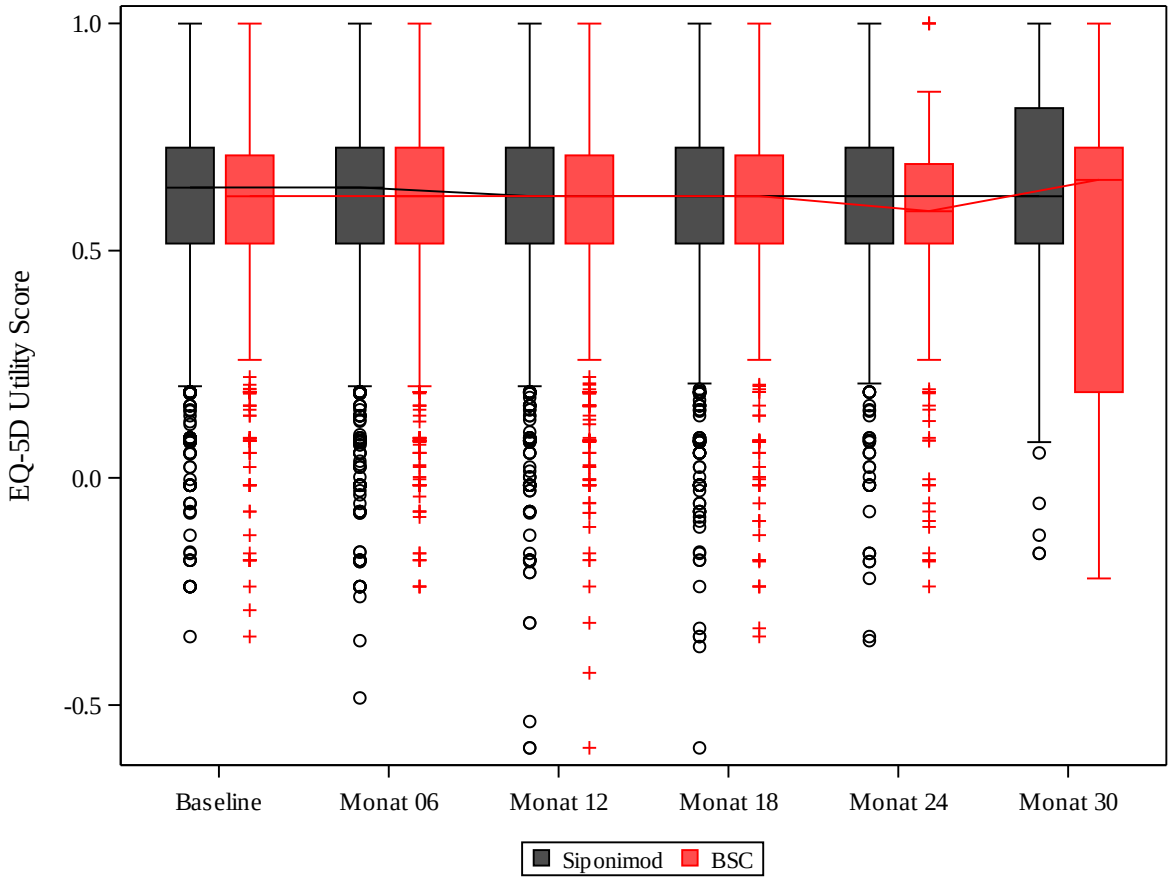


Abbildung 16: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-Indexwert)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

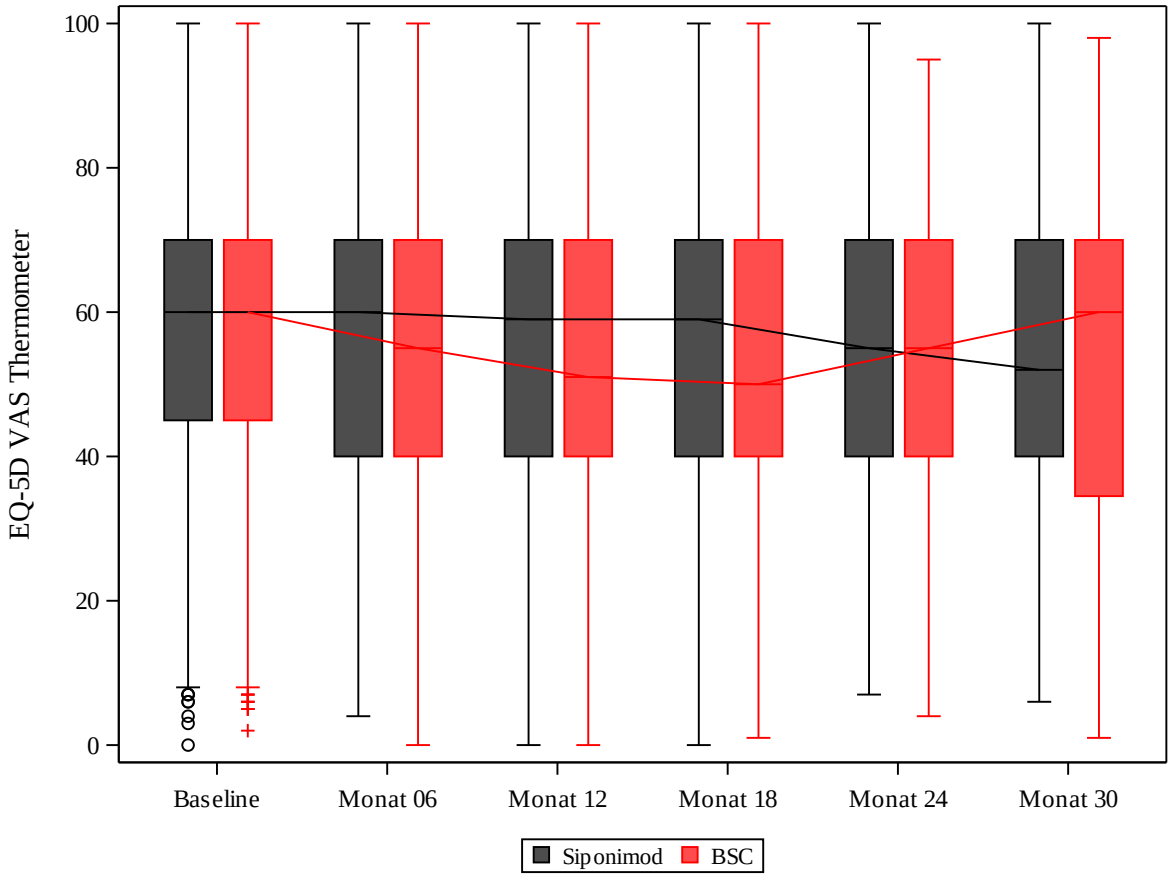


Abbildung 17: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-VAS)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Tabelle 4-97: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
EQ-5D-Indexwert – Veränderung seit Studienbeginn				
N'	118	58		
Mittelwert zu Studienbeginn (SD)	0,62 (0,24)	0,59 (0,27)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-0,03 (0,02)	-0,03 (0,03)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,01 (0,02)	-0,06 (0,03)	0,04 [-0,02; 0,11]	0,188
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-0,04 (0,02)	-0,06 (0,04)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,05 (0,03)	-0,10 (0,05)	0,05 [-0,07; 0,17]	0,413
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-0,04 (0,03)	-0,03 (0,08)		
EQ-5D-VAS – Veränderung seit Studienbeginn				
N'	117	58		
Mittelwert zu Studienbeginn (SD)	58,76 (18,96)	56,52 (20,22)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-1,93 (1,67)	1,96 (2,38)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-2,27 (1,68)	-0,16 (2,48)	-2,11 [-8,02; 3,80]	0,482
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-4,27 (1,85)	-1,29 (2,93)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-6,01 (2,27)	-1,68 (3,49)	-4,33 [-12,61; 3,96]	0,302
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-1,16 (3,65)	0,14 (7,81)		

EQ-5D: EuroQol – 5 Dimensionen; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation [*Standardabweichung*]; SE: Standard Error; VAS: Visuelle Analogskala.

a: Positive Werte zeigen eine Verbesserung an.

Quelle: Anhang 2 zu Modul 4, Tabellen 10.3.1 und 10.4.1.

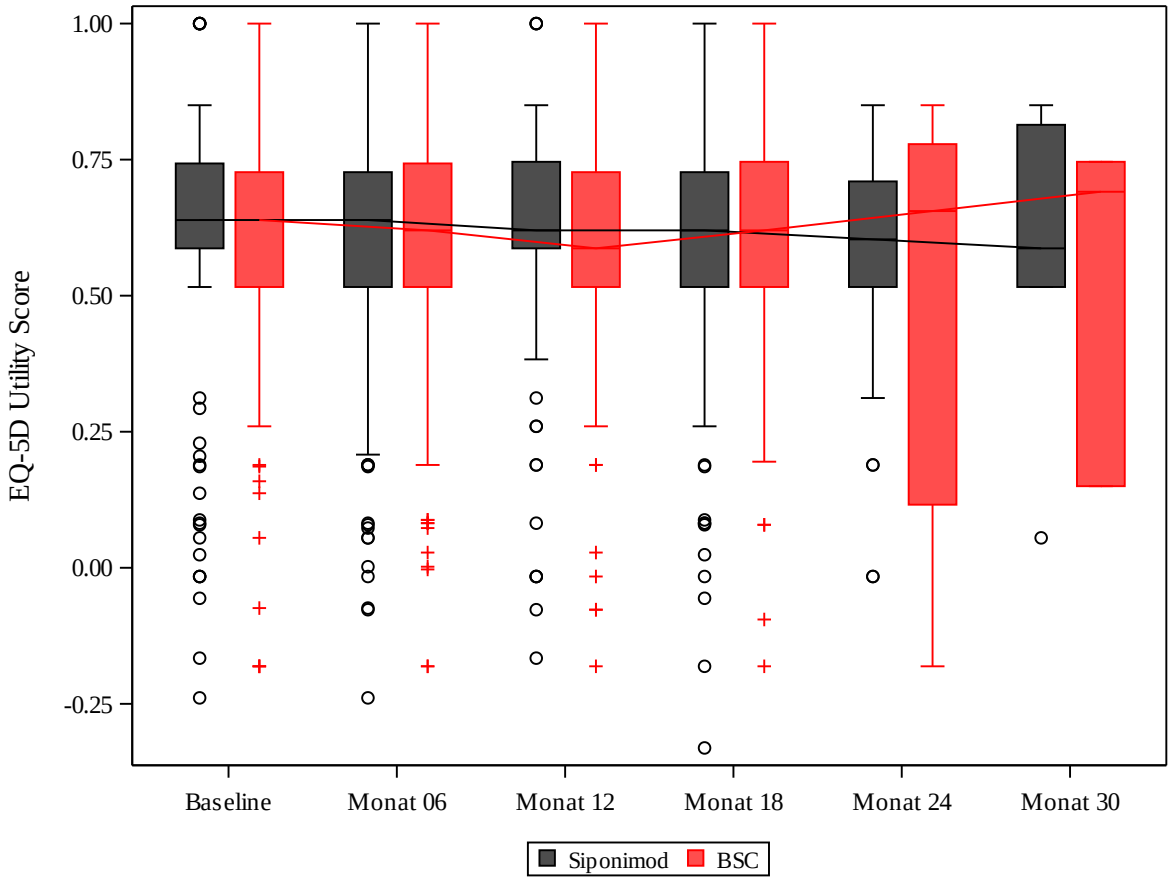


Abbildung 18: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-Indexwert)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

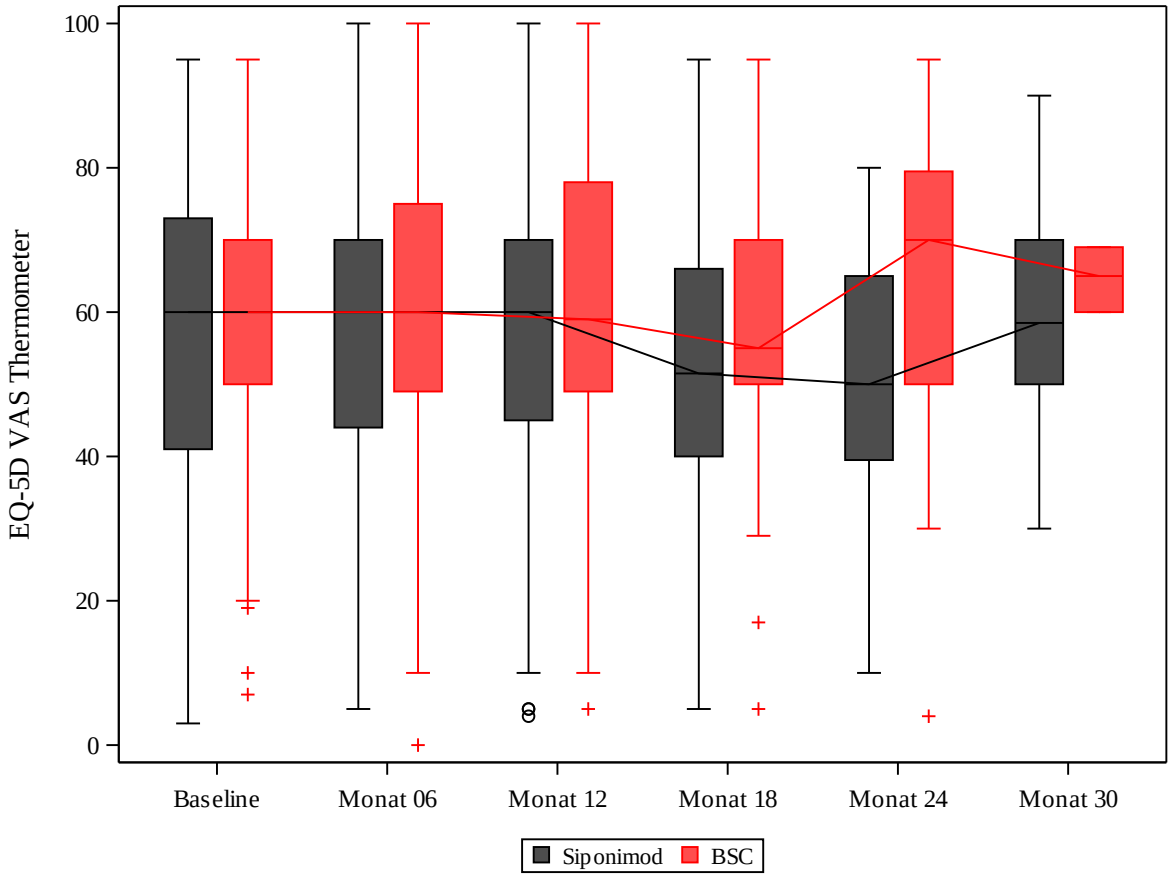


Abbildung 19: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-VAS)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

Die Auswertungen des EQ-5D zeigten in der Gesamtpopulation einen statistisch signifikanten Vorteil von Siponimod, und zwar im Indexwert zu Monat 12 (Mittelwert-Differenz 0,03; 95 %-KI [0,00; 0,05]). Unter BSC verschlechterte sich hierbei der Gesundheitszustand stärker als unter Siponimod. Der Gruppenunterschied blieb dabei unter der klinischen Relevanzschwelle von 0,074 Punkten (Tabelle 4-96 und Tabelle 4-97).

In der Supopulation B wurden keine statistisch signifikanten Unterschiede zwischen den Behandlungsgruppen festgestellt.

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

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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.

4.3.1.3.11 Unerwünschte Ereignisse – RCT

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

Tabelle 4-98: Operationalisierung des Endpunkts *Unerwünschte Ereignisse*

Studie	Operationalisierung
EXPAND	Die Untersuchungen zur Sicherheit bestanden aus dem Erfassen aller Unerwünschten (UE) und aller Schwerwiegenden Unerwünschten Ereignisse (SUE) sowie von Schwangerschaften. Das Auftreten von (S)UE wurde bei jeder Visite mittels nicht-direktiver Befragung des Patienten durch den Arzt ermittelt. Zudem konnten (S)UE auch jederzeit durch den Patienten selbst spontan berichtet oder anhand von Untersuchungen wie Labormessungen nach definierten Kriterien festgestellt werden. Laborparameter (hämatologisch, biochemisch und Urinanalyse) wurden regelmäßig durch ein Zentrallabor überwacht. Vitalzeichen, Körpergewicht und körperliche Verfassung des Patienten wurden regelmäßig untersucht. Im Prüfplan waren Kriterien für klinisch bemerkenswerte abnormale Werte definiert. (S)UE wurden nach MedDRA kodiert. Im Dossier berichtet werden: • Inzidenz von Unerwünschten Ereignissen (UE) insgesamt und unter Ausschluss krankheitsspezifischer UE. • Inzidenz von Schwerwiegenden Unerwünschten Ereignissen (SUE) insgesamt und unter Ausschluss krankheitsspezifischer UE. • Inzidenz von schweren UE (CTCAE-Grad ≥ 3) insgesamt und unter Ausschluss krankheitsspezifischer UE. • Inzidenz von Studienabbrüchen wegen UE. • Inzidenz von Behandlungsabbrüchen wegen UE. • Inzidenz von UE nach Schweregrad (CTCAE-Grade 1 bis 4). • Inzidenz von UE nach System Organ Class (SOC) und Preferred Term (PT): o nur SOC/PT mit einer Häufigkeit von ≥ 10 Patienten und ≥ 1 % in mindestens einem Behandlungsarm und

- nur für UE mit einer Häufigkeit von ≥ 10 % in mindestens einem Behandlungsarm.

- Inzidenz von SUE nach System Organ Class (SOC) und Preferred Term (PT):
  - nur SOC/PT mit einer Häufigkeit von ≥ 10 Patienten und ≥ 1 % in mindestens einem Behandlungsarm und
  - nur für SUE mit einer Häufigkeit von ≥ 5 % in mindestens einem Behandlungsarm.
- Inzidenz von UE nach SOC und PT und Schweregrad (nicht schwer = CTCAE-Grade 1 oder 2; schwer = CTCAE-Grad ≥ 3):
  - nur SOC/PT für Ereignisse mit einer Häufigkeit von ≥ 10 Patienten und ≥ 1 % in mindestens einem Behandlungsarm und
  - nur für UE mit einer Häufigkeit von ≥ 5 % in mindestens einem Behandlungsarm.
- Inzidenz von Studienabbrüchen wegen UE nach SOC und PT.
- Inzidenz von Behandlungsabbrüchen wegen UE nach SOC und PT.
- Inzidenz von UE von besonderem Interesse nach Schweregrad bzw. Einstufung als SUE (nicht schwer = CTCAE-Grade 1 oder 2; schwer = CTCAE-Grad ≥3; schwerwiegend = klassifiziert als SUE). Die Selektion der UE von besonderem Interesse erfolgte gemäß Studienbericht (Tabellen 14.3.1-4.1 und -4.2).
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für UE insgesamt und unter Ausschluss krankheitsspezifischer UE.
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für SUE insgesamt und unter Ausschluss krankheitsspezifischer UE.
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für Schwere UE (CTCAE-Grad ≥ 3) und unter Ausschluss krankheitsspezifischer UE.
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für Studienabbrüche wegen UE.
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für Behandlungsabbrüche wegen UE.
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für UE nach Schweregrad (Kategorien: CTCAE-Grade 1 bis 4; nicht schwer = CTCAE-Grade 1 oder 2; schwer = CTCAE-Grad ≥ 3).
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für UE nach SOC und PT:
  - nur SOC/PT für Ereignisse mit einer Häufigkeit von ≥ 10 Patienten und ≥ 1 % in mindestens einem Behandlungsarm und
  - nur für UE mit einer Häufigkeit von ≥ 10 % in mindestens einem Behandlungsarm.
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für SUE nach SOC und PT:
  - nur SOC/PT für Ereignisse mit einer Häufigkeit von ≥ 10 Patienten und ≥ 1 % in mindestens einem Behandlungsarm und
  - nur für SUE mit einer Häufigkeit von ≥ 5 % in mindestens einem Behandlungsarm.
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für UE nach SOC und PT und Schweregrad (nicht schwer = CTCAE-Grade 1 oder 2; schwer = CTCAE-Grad ≥3):
  - nur SOC/PT für Ereignisse mit einer Häufigkeit von ≥ 10 Patienten und ≥ 1 % in mindestens einem Behandlungsarm und
  - nur für UE mit einer Häufigkeit von ≥ 5 % in mindestens einem Behandlungsarm.
- Zeit bis zu einem Ereignis aus einer Kaplan-Meier-Schätzung für UE von besonderem Interesse nach Schweregrad bzw. Einstufung als SUE (nicht schwer = CTCAE-Grade 1 oder 2; schwer = CTCAE-Grad ≥ 3; schwerwiegend = klassifiziert als SUE). Die Selektion

der UE von besonderem Interesse erfolgte gemäß Studienbericht (Tabellen 14.3.1-4.1 und -4.2).

Neben der Inzidenz werden jeweils die Risikoschätzer OR, RR und RD mit den jeweils dazugehörigen Konfidenzintervallen und p-Werten berichtet. Eine Ausnahme hiervon bilden die Inzidenz von Studienabbrüchen wegen UE nach SOC und PT bzw. die Inzidenz der Behandlungsabbrüche wegen UE nach SOC und PT; diese werden gemäß neuer Dossier-Vorlage nur deskriptiv dargestellt. Für die Inzidenz von UE, SUE, Schweren UE und UE nach Schweregrad nach SOC und PT werden nur diejenigen im Modul 4 direkt dargestellt, für die die RR einen statistisch signifikanten Gruppenunterschied ergab. Eine Darstellung aller SOC/PT ist in Anhang 1 bzw. Anhang 2 zum Modul 4 zu finden.

Zu den Ereigniszeitanalysen werden angegeben: der Kaplan-Meier-Schätzer für den Anteil der Patienten ohne Ereignis und der Schätzer für die HR mit dem dazugehörigen Konfidenzintervall und p-Wert, basierend auf einer Cox-Regressionsanalyse. Von den Ereigniszeitanalysen werden nur die Gesamtraten der UE, SUE, Schweren UE, Studienabbrüche aufgrund von UE und Behandlungsabbrüche aufgrund von UE im Modul 4 direkt dargestellt. Die Auswertungen der Zeit bis zum Ereignis nach Schweregrad, von Unerwünschten Ereignissen von besonderem Interesse und Auswertungen nach SOC/PT werden nur in Anhang 1 bzw. Anhang 2 dargestellt.

Eine Liste der als krankheitsspezifisch definierten UE ist in den Anhängen 1 und 2 zu finden. Die Selektion krankheitsspezifischer UE und SUE, die auf die Grunderkrankungen Multiple Sklerose zurückzuführen sind, erfolgte systematisch auf folgende Weise: Als auf die Multiple Sklerose zurückzuführend wurden UE und SUE eingestuft, wenn sie in die vom G-BA empfohlenen Endpunktkategorien fielen. Die Einstufung wurde unabhängig durch zwei Mitarbeiter aus der Medizin, die mehrjährige Erfahrung im Indikationsgebiet Multiple Sklerose haben, durchgeführt. Nur UE/SUE, die durch beide Mitarbeiter als auf die Multiple Sklerose zurückzuführen eingestuft wurden, wurden aus der Liste entfernt.

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 4-99: Bewertung des Verzerrungspotenzials für den Endpunkt *Unerwünschte Ereignisse* in RCT mit dem zu bewertenden Arzneimittel

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

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Unerwünschte Ereignisse* als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkterheber (Prüfarzt) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde und eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war. Die Kriterien für Unerwünschte Ereignisse waren im Prüfplan a priori definiert. Die Beobachtungsdauer war in den Studienarmen in der Gesamtpopulation nicht unterschiedlich. In der Subpopulation B lag der Unterschied in der Beobachtungsdauer bei etwa 50 Tagen. Bezogen auf die Gesamtdauer von über 600 Tagen erscheint dieser Unterschied von geringer Relevanz. Insgesamt kann von einem objektivierbaren Endpunkt ausgegangen werden und es ist nur mit einem niedrigen daraus resultierenden Verzerrungspotenzial zu rechnen.

Aufgrund des Unterschieds der Beobachtungsdauer in der Subpopulation wurden Ereigniszeitanalysen für beide Auswertungspopulationen durchgeführt und im Modul 4A bzw. im Anhang zu Modul 4A dargestellt. Aufgrund der geringen Relevanz des Unterschieds werden zur Nutzenbewertung die binären Auswertungen zur Häufigkeit herangezogen.

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

Inzidenz von UE insgesamt sowie ohne krankheitsspezifische UE

Tabelle 4-100: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 1.099)	**BSC (N = 546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a,c] p-Wert**	**RD [95%-KI][a] p-Wert**
Unerwünschte Ereignisse					
N'	1.099	546			
UE – n (%)	975 (88,7)	445 (81,5)	1,78 [1,34; 2,37] < 0,001	1,09 [1,04; 1,14] < 0,001	0,07 [0,03; 0,11] < 0,001
UE – n (%) ohne krankheits-spezifische UE	918 (83,5)	418 (76,6)	1,55 [1,20; 2,00] 0,001	1,09 [1,03; 1,15] 0,001	0,07 [0,03; 0,11] 0,001
SUE – n (%)	186 (16,9)	74 (13,6)	1,30 [0,97; 1,74] 0,078	1,25 [0,97; 1,60] 0,080	0,03 [-0,00; 0,07] 0,069
SUE – n (%) ohne krankheits-spezifische SUE	146 (13,3)	53 (9,7)	1,43 [1,02; 1,99] 0,036	1,37 [1,02; 1,84] 0,038	0,04 [0,00; 0,07] 0,028

Schwere UE[b] – n (%)	139 (12,6)	61 (11,2)	1,15 [0,84; 1,59] 0,388	1,13 [0,85; 1,50] 0,390	0,01 [-0,02; 0,05] 0,380
Schwere UE – n (%) ohne krankheits-spezifische UE	110 (10,0)	44 (8,1)	1,27 [0,88; 1,83] 0,201	1,24 [0,89; 1,73] 0,204	0,02 [-0,01; 0,05] 0,186
Studienabbruch aufgrund von UE – n (%)	46 (4,2)	16 (2,9)	1,45 [0,81; 2,58] 0,208	1,43 [0,82; 2,50] 0,212	0,01 [-0,01; 0,03] 0,182
Behandlungs-abbruch aufgrund von UE – n (%)	84 (7,6)	28 (5,1)	1,53 [0,99; 2,38] 0,057	1,49 [0,98; 2,26] 0,060	0,03 [0,00; 0,05] 0,042

CTCAE: Common Terminology Criteria for Adverse Events; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; SUE: Schwerwiegendes Unerwünschtes Ereignis; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: CTCAE-Grad ≥ 3.

c: Für die Bewertung der Unerwünschten Ereignisse wird das relative Risiko im Vergleich der Behandlungsgruppen herangezogen.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.1.1.

Tabelle 4-101: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N = 61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a,c] p-Wert**	**RD [95%-KI][a] p-Wert**
Unerwünschte Ereignisse					
N'	127	61			
UE – n (%)	114 (89,8)	47 (77,0)	2,61 [1,14; 5,98] 0,020	1,17 [1,00; 1,35] 0,045	0,13 [0,01; 0,25] 0,035
UE – n (%) ohne krankheits-spezifische UE	108 (85,0)	46 (75,4)	1,85 [0,87; 3,96] 0,108	1,13 [0,96; 1,32] 0,143	0,10 [-0,03; 0,22] 0,130
SUE – n (%)	16 (12,6)	11 (18,0)	0,66 [0,28; 1,51] 0,320	0,70 [0,35; 1,41] 0,318	-0,05 [-0,17; 0,06] 0,343
SUE – n (%) ohne krankheits-spezifische SUE	12 (9,4)	8 (13,1)	0,69 [0,27; 1,79] 0,445	0,72 [0,31; 1,67] 0,445	-0,04 [-0,14; 0,06] 0,467
Schwere UE[b] – n (%)	11 (8,7)	10 (16,4)	0,48 [0,19; 1,21] 0,115	0,53 [0,24; 1,18] 0,118	-0,08 [-0,18; 0,03] 0,149

Schwere UE – n (%) ohne krankheits-spezifische UE	9 (7,1)	8 (13,1)	0,51 [0,18; 1,38] 0,177	0,54 [0,22; 1,33] 0,181	-0,06 [-0,16; 0,04] 0,217
Studienabbruch aufgrund von UE – n (%)	2 (1,6)	3 (4,9)	0,31 [0,05; 1,90] 0,182	0,32 [0,05; 1,87] 0,205	-0,03 [-0,09; 0,02] 0,262
Behandlungs-abbruch aufgrund von UE – n (%)	6 (4,7)	4 (6,6)	0,71 [0,19; 2,60] 0,600	0,72 [0,21; 2,46] 0,601	-0,02 [-0,09; 0,05] 0,619

CTCAE: Common Terminology Criteria for Adverse Events; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; SUE: Schwerwiegendes Unerwünschtes Ereignis; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: CTCAE-Grad ≥ 3.

c: Für die Bewertung der Unerwünschten Ereignisse wird das relative Risiko im Vergleich der Behandlungsgruppen herangezogen.

Quelle: Anhang 2 zu Modul 4, Tabelle 11.1.1.

Unerwünschte Ereignisse waren insgesamt in der Gesamtpopulation unter Siponimod statistisch signifikant häufiger, der Gruppenunterschied war aber nur gering (RR = 1,09; p < 0,001). Bei der Auswertung der SUE wurde keine statistische Signifikanz beobachtet. Bei Nichtberücksichtigung krankheitsspezifischer UE war die Häufigkeit von UE und SUE insgesamt unter Siponimod höher als unter BSC (RR = 1,09; p = 0,001 bzw. RR = 1,37; p = 0,038).

In der Subpopulation B waren nur die UE insgesamt unter Siponimod statistisch signifikant häufiger, der Unterschied zwischen den Gruppen aber nur gering (RR = 1,17; p = 0,045). Die Häufigkeit von SUE war nicht statistisch signifikant unterschiedlich zwischen den Gruppen.

Inzidenz von UE nach Schweregrad

Tabelle 4-102: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse nach Schweregrad* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 1.099)	**BSC (N = 546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a,e] p-Wert**	**RD [95%-KI][a] p-Wert**
Unerwünschte Ereignisse nach Schweregrad[b]					
N'	1.099	546			
CTCAE-Grad 1 – n (%)	900 (81,9)	405 (74,2)	1,57 [1,23; 2,01] < 0,001	1,10 [1,04; 1,17] 0,001	0,08 [0,03; 0,12] < 0,001

CTCAE-Grad 2 – n (%)	520 (47,3)	239 (43,8)	1,15 [0,94; 1,42] 0,175	1,08 [0,96; 1,21] 0,180	0,04 [-0,02; 0,09] 0,174
CTCAE-Grad 3 – n (%)	121 (11,0)	55 (10,1)	1,10 [0,79; 1,55] 0,563	1,09 [0,81; 1,48] 0,564	0,01 [-0,02; 0,04] 0,558
CTCAE-Grad 4 – n (%)	23 (2,1)	7 (1,3)	1,65 [0,70; 3,86] 0,247	1,63 [0,70; 3,78] 0,253	0,01 [-0,00; 0,02] 0,210
Schwere UE[c]– n (%)	139 (12,6)	61 (11,2)	1,15 [0,84; 1,59] 0,388	1,13 [0,85; 1,50] 0,390	0,01 [-0,02; 0,05] 0,380
Nicht schwere UE[d] – n (%)	965 (87,8)	438 (80,2)	1,78 [1,35; 2,34] < 0,001	1,09 [1,04; 1,15] < 0,001	0,08 [0,04; 0,11] < 0,001

CTCAE: Common Terminology Criteria for Adverse Events; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Die Summe der Häufigkeiten in den einzelnen Schweregraden liegt höher als die Gesamtrate der UE, da für jeden Grad alle Patienten gezählt wurden, die mindestens ein UE des jeweiligen Schweregrads hatten.

c: CTCAE-Grad ≥ 3.

d: CTCAE-Grad < 3.

e: Für die Bewertung der Unerwünschten Ereignisse wird das relative Risiko im Vergleich der Behandlungsgruppen herangezogen.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.1.2.

Tabelle 4-103: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse nach Schweregrad* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N = 61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a,e] p-Wert**	**RD [95%-KI][a] p-Wert**
Unerwünschte Ereignisse nach Schweregrad[b]					
N'	127	61			
CTCAE-Grad 1 – n (%)	108 (85,0)	40 (65,6)	2,98 [1,45; 6,12] 0,002	1,30 [1,07; 1,58] 0,009	0,19 [0,06; 0,33] 0,005
CTCAE-Grad 2 – n (%)	60 (47,2)	28 (45,9)	1,06 [0,57; 1,95] 0,863	1,03 [0,74; 1,43] 0,863	0,01 [-0,14; 0,17] 0,863
CTCAE-Grad 3 – n (%)	10 (7,9)	10 (16,4)	0,44 [0,17; 1,11] 0,076	0,48 [0,21; 1,09] 0,080	-0,09 [-0,19; 0,02] 0,109

CTCAE-Grad 4 – n (%)	3 (2,4)	0 (0,0)	3,46 [0,18; 68,00] 0,226	3,39 [0,18; 64,63] 0,417	0,02 [-0,00; 0,05] 0,080
Schwere UE[c] – n (%)	11 (8,7)	10 (16,4)	0,48 [0,19; 1,21] 0,115	0,53 [0,24; 1,18] 0,118	-0,08 [-0,18; 0,03] 0,149
Nicht schwere UE[d] – n (%)	113 (89,0)	47 (77,0)	2,40 [1,06; 5,43] 0,032	1,15 [0,99; 1,34] 0,060	0,12 [0,00; 0,24] 0,049

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Die Summe der Häufigkeiten in den einzelnen Schweregraden liegt höher als die Gesamtrate der UE, da für jeden Grad alle Patienten gezählt wurden, die mindestens ein UE des jeweiligen Schweregrads hatten.

c: CTCAE-Grad ≥ 3.

d: CTCAE-Grad < 3.

e: Für die Bewertung der Unerwünschten Ereignisse wird das relative Risiko im Vergleich der Behandlungsgruppen herangezogen.

Quelle: Anhang 2 zu Modul 4, Tabelle 11.1.2.

Die Mehrheit der UE waren von leichter Ausprägung, schwere UE (CTCAE-Grad ≥ 3) traten nur bei rund 10 % der Patienten auf. Leichte und moderate UE (CTCAE-Grad < 3) waren in der Gesamtpopulation unter Siponimod statistisch signifikant häufiger, der Unterschied war aber gering (RR = 1,09; p < 0,001). Dies gilt auch für die Betrachtung der UE von CTCAE-Grad 1 alleine (RR = 1,10; p = 0,001).

In der Subpopulation B waren UE vom CTCAE-Grad 1 unter Siponimod statistisch signifikant häufiger (RR = 1,30; p = 0,009). Auch hier traten UE schwerer Ausprägung unter Siponimod bei unter 10 % der Patienten auf.

Inzidenz von UE von besonderem Interesse nach Schweregrad

Tabelle 4-104: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse von besonderem Interesse nach Schweregrad* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 1.099)	BSC (N = 546)	OR [95%-KI][a] p-Wert	RR [95%-KI][a,c] p-Wert	RD [95%-KI][a] p-Wert
Unerwünschte Ereignisse von besonderem Interesse nach Schweregrad[b]					
N'	1.099	546			
Infektionen / parasitäre Erkrankungen – n (%)	539 (49,0)	268 (49,1)	1,00 [0,81; 1,23] 0,988	1,00 [0,90; 1,11] 0,988	-0,00 [-0,05; 0,05] 0,988
Nicht schwer – n (%)	528 (48,0)	262 (48,0)	1,00 [0,82; 1,23] 0,982	1,00 [0,90; 1,11] 0,982	0,00 [-0,05; 0,05] 0,982
Schwer – n (%)	24 (2,2)	11 (2,0)	1,09 [0,53; 2,23] 0,823	1,08 [0,53; 2,20] 0,823	0,00 [-0,01; 0,02] 0,821
Schwerwiegend – n (%)	33 (3,0)	15 (2,7)	1,10 [0,59; 2,04] 0,772	1,09 [0,60; 1,99] 0,772	0,00 [-0,01; 0,02] 0,769
Herpesvirusinfektionen – n (%)	25 (2,3)	11 (2,0)	1,13 [0,55; 2,32] 0,734	1,13 [0,56; 2,28] 0,734	0,00 [-0,01; 0,02] 0,729
Nicht schwer – n (%)	25 (2,3)	11 (2,0)	1,13 [0,55; 2,32] 0,734	1,13 [0,56; 2,28] 0,734	0,00 [-0,01; 0,02] 0,729
Schwer – n (%)	1 (0,1)	0 (0,0)	1,49 [0,06; 36,70] 0,481	1,49 [0,06; 36,56] 0,806	0,00 [-0,00; 0,00] 0,317
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Varizella-Zoster-Infektion – n (%)	33 (3,0)	4 (0,7)	4,19 [1,48; 11,90] 0,003	4,10 [1,46; 11,51] 0,007	0,02 [0,01; 0,04] < 0,001
Nicht schwer – n (%)	30 (2,7)	4 (0,7)	3,80 [1,33; 10,85] 0,007	3,73 [1,32; 10,52] 0,013	0,02 [0,01; 0,03] 0,001
Schwer – n (%)	3 (0,3)	0 (0,0)	3,49 [0,18; 67,66] 0,222	3,48 [0,18; 67,27] 0,409	0,00 [-0,00; 0,01] 0,083
Schwerwiegend – n (%)	2 (0,2)	0 (0,0)	2,49 [0,12; 51,95] 0,319	2,49 [0,12; 51,70] 0,556	0,00 [-0,00; 0,00] 0,157

Lymphopenie – n (%)	14 (1,3)	0 (0,0)	14,60 [0,87; 245,21] 0,008	14,42 [0,86; 241,29] 0,063	0,01 [0,01; 0,02] < 0,001
Nicht schwer – n (%)	10 (0,9)	0 (0,0)	10,53 [0,62; 180,09] 0,025	10,44 [0,61; 177,88] 0,105	0,01 [0,00; 0,01] 0,001
Schwer – n (%)	4 (0,4)	0 (0,0)	4,49 [0,24; 83,54] 0,158	4,48 [0,24; 82,98] 0,314	0,00 [0,00; 0,01] 0,045
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Makulaödem – n (%)	20 (1,8)	1 (0,2)	10,10 [1,35; 75,47] 0,005	9,94 [1,34; 73,84] 0,025	0,02 [0,01; 0,03] < 0,001
Nicht schwer – n (%)	15 (1,4)	0 (0,0)	15,62 [0,93; 261,57] 0,006	15,42 [0,92; 257,15] 0,057	0,01 [0,01; 0,02] < 0,001
Schwer – n (%)	5 (0,5)	1 (0,2)	2,49 [0,29; 21,37] 0,389	2,48 [0,29; 21,21] 0,406	0,00 [-0,00; 0,01] 0,320
Schwerwiegend – n (%)	3 (0,3)	0 (0,0)	3,49 [0,18; 67,66] 0,222	3,48 [0,18; 67,27] 0,409	0,00 [-0,00; 0,01] 0,083
Erhöhte Leberwerte – n (%)	135 (12,3)	21 (3,8)	3,50 [2,18; 5,61] < 0,001	3,19 [2,04; 5,00] < 0,001	0,08 [0,06; 0,11] < 0,001
Nicht schwer – n (%)	125 (11,4)	19 (3,5)	3,56 [2,17; 5,83] < 0,001	3,27 [2,04; 5,24] < 0,001	0,08 [0,05; 0,10] < 0,001
Schwer – n (%)	13 (1,2)	3 (0,5)	2,17 [0,61; 7,64] 0,218	2,15 [0,62; 7,52] 0,230	0,01 [-0,00; 0,02] 0,163
Schwerwiegend – n (%)	14 (1,3)	3 (0,5)	2,34 [0,67; 8,16] 0,171	2,32 [0,67; 8,03] 0,185	0,01 [-0,00; 0,02] 0,118
Malignitäten – n (%)	20 (1,8)	14 (2,6)	0,70 [0,35; 1,41] 0,318	0,71 [0,36; 1,39] 0,320	-0,01 [-0,02; 0,01] 0,345
Nicht schwer – n (%)	11 (1,0)	7 (1,3)	0,78 [0,30; 2,02] 0,606	0,78 [0,30; 2,00] 0,607	-0,00 [-0,01; 0,01] 0,620
Schwer – n (%)	11 (1,0)	7 (1,3)	0,78 [0,30; 2,02] 0,606	0,78 [0,30; 2,00] 0,607	-0,00 [-0,01; 0,01] 0,620
Schwerwiegend – n (%)	19 (1,7)	13 (2,4)	0,72 [0,35; 1,47] 0,367	0,73 [0,36; 1,46] 0,369	-0,01 [-0,02; 0,01] 0,392

Hypertonie – n (%)	137 (12,5)	50 (9,2)	1,41 [1,00; 1,99] 0,047	1,36 [1,00; 1,85] 0,049	0,03 [0,00; 0,06] 0,037
Nicht schwer – n (%)	137 (12,5)	49 (9,0)	1,44 [1,02; 2,04] 0,035	1,39 [1,02; 1,89] 0,038	0,03 [0,00; 0,07] 0,027
Schwer – n (%)	1 (0,1)	1 (0,2)	0,50 [0,03; 7,95] 0,613	0,50 [0,03; 7,93] 0,621	-0,00 [-0,00; 0,00] 0,652
Schwerwiegend – n (%)	1 (0,1)	0 (0,0)	1,49 [0,06; 36,70] 0,481	1,49 [0,06; 36,56] 0,806	0,00 [-0,00; 0,00] 0,317
Krampfanfälle – n (%)	19 (1,7)	2 (0,4)	4,79 [1,11; 20,62] 0,020	4,72 [1,10; 20,19] 0,036	0,01 [0,00; 0,02] 0,004
Nicht schwer – n (%)	16 (1,5)	2 (0,4)	4,02 [0,92; 17,54] 0,045	3,97 [0,92; 17,22] 0,065	0,01 [0,00; 0,02] 0,014
Schwer – n (%)	5 (0,5)	0 (0,0)	5,49 [0,30; 99,51] 0,114	5,47 [0,30; 98,74] 0,250	0,00 [0,00; 0,01] 0,025
Schwerwiegend – n (%)	9 (0,8)	0 (0,0)	9,52 [0,55; 163,90] 0,034	9,45 [0,55; 162,03] 0,121	0,01 [0,00; 0,01] 0,003
Thromboembolische Ereignisse – n (%)	33 (3,0)	15 (2,7)	1,10 [0,59; 2,04] 0,772	1,09 [0,60; 1,99] 0,772	0,00 [-0,01; 0,02] 0,769
Nicht schwer – n (%)	25 (2,3)	12 (2,2)	1,04 [0,52; 2,08] 0,921	1,04 [0,52; 2,04] 0,921	0,00 [-0,01; 0,02] 0,921
Schwer – n (%)	9 (0,8)	5 (0,9)	0,89 [0,30; 2,68] 0,840	0,89 [0,30; 2,66] 0,841	-0,00 [-0,01; 0,01] 0,843
Schwerwiegend – n (%)	15 (1,4)	7 (1,3)	1,07 [0,43; 2,63] 0,890	1,06 [0,44; 2,60] 0,891	0,00 [-0,01; 0,01] 0,889
Atemwegserkrankungen – n (%)	117 (10,6)	71 (13,0)	0,80 [0,58; 1,09] 0,157	0,82 [0,62; 1,08] 0,156	-0,02 [-0,06; 0,01] 0,169
Nicht schwer – n (%)	114 (10,4)	70 (12,8)	0,79 [0,57; 1,08] 0,138	0,81 [0,61; 1,07] 0,137	-0,02 [-0,06; 0,01] 0,150
Schwer – n (%)	3 (0,3)	1 (0,2)	1,49 [0,15; 14,37] 0,728	1,49 [0,16; 14,30] 0,729	0,00 [-0,00; 0,01] 0,710
Schwerwiegend – n (%)	2 (0,2)	1 (0,2)	0,99 [0,09; 10,98] 0,996	0,99 [0,09; 10,93] 0,996	-0,00 [-0,00; 0,00] 0,996

Periphere Ödeme / Schwellungen – n (%)	74 (6,7)	18 (3,3)	2,12 [1,25; 3,58] 0,004	2,04 [1,23; 3,38] 0,006	0,03 [0,01; 0,06] 0,001
Nicht schwer – n (%)	74 (6,7)	18 (3,3)	2,12 [1,25; 3,58] 0,004	2,04 [1,23; 3,38] 0,006	0,03 [0,01; 0,06] 0,001
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Asthma / Broncho-spasmen – n (%)	8 (0,7)	2 (0,4)	1,99 [0,42; 9,42] 0,374	1,99 [0,42; 9,33] 0,384	0,00 [-0,00; 0,01] 0,321
Nicht schwer – n (%)	8 (0,7)	2 (0,4)	1,99 [0,42; 9,42] 0,374	1,99 [0,42; 9,33] 0,384	0,00 [-0,00; 0,01] 0,321
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwangerschaft / Neonatale Ereignisse – n (%)	6 (0,5)	0 (0,0)	6,50 [0,37; 115,54] 0,084	6,46 [0,36; 114,54] 0,203	0,01 [0,00; 0,01] 0,014
Nicht schwer – n (%)	6 (0,5)	0 (0,0)	6,50 [0,37; 115,54] 0,084	6,46 [0,36; 114,54] 0,203	0,01 [0,00; 0,01] 0,014
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Bradyarrhythmie/ Bradykardie – n (%)	103 (9,4)	31 (5,7)	1,72 [1,13; 2,60] 0,010	1,65 [1,12; 2,43] 0,011	0,04 [0,01; 0,06] 0,005
Nicht schwer – n (%)	101 (9,2)	30 (5,5)	1,74 [1,14; 2,65] 0,009	1,67 [1,13; 2,48] 0,011	0,04 [0,01; 0,06] 0,005
Schwer – n (%)	5 (0,5)	1 (0,2)	2,49 [0,29; 21,37] 0,389	2,48 [0,29; 21,21] 0,406	0,00 [-0,00; 0,01] 0,320
Schwerwiegend – n (%)	10 (0,9)	1 (0,2)	5,00 [0,64; 39,20] 0,089	4,97 [0,64; 38,71] 0,126	0,01 [0,00; 0,01] 0,032

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Die Summe der Häufigkeiten in den einzelnen Schweregraden liegt höher als die Gesamtrate der UE, da für jeden Grad alle Patienten gezählt wurden, die mindestens ein UE des jeweiligen Schweregrads hatten.

c: Für die Bewertung der Unerwünschten Ereignisse wird das relative Risiko im Vergleich der Behandlungsgruppen herangezogen.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.1.8.

Tabelle 4-105: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse von besonderem Interesse nach Schweregrad* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N = 61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a,c] p-Wert**	**RD [95%-KI][a] p-Wert**
Unerwünschte Ereignisse von besonderem Interesse nach Schweregrad[b]					
N'	127	61			
Infektionen / parasitäre Erkrankungen – n (%)	65 (51,2)	26 (42,6)	1,41 [0,76; 2,61] 0,272	1,20 [0,86; 1,68] 0,287	0,09 [-0,07; 0,24] 0,268
Nicht schwer – n (%)	64 (50,4)	26 (42,6)	1,37 [0,74; 2,53] 0,318	1,18 [0,84; 1,66] 0,332	0,08 [-0,07; 0,23] 0,315
Schwer – n (%)	4 (3,1)	2 (3,3)	0,96 [0,17; 5,39] 0,962	0,96 [0,18; 5,10] 0,962	-0,00 [-0,06; 0,05] 0,963
Schwerwiegend – n (%)	5 (3,9)	2 (3,3)	1,21 [0,23; 6,42] 0,823	1,20 [0,24; 6,01] 0,824	0,01 [-0,05; 0,06] 0,818
Herpesvirusinfektionen – n (%)	4 (3,1)	1 (1,6)	1,95 [0,21; 17,84] 0,547	1,92 [0,22; 16,83] 0,555	0,02 [-0,03; 0,06] 0,501
Nicht schwer – n (%)	4 (3,1)	1 (1,6)	1,95 [0,21; 17,84] 0,547	1,92 [0,22; 16,83] 0,555	0,02 [-0,03; 0,06] 0,501
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-

Varizella-Zoster-Infektion – n (%)	1 (0,8)	0 (0,0)	1,46 [0,06; 36,32] 0,487	1,45 [0,06; 35,16] 0,818	0,01 [-0,01; 0,02] 0,315
Nicht schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwer – n (%)	1 (0,8)	0 (0,0)	1,46 [0,06; 36,32] 0,487	1,45 [0,06; 35,16] 0,818	0,01 [-0,01; 0,02] 0,315
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Lymphopenie – n (%)	2 (1,6)	0 (0,0)	2,45 [0,12; 51,82] 0,324	2,42 [0,12; 49,69] 0,566	0,02 [-0,01; 0,04] 0,154
Nicht schwer – n (%)	2 (1,6)	0 (0,0)	2,45 [0,12; 51,82] 0,324	2,42 [0,12; 49,69] 0,566	0,02 [-0,01; 0,04] 0,154
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Makulaödem – n (%)	2 (1,6)	1 (1,6)	0,96 [0,09; 10,80] 0,974	0,96 [0,09; 10,39] 0,974	-0,00 [-0,04; 0,04] 0,974
Nicht schwer – n (%)	1 (0,8)	0 (0,0)	1,46 [0,06; 36,32] 0,487	1,45 [0,06; 35,16] 0,818	0,01 [-0,01; 0,02] 0,315
Schwer – n (%)	1 (0,8)	1 (1,6)	0,48 [0,03; 7,74] 0,594	0,48 [0,03; 7,55] 0,602	-0,01 [-0,04; 0,03] 0,637
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Erhöhte Leberenzymwerte – n (%)	19 (15,0)	4 (6,6)	2,51 [0,81; 7,72] 0,100	2,28 [0,81; 6,42] 0,118	0,08 [-0,00; 0,17] 0,061
Nicht schwer – n (%)	18 (14,2)	2 (3,3)	4,87 [1,09; 21,72] 0,023	4,32 [1,04; 18,04] 0,045	0,11 [0,03; 0,18] 0,005
Schwer – n (%)	1 (0,8)	2 (3,3)	0,23 [0,02; 2,63] 0,202	0,24 [0,02; 2,60] 0,240	-0,02 [-0,07; 0,02] 0,301
Schwerwiegend – n (%)	2 (1,6)	2 (3,3)	0,47 [0,06; 3,43] 0,448	0,48 [0,07; 3,33] 0,458	-0,02 [-0,07; 0,03] 0,501

Malignitäten – n (%)	0 (0,0)	1 (1,6)	0,16 [0,01; 3,94] 0,148	0,16 [0,01; 3,91] 0,262	-0,02 [-0,05; 0,02] 0,313
Nicht schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwer – n (%)	0 (0,0)	1 (1,6)	0,16 [0,01; 3,94] 0,148	0,16 [0,01; 3,91] 0,262	-0,02 [-0,05; 0,02] 0,313
Schwerwiegend – n (%)	0 (0,0)	1 (1,6)	0,16 [0,01; 3,94] 0,148	0,16 [0,01; 3,91] 0,262	-0,02 [-0,05; 0,02] 0,313
Hypertonie – n (%)	17 (13,4)	5 (8,2)	1,73 [0,61; 4,94] 0,300	1,63 [0,63; 4,22] 0,311	0,05 [-0,04; 0,14] 0,263
Nicht schwer – n (%)	17 (13,4)	5 (8,2)	1,73 [0,61; 4,94] 0,300	1,63 [0,63; 4,22] 0,311	0,05 [-0,04; 0,14] 0,263
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Krampfanfälle – n (%)	0 (0,0)	0 (0,0)	-	-	-
Nicht schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Thromboembolische Ereignisse – n (%)	5 (3,9)	1 (1,6)	2,46 [0,28; 21,52] 0,401	2,40 [0,29; 20,11] 0,419	0,02 [-0,02; 0,07] 0,332
Nicht schwer – n (%)	5 (3,9)	0 (0,0)	5,52 [0,30; 101,50] 0,116	5,33 [0,30; 94,83] 0,255	0,04 [0,01; 0,07] 0,023
Schwer – n (%)	0 (0,0)	1 (1,6)	0,16 [0,01; 3,94] 0,148	0,16 [0,01; 3,91] 0,262	-0,02 [-0,05; 0,02] 0,313
Schwerwiegend – n (%)	1 (0,8)	1 (1,6)	0,48 [0,03; 7,74] 0,594	0,48 [0,03; 7,55] 0,602	-0,01 [-0,04; 0,03] 0,637

Atemwegs-erkrankungen – n (%)	22 (17,3)	11 (18,0)	0,95 [0,43; 2,12] 0,905	0,96 [0,50; 1,85] 0,905	-0,01 [-0,12; 0,11] 0,905
Nicht schwer – n (%)	22 (17,3)	11 (18,0)	0,95 [0,43; 2,12] 0,905	0,96 [0,50; 1,85] 0,905	-0,01 [-0,12; 0,11] 0,905
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Periphere Ödeme / Schwellungen – n (%)	6 (4,7)	1 (1,6)	2,98 [0,35; 25,28] 0,296	2,88 [0,35; 23,41] 0,322	0,03 [-0,02; 0,08] 0,215
Nicht schwer – n (%)	6 (4,7)	1 (1,6)	2,98 [0,35; 25,28] 0,296	2,88 [0,35; 23,41] 0,322	0,03 [-0,02; 0,08] 0,215
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Asthma / Broncho-spasmen – n (%)	1 (0,8)	0 (0,0)	1,46 [0,06; 36,32] 0,487	1,45 [0,06; 35,16] 0,818	0,01 [-0,01; 0,02] 0,315
Nicht schwer – n (%)	1 (0,8)	0 (0,0)	1,46 [0,06; 36,32] 0,487	1,45 [0,06; 35,16] 0,818	0,01 [-0,01; 0,02] 0,315
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwangerschaft / Neonatale Ereignisse – n (%)	1 (0,8)	0 (0,0)	1,46 [0,06; 36,32] 0,487	1,45 [0,06; 35,16] 0,818	0,01 [-0,01; 0,02] 0,315
Nicht schwer – n (%)	1 (0,8)	0 (0,0)	1,46 [0,06; 36,32] 0,487	1,45 [0,06; 35,16] 0,818	0,01 [-0,01; 0,02] 0,315
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-

Bradyarrhythmie/ Bradykardie – n (%)	11 (8,7)	1 (1,6)	5,69 [0,72; 45,12] 0,065	5,28 [0,70; 40,00] 0,107	0,07 [0,01; 0,13] 0,018
Nicht schwer – n (%)	11 (8,7)	1 (1,6)	5,69 [0,72; 45,12] 0,065	5,28 [0,70; 40,00] 0,107	0,07 [0,01; 0,13] 0,018
Schwer – n (%)	0 (0,0)	0 (0,0)	-	-	-
Schwerwiegend – n (%)	0 (0,0)	0 (0,0)	-	-	-

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Die Summe der Häufigkeiten in den einzelnen Schweregraden liegt höher als die Gesamtrate der UE, da für jeden Grad alle Patienten gezählt wurden, die mindestens ein UE des jeweiligen Schweregrads hatten.

c: Für die Bewertung der Unerwünschten Ereignisse wird das relative Risiko im Vergleich der Behandlungsgruppen herangezogen.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.1.8.

Unter den UE von besonderem Interesse über alle Schweregrade traten in der Gesamtpopulation unter einer Therapie mit Siponimod Varizella-Zoster-Infektionen, Makulaödeme, erhöhte Leberenzymwerte, Hypertonie, Krampfanfälle, periphere Ödeme bzw. Schwellungen und Bradyarrhythmien statistisch signifikant häufiger als unter BSC auf. Dabei handelte es sich überwiegend um leichte und moderate UE (CTCAE-Grad < 3). Die Betrachtung der UE von besonderem Interesse nach Schweregrad zeigte eine statistisch signifikant höhere Häufigkeit bei Varizella-Zoster-Infektionen, erhöhten Leberenzymwerten, Hypertonie, periphere Ödeme bzw. Schwellungen und Bradyarrhythmien, jeweils in der Kategorie der nicht schweren UE (CTCAE-Grad < 3). Bei den erhöhten Leberenzymwerten handelt es sich um reine Laborparameter, die nicht als patientenrelevant gewertet werden.

Bei den Krampfanfällen handelte es sich in etwa der Hälfte der Fälle um SUE. Von den insgesamt 19 Patienten mit Krampfanfällen unter Siponimod hat sich in zwei Fällen später herausgestellt, dass es sich lediglich um Beinkrämpfe gehandelt hatte. Von den letztlich 17 Patienten mit Epilepsie bzw. epileptischen Anfällen war bei sieben Patienten eine relevante medizinische Vorgeschichte bekannt (9). Neun Fälle wurden als SUE klassifiziert. In einem Fall führte das SUE zum Abbruch der Studienmedikation, in allen anderen Fällen wurde die Behandlung gar nicht oder nur vorübergehend unterbrochen. Bei vier Patienten handelte es sich bei den SUE um schwere Ereignisse mit einem CTCAE-Grad von 3 bzw. 4.

In der Subpopulation B zeigte das relative Risiko für das Auftreten von UE von besonderem Interesse über alle Schweregrade keine statistisch signifikant erhöhte Häufigkeit unter einer Therapie mit Siponimod im Vergleich zu BSC. Die Betrachtung der UE von besonderem Interesse nach Schweregrad zeigte eine statistisch signifikant höhere Häufigkeit bei erhöhten Leberenzymwerten in der Kategorie der nicht schweren UE (CTCAE-Grad < 3). Bei den

erhöhten Leberenzymwerten handelt es sich um reine Laborparameter, die nicht als patientenrelevant gewertet werden.

Inzidenz von UE und SUE nach SOC und PT nach Schweregrad

Tabelle 4-106: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse (nach SOC und PT)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 1.099)	BSC (N = 546)	OR [95%-KI][a] p-Wert	RR [95%-KI][a,c] p-Wert	RD [95%-KI][a] p-Wert
N'	1.099	546			
Unerwünschte Ereignisse[b]					
Erkrankungen des Blutes und des Lymphsystems (SOC) – n (%)	46 (4,2)	10 (1,8)	2,34 [1,17; 4,68] 0,013	2,29 [1,16; 4,49] 0,017	0,02 [0,01; 0,04] 0,005
Vertigo – n (%)	25 (2,3)	26 (4,8)	0,47 [0,27; 0,81] 0,006	0,48 [0,28; 0,82] 0,007	-0,02 [-0,04; -0,00] 0,014
Makulaödem – n (%)	18 (1,6)	1 (0,2)	9,07 [1,21; 68,16] 0,009	8,94 [1,20; 66,81] 0,033	0,01 [0,01; 0,02] 0,001
Übelkeit – n (%)	74 (6,7)	19 (3,5)	2,00 [1,20; 3,35] 0,007	1,93 [1,18; 3,17] 0,009	0,03 [0,01; 0,05] 0,003
Gangstörung – n (%)	24 (2,2)	22 (4,0)	0,53 [0,30; 0,96] 0,033	0,54 [0,31; 0,96] 0,035	-0,02 [-0,04; 0,00] 0,052
Peripheres Ödem – n (%)	50 (4,5)	13 (2,4)	1,95 [1,05; 3,63] 0,031	1,91 [1,05; 3,49] 0,035	0,02 [0,00; 0,04] 0,017
Leber- und Gallenerkrankungen (SOC) – n (%)	27 (2,5)	4 (0,7)	3,41 [1,19; 9,80] 0,015	3,35 [1,18; 9,54] 0,023	0,02 [0,01; 0,03] 0,004
Herpes zoster – n (%)	25 (2,3)	4 (0,7)	3,15 [1,09; 9,11] 0,025	3,11 [1,09; 8,88] 0,035	0,02 [0,00; 0,03] 0,008
Untersuchungen (SOC) – n (%)	263 (23,9)	80 (14,7)	1,83 [1,39; 2,41] < 0,001	1,63 [1,30; 2,05] < 0,001	0,09 [0,05; 0,13] < 0,001
ALT erhöht – n (%)	65 (5,9)	8 (1,5)	4,23 [2,01; 8,88] < 0,001	4,04 [1,95; 8,35] < 0,001	0,04 [0,03; 0,06] < 0,001

Gamma-GT erhöht – n (%)	43 (3,9)	6 (1,1)	3,66 [1,55; 8,66] 0,002	3,56 [1,52; 8,31] 0,003	0,03 [0,01; 0,04] < 0,001
Leberenzym erhöht – n (%)	16 (1,5)	1 (0,2)	8,05 [1,07; 60,87] 0,016	7,95 [1,06; 59,78] 0,044	0,01 [0,00; 0,02] 0,002
Erkrankungen des Nervensystems (SOC) – n (%)	416 (37,9)	175 (32,1)	1,29 [1,04; 1,60] 0,021	1,18 [1,02; 1,36] 0,023	0,06 [0,01; 0,11] 0,019
Parästhesie – n (%)	18 (1,6)	20 (3,7)	0,44 [0,23; 0,83] 0,010	0,45 [0,24; 0,84] 0,012	-0,02 [-0,04; -0,00] 0,023
Gefäßerkrankungen (SOC) – n (%)	160 (14,6)	60 (11,0)	1,38 [1,01; 1,89] 0,045	1,32 [1,00; 1,75] 0,048	0,04 [0,00; 0,07] 0,037
Schwerwiegende Unerwünschte Ereignisse[b]					
Untersuchungen (SOC) – n (%)	21 (1,9)	3 (0,5)	3,53 [1,05; 11,87] 0,030	3,48 [1,04; 11,61] 0,043	0,01 [0,00; 0,02] 0,009
Unerwünschte Ereignisse nach Schweregrad[b]					
Erkrankungen des Blutes und des Lymphsystems (SOC) (CTCAE-Grad < 3)– n (%)	41 (3,7)	8 (1,5)	2,61 [1,21; 5,60] 0,011	2,55 [1,20; 5,39] 0,015	0,02 [0,01; 0,04] 0,003
Vertigo (CTCAE-Grad < 3) – n (%)	25 (2,3)	25 (4,6)	0,49 [0,28; 0,85] 0,010	0,50 [0,29; 0,86] 0,012	-0,02 [-0,04; -0,00] 0,021
Erkrankungen des Gastrointestinal-trakts (SOC) (CTCAE-Grad < 3)– n (%)	268 (24,4)	107 (19,6)	1,32 [1,03; 1,70] 0,029	1,24 [1,02; 1,52] 0,032	0,05 [0,01; 0,09] 0,025
Übelkeit (CTCAE-Grad < 3) – n (%)	73 (6,6)	19 (3,5)	1,97 [1,18; 3,30] 0,009	1,91 [1,16; 3,13] 0,010	0,03 [0,01; 0,05] 0,004
Peripheres Ödem (CTCAE-Grad < 3) – n (%)	50 (4,5)	13 (2,4)	1,95 [1,05; 3,63] 0,031	1,91 [1,05; 3,49] 0,035	0,02 [0,00; 0,04] 0,017
Untersuchungen (SOC) (CTCAE-Grad < 3) – n (%)	252 (22,9)	77 (14,1)	1,81 [1,37; 2,40] < 0,001	1,63 [1,29; 2,05] < 0,001	0,09 [0,05; 0,13] < 0,001
ALT erhöht (CTCAE-Grad < 3) – n (%)	59 (5,4)	7 (1,3)	4,37 [1,98; 9,63] < 0,001	4,19 [1,93; 9,11] < 0,001	0,04 [0,02; 0,06] < 0,001
Gamma-GT erhöht (CTCAE-Grad < 3) – n (%)	42 (3,8)	6 (1,1)	3,58 [1,51; 8,47] 0,002	3,48 [1,49; 8,13] 0,004	0,03 [0,01; 0,04] <0,001

Stoffwechsel- und Ernährungs-störungen (SOC) (CTCAE-Grad < 3) – n (%)	83 (7,6)	27 (4,9)	1,57 [1,00; 2,46] 0,046	1,53 [1,00; 2,33] 0,049	0,03 [0,00; 0,05] 0,033
Erkrankungen des Nervensystems (SOC) (CTCAE-Grad < 3) – n (%)	408 (37,1)	171 (31,3)	1,29 [1,04; 1,61] 0,020	1,19 [1,02; 1,37] 0,023	0,06 [0,01; 0,11] 0,018
Parästhesie (CTCAE-Grad < 3) – n (%)	18 (1,6)	20 (3,7)	0,44 [0,23; 0,83] 0,010	0,45 [0,24; 0,84] 0,012	-0,02 [-0,04; -0,00] 0,023
Gefäßerkrankungen (SOC) (CTCAE-Grad < 3) – n (%)	160 (14,6)	59 (10,8)	1,41 [1,02; 1,93] 0,035	1,35 [1,02; 1,78] 0,037	0,04 [0,00; 0,07] 0,027
Hypertonie (CTCAE-Grad < 3) – n (%)	115 (10,5)	40 (7,3)	1,48 [1,02; 2,15] 0,040	1,43 [1,01; 2,02] 0,043	0,03 [0,00; 0,06] 0,030

ALT: Alanin-Aminotransferase; CTCAE: Common Terminology Criteria for Adverse Events;GT: Glutamyltransferase; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N‘: Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; SOC: System Organ Class; SUE: Schwerwiegendes Unerwünschtes Ereignis; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Nur Ereignisse mit statistisch signifikantem Gruppenunterschied werden dargestellt. Eine vollständige Darstellung ist in Anhang 1 zu finden.

c: Für die Bewertung der Unerwünschten Ereignisse wird das relative Risiko im Vergleich der Behandlungsgruppen herangezogen.

Quelle: Anhang 1 zu Modul 4, Tabellen 11.1.3, 11.1.4 und 11.1.5.

Tabelle 4-107: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse (nach SOC und PT)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N = 61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a,c] p-Wert**	**RD [95%-KI][a] p-Wert**
N'	127	61			
Unerwünschte Ereignisse[b]					
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) – n (%)	31 (24,4)	6 (9,8)	2,96 [1,16; 7,54] 0,019	2,48 [1,09; 5,63] 0,030	0,15 [0,04; 0,25] 0,007
Untersuchungen (SOC) – n (%)	41 (32,3)	9 (14,8)	2,75 [1,24; 6,13] 0,011	2,19 [1,14; 4,21] 0,019	0,18 [0,05; 0,30] 0,004

Schwerwiegende Unerwünschte Ereignisse[b]					
Nicht zutreffend					
Unerwünschte Ereignisse nach Schweregrad[b]					
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort (SOC) (CTCAE-Grad < 3) – n (%)	30 (23,6)	6 (9,8)	2,84 [1,11; 7,24] 0,025	2,40 [1,06; 5,46] 0,037	0,14 [0,03; 0,24] 0,010
Untersuchungen (SOC) (CTCAE-Grad < 3) – n (%)	40 (31,5)	8 (13,1)	3,05 [1,33; 7,00] 0,007	2,40 [1,20; 4,81] 0,013	0,18 [0,07; 0,30] 0,002

CTCAE: Common Terminology Criteria for Adverse Events; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; SOC: System Organ Class.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Nur Ereignisse mit statistisch signifikantem Gruppenunterschied werden dargestellt. Eine vollständige Darstellung ist in Anhang 1 zu finden.

c: Für die Bewertung der Unerwünschten Ereignisse wird das relative Risiko im Vergleich der Behandlungsgruppen herangezogen.

Quelle: Anhang 2 zu Modul 4, Tabellen 11.1.3, 11.1.4 und 11.1.5.

Die Auswertungen der UE und SUE nach SOC bzw. PT in der Gesamtpopulation sowie in der Subpopulation B zeigen, dass die unter Siponimod statistisch signifikant häufiger auftretenden UE überwiegend nicht schwerwiegend und jeweils nur von geringem CTCAE-Grad waren.

Zeit bis zum ersten UE (insgesamt sowie ohne krankheitsspezifische UE)

Tabelle 4-108: Ergebnisse für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**HR [95%-KI][b]**	**p-Wert**
N'	1.099	546		
UE				
Anzahl der Ereignisse	975	445		
Median (in Wochen) [95%-KI][a]	4,1 [3,6; 5,3]	7,3 [5,0; 11,1]		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	6,82 [4,36; 9,27]	13,39 [9,40; 17,37]		
			1,26 [1,13; 1,41]	<0,001

UE ohne krankheitsspezifische Ereignisse				
Anzahl der Ereignisse	918	418		
Median (in Wochen) [95%-KI][a]	11,0 [9,0; 12,6]	13,0 [10,9; 17,0]		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	10,14 [7,33; 12,95]	16,95 [12,46; 21,44]		
			1,22 [1,09; 1,37]	0,001
SUE				
Anzahl der Ereignisse	186	74		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	75,71 [70,68; 80,74]	79,69 [71,61; 87,78]		
			1,26 [0,96; 1,65]	0,095
SUE ohne krankheitsspezifische Ereignisse				
Anzahl der Ereignisse	146	53		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	82,01 [78,87; 85,15]	84,05 [76,06; 92,03]		
			1,37 [1,00; 1,87]	0,050
Schwere UE (CTCAE-Grad ≥ 3)				
Anzahl der Ereignisse	139	61		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	83,41 [80,29; 86,53]	85,27 [81,28; 89,25]		
			1,13 [0,84; 1,53]	0,423
Schwere UE ohne krankheitsspezifische Ereignisse				
Anzahl der Ereignisse	110	44		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	86,19 [83,17; 89,22]	88,75 [85,04; 92,45]		
			1,23 [0,87; 1,75]	0,244
Studienabbrüche aufgrund von UE				
Anzahl der Ereignisse	46	16		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	95,50 [94,20; 96,79]	96,24 [94,18; 98,30]		
			1,43 [0,81; 2,53]	0,216

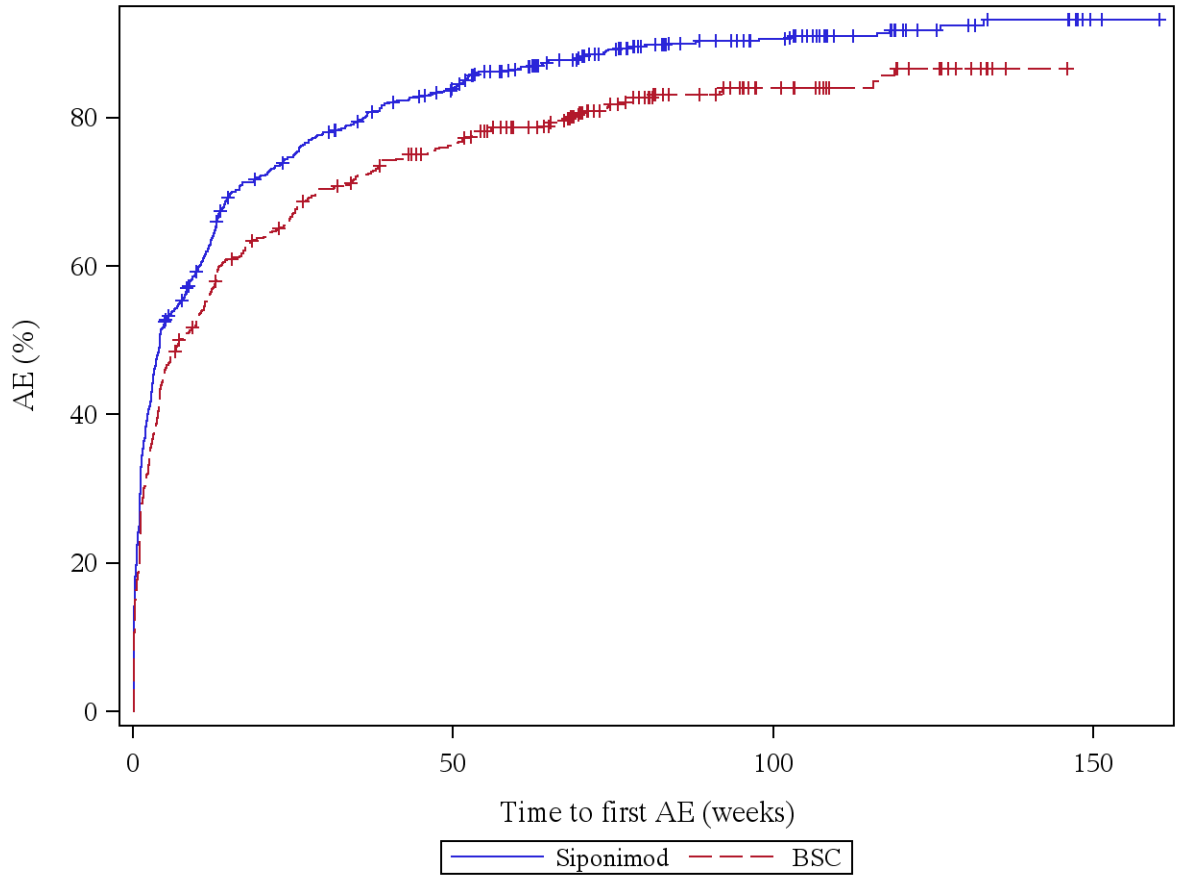
Behandlungsabbrüche aufgrund von UE				
Anzahl der Ereignisse	84	28		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	91,70 [89,97; 93,43]	93,52 [90,99; 96,06]		
			1,49 [0,97; 2,29]	0,066

CTCAE: Common Terminology Criteria for Adverse Events [*Allgemeine Terminologie und Merkmale unerwünschter Ereignisse*]; HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar; UE: Unerwünschtes Ereignis; SUE: Schwerwiegendes Unerwünschtes Ereignis.

a: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

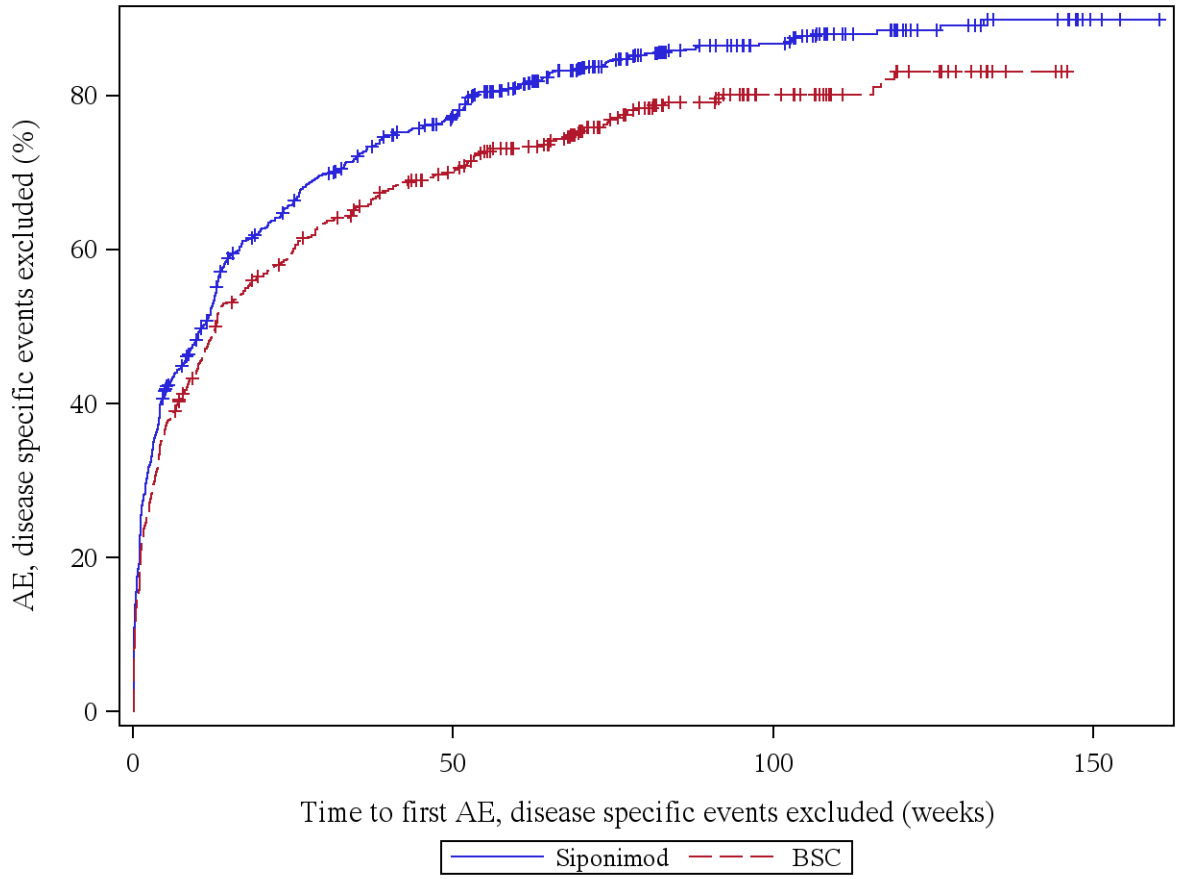
b: Cox proportional hazards model.

Quelle: Anhang 1 zu Modul 4, Tabelle 12.1.1.



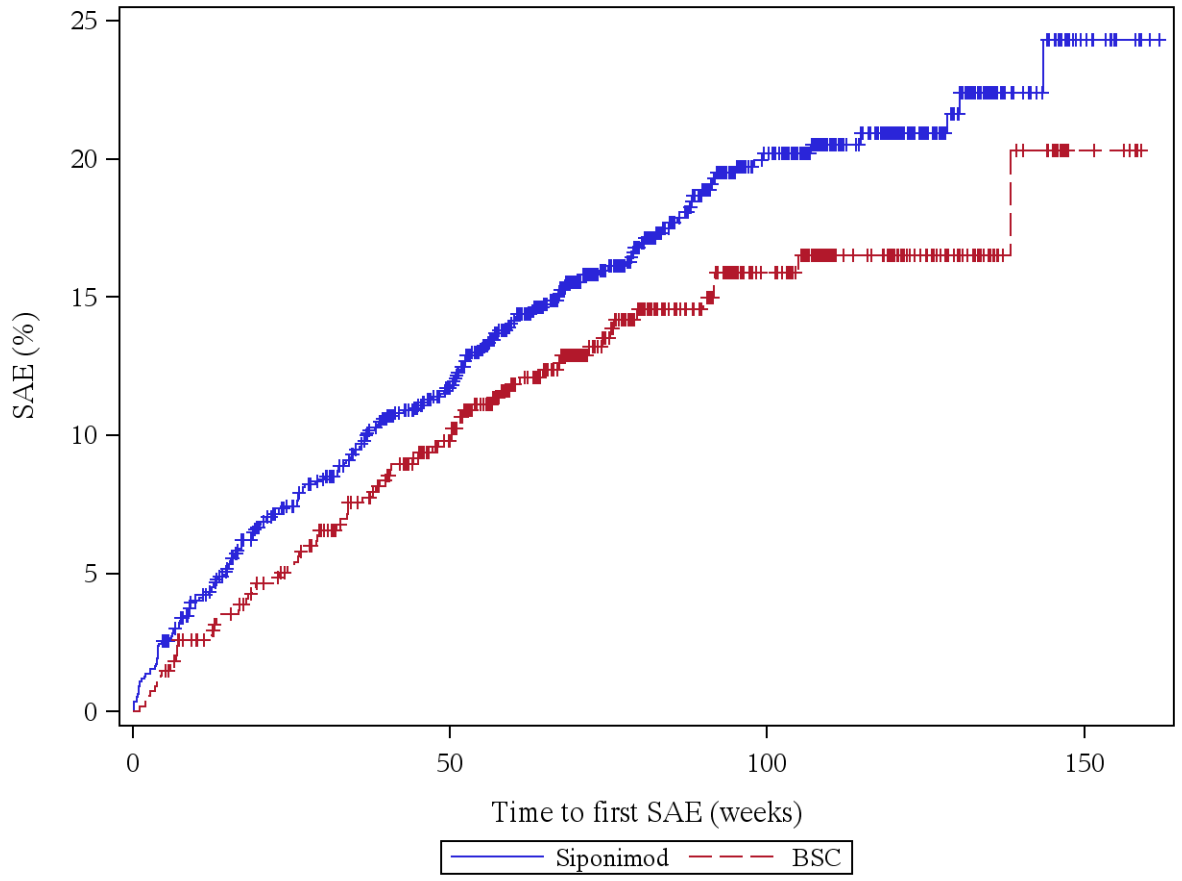
Patients at risk (Siponimod/BSC): Week 0: 1099/546, Week 4: 560/319, Week 12: 402/236, Week 24: 273/182, Week 36: 210/144, Week 48: 169/120, Week 60: 125/96, Week 72: 85/65, Week 84: 56/40, Week 96: 46/31, Week 108: 28/21, Week 120: 17/13, Week 132: 10/6, Week 144: 8/1, Week 156: 1/0, Week 168: 0/0

Abbildung 20: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)



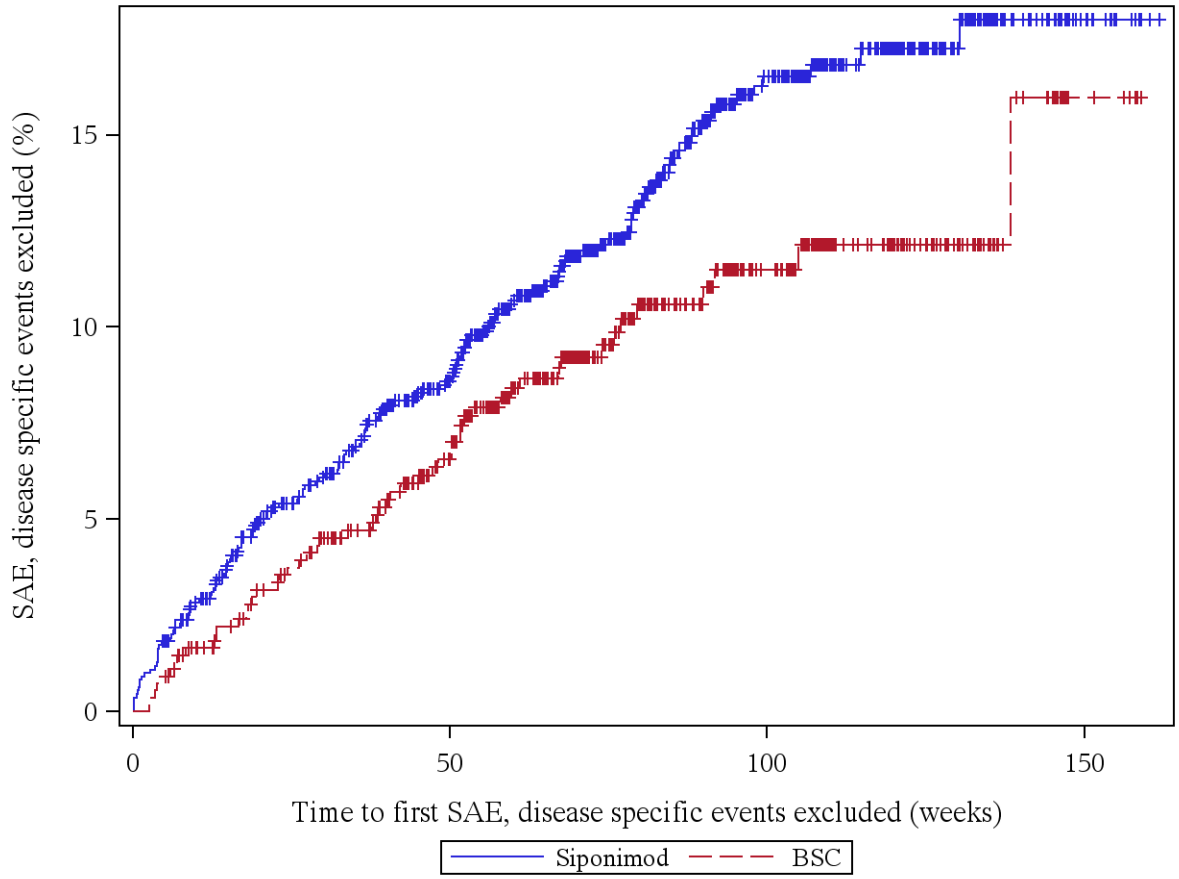
Patients at risk (Siponimod/BSC): Week 0: 1099/546, Week 4: 680/364, Week 12: 521/278, Week 24: 366/217, Week 36: 281/176, Week 48: 232/148, Week 60: 165/118, Week 72: 111/82, Week 84: 72/50, Week 96: 58/37, Week 108: 36/25, Week 120: 23/15, Week 132: 15/8, Week 144: 10/3, Week 156: 1/0, Week 168: 0/0

Abbildung 21: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)



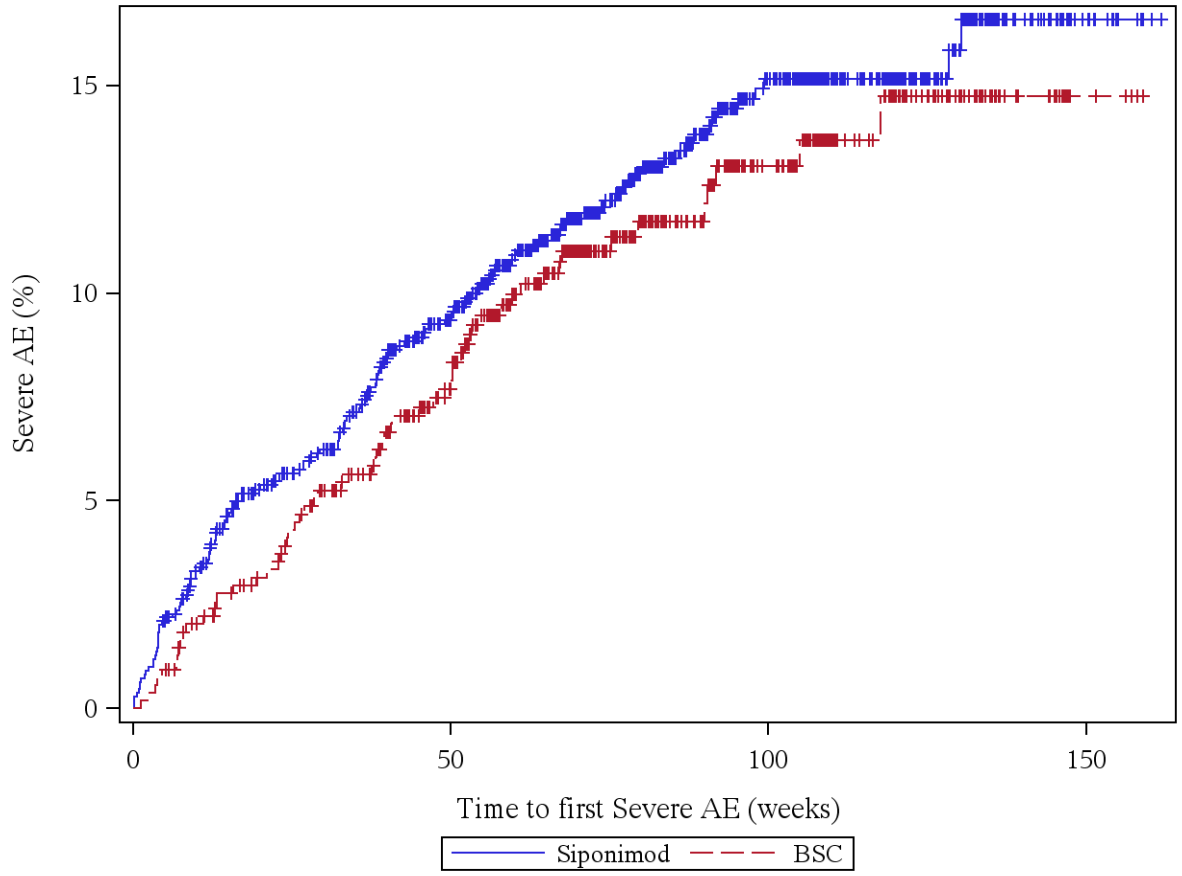
Patients at risk (Siponimod/BSC): Week 0: 1099/546, Week 4: 1073/540, Week 12: 1024/522, Week 24: 959/496, Week 36: 912/465, Week 48: 845/423, Week 60: 730/354, Week 72: 564/277, Week 84: 437/211, Week 96: 337/157, Week 108: 242/112, Week 120: 164/72, Week 132: 88/38, Week 144: 40/19, Week 156: 7/5, Week 168: 0/0

Abbildung 22: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)



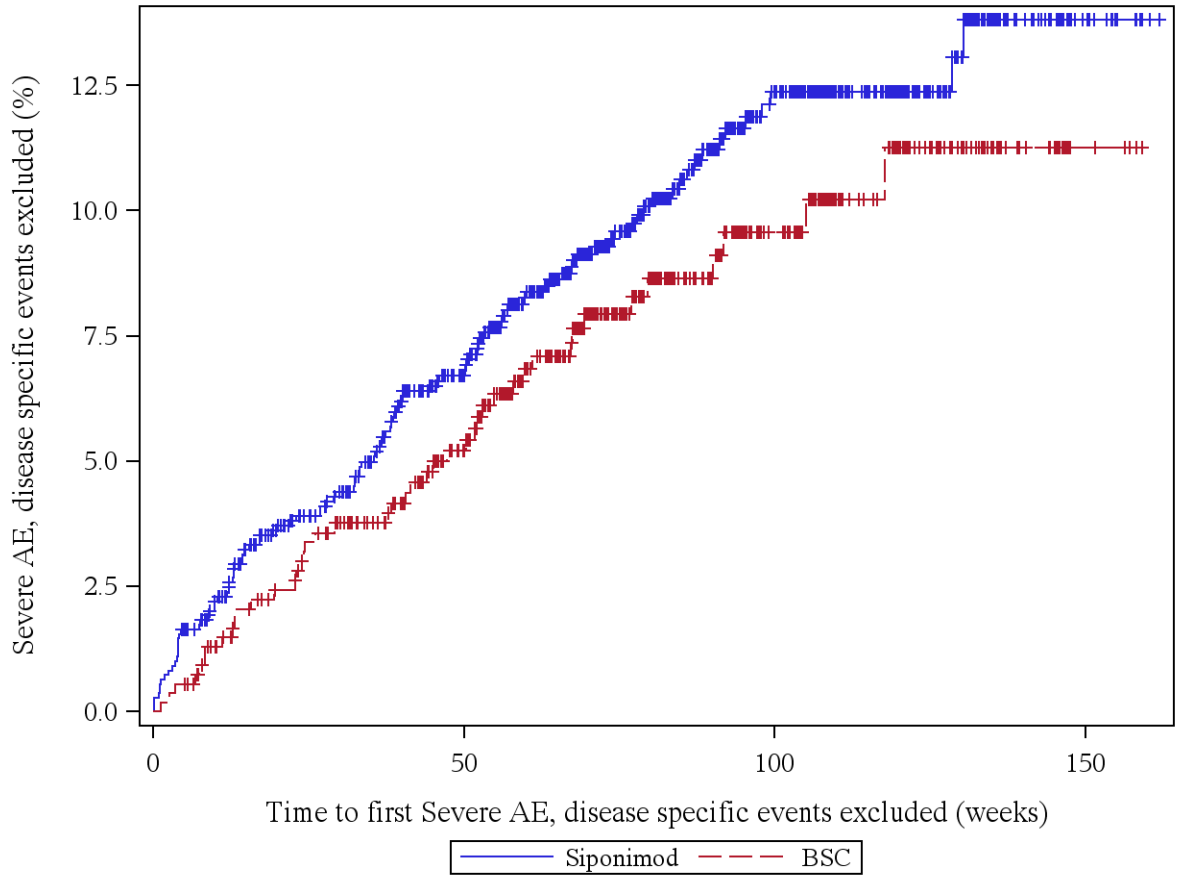
Patients at risk (Siponimod/BSC): Week 0: 1099/546, Week 4: 1081/542, Week 12: 1036/526, Week 24: 975/503, Week 36: 935/477, Week 48: 870/435, Week 60: 755/363, Week 72: 588/283, Week 84: 455/217, Week 96: 352/162, Week 108: 252/114, Week 120: 171/73, Week 132: 95/39, Week 144: 44/20, Week 156: 8/5, Week 168: 0/0

Abbildung 23: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)



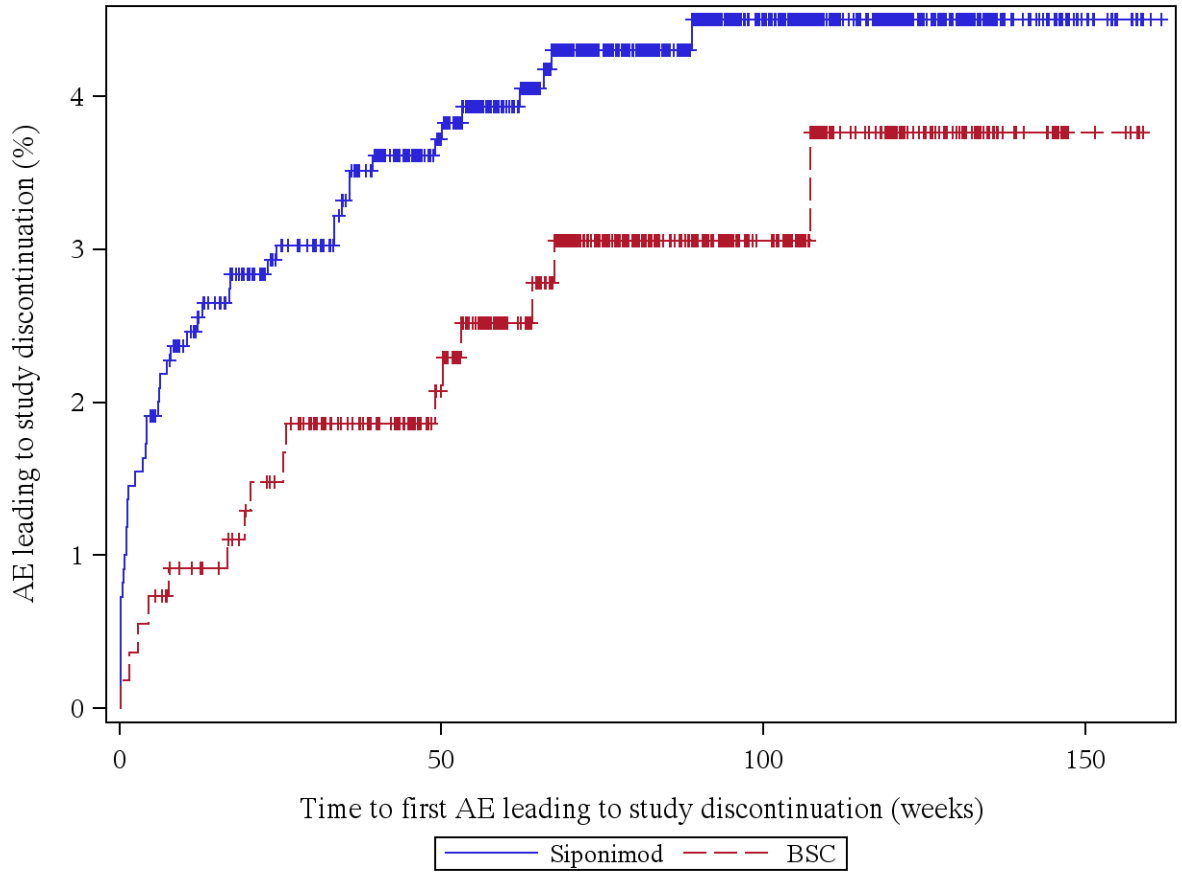
Patients at risk (Siponimod/BSC): Week 0: 1099/546, Week 4: 1079/542, Week 12: 1028/525, Week 24: 974/504, Week 36: 934/476, Week 48: 863/433, Week 60: 753/361, Week 72: 594/281, Week 84: 462/216, Week 96: 356/161, Week 108: 258/112, Week 120: 179/70, Week 132: 99/35, Week 144: 45/17, Week 156: 8/4, Week 168: 0/0

Abbildung 24: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten schweren UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)



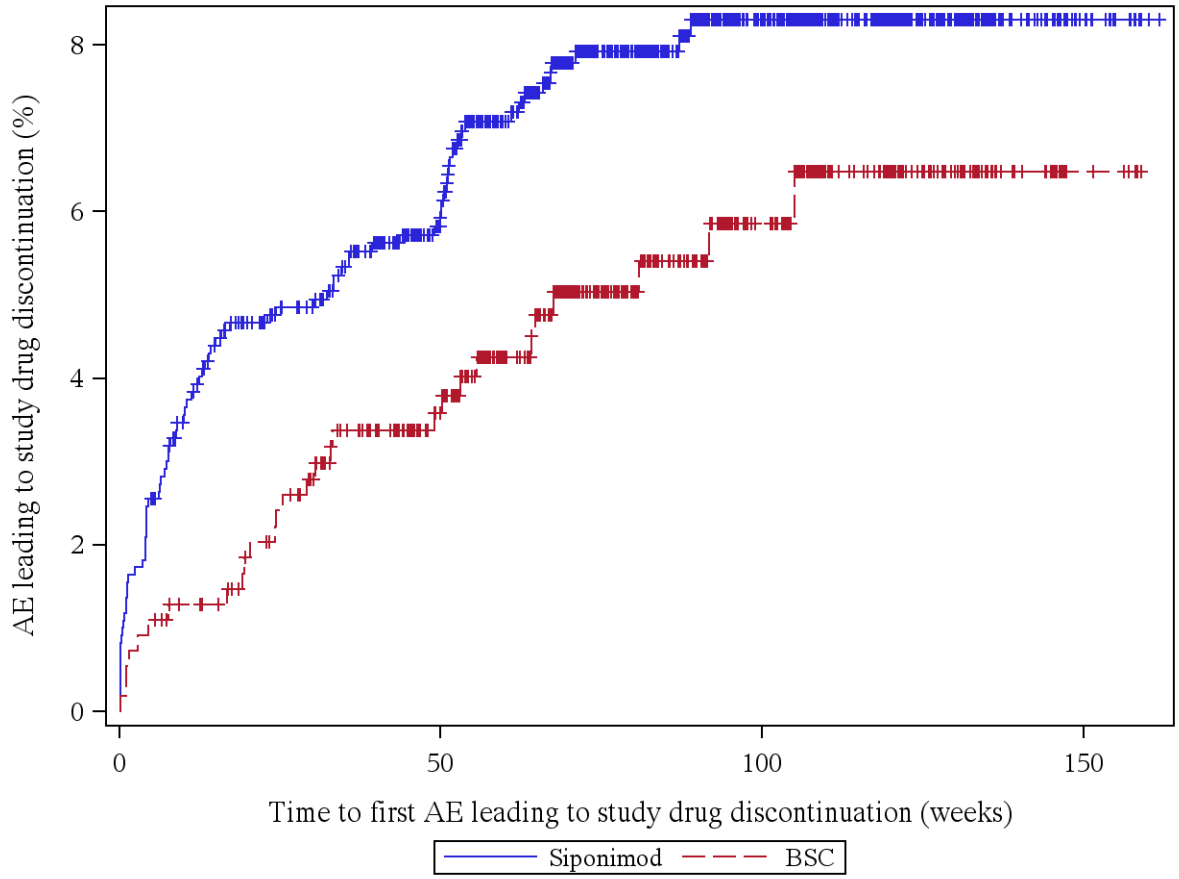
Patients at risk (Siponimod/BSC): Week 0: 1099/546, Week 4: 1083/543, Week 12: 1041/527, Week 24: 989/507, Week 36: 952/482, Week 48: 883/440, Week 60: 770/370, Week 72: 608/287, Week 84: 475/222, Week 96: 365/164, Week 108: 265/115, Week 120: 183/73, Week 132: 101/38, Week 144: 45/19, Week 156: 8/4, Week 168: 0/0

Abbildung 25: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten schweren UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)



Patients at risk (Siponimod/BSC): Week 0: 1099/546, Week 4: 1080/543, Week 12: 1051/534, Week 24: 1012/519, Week 36: 982/497, Week 48: 922/461, Week 60: 816/387, Week 72: 650/304, Week 84: 515/237, Week 96: 400/176, Week 108: 291/125, Week 120: 200/80, Week 132: 113/40, Week 144: 51/20, Week 156: 10/5, Week 168: 0/0

Abbildung 26: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum Studienabbruch aufgrund von UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)



Patients at risk (Siponimod/BSC): Week 0: 1099/546, Week 4: 1076/541, Week 12: 1044/534, Week 24: 1008/519, Week 36: 980/493, Week 48: 921/459, Week 60: 813/386, Week 72: 646/304, Week 84: 514/236, Week 96: 400/175, Week 108: 291/125, Week 120: 200/80, Week 132: 113/40, Week 144: 51/20, Week 156: 10/5, Week 168: 0/0

Abbildung 27: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum Behandlungsabbruch aufgrund von UE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

Tabelle 4-109: Ergebnisse für den Endpunkt *Zeit bis zum Unerwünschten Ereignis* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**HR [95%-KI][b]**	**p-Wert**
N'	127	61		
UE				
Anzahl der Ereignisse	114	47		
Median (in Wochen) [95%-KI][a]	3,6 [2,4; 8,4]	9,7 [4,3; 22,4]		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	7,50 [1,99; 13,02]	21,47 [10,90; 32,04]		
			1,48 [1,05; 2,08]	0,024
UE ohne krankheitsspezifische Ereignisse				
Anzahl der Ereignisse	108	46		
Median (in Wochen) [95%-KI][a]	11,0 [6,4; 13,7]	13,1 [7,3; 25,0]		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	10,03 [3,69; 16,38]	17,26 [4,17; 30,36]		
			1,31 [0,93; 1,86]	0,122
SUE				
Anzahl der Ereignisse	16	11		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	84,56 [77,19; 91,93]	78,99 [66,75; 91,22]		
			0,64 [0,30; 1,38]	0,253
SUE ohne krankheitsspezifische Ereignisse				
Anzahl der Ereignisse	12	8		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	89,00 [82,90; 95,10]	83,60 [71,81; 95,38]		
			0,67 [0,27; 1,64]	0,382
Schwere UE (CTCAE-Grad ≥ 3)				
Anzahl der Ereignisse	11	10		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	89,27 [83,06; 95,48]	80,61 [68,74; 92,49]		
			0,48 [0,20; 1,13]	0,093

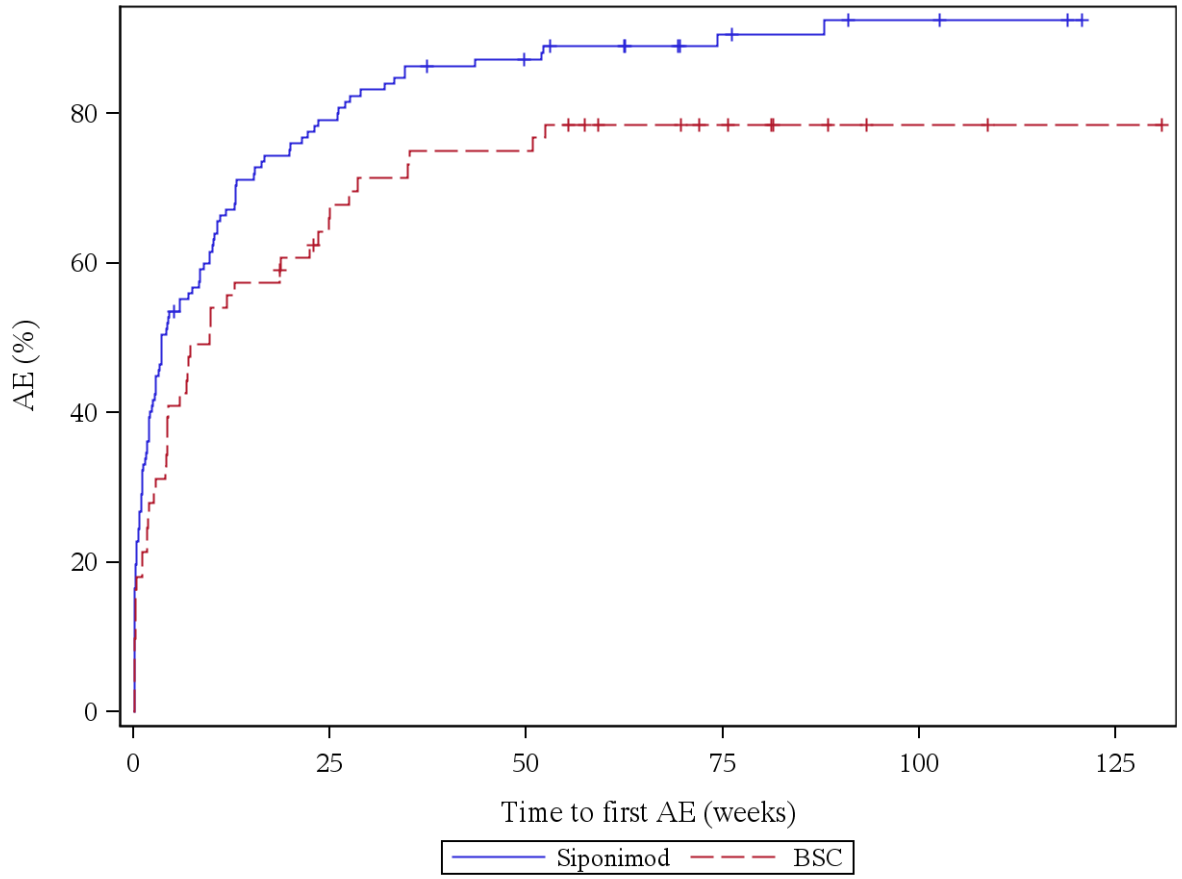
Schwere UE ohne krankheitsspezifische Ereignisse				
Anzahl der Ereignisse	9	8		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	90,89 [85,00; 96,77]	83,58 [72,00; 95,15]		
			0,48 [0,19; 1,25]	0,134
Studienabbrüche aufgrund von UE				
Anzahl der Ereignisse	2	3		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	98,41 [96,23; 100,0]	94,66 [88,76; 100,0]		
			0,32 [0,05; 1,90]	0,208
Behandlungsabbrüche aufgrund von UE				
Anzahl der Ereignisse	6	4		
Median (in Wochen) [95%-KI][a]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][a]	94,34 [89,78; 98,91]	91,16 [82,34; 99,98]		
			0,67 [0,19; 2,39]	0,539

CTCAE: Common Terminology Criteria for Adverse Events ; HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar; UE: Unerwünschtes Ereignis; SUE: Schwerwiegendes Unerwünschtes Ereignis.

a: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

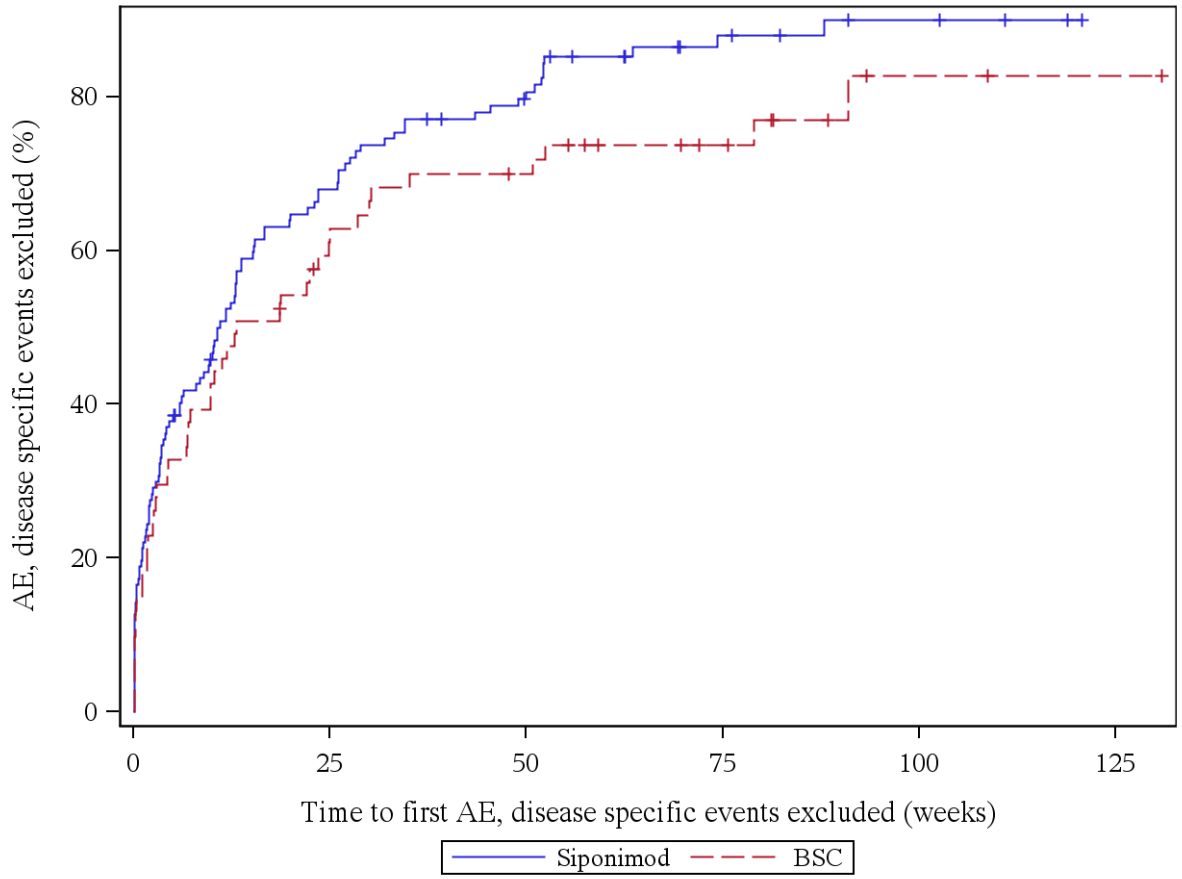
b: Cox proportional hazards model.

Quelle: Anhang 2 zu Modul 4, Tabelle 12.1.1.



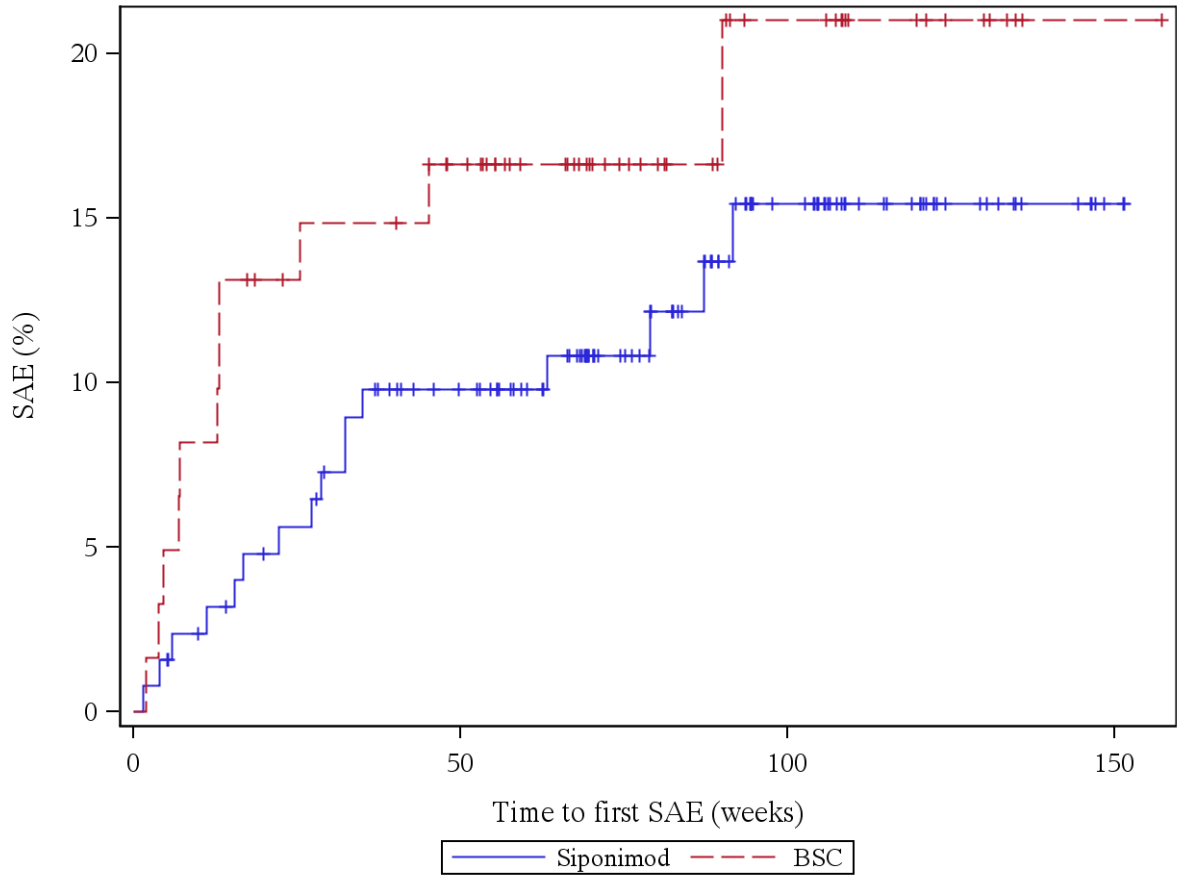
Patients at risk (Siponimod/BSC): Week 0: 127/61, Week 4: 63/41, Week 12: 41/27, Week 24: 26/20, Week 36: 17/14, Week 48: 15/14, Week 60: 11/9, Week 72: 7/7, Week 84: 5/4, Week 96: 3/2, Week 108: 2/2, Week 120: 1/1, Week 132: 0/0

Abbildung 28: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)



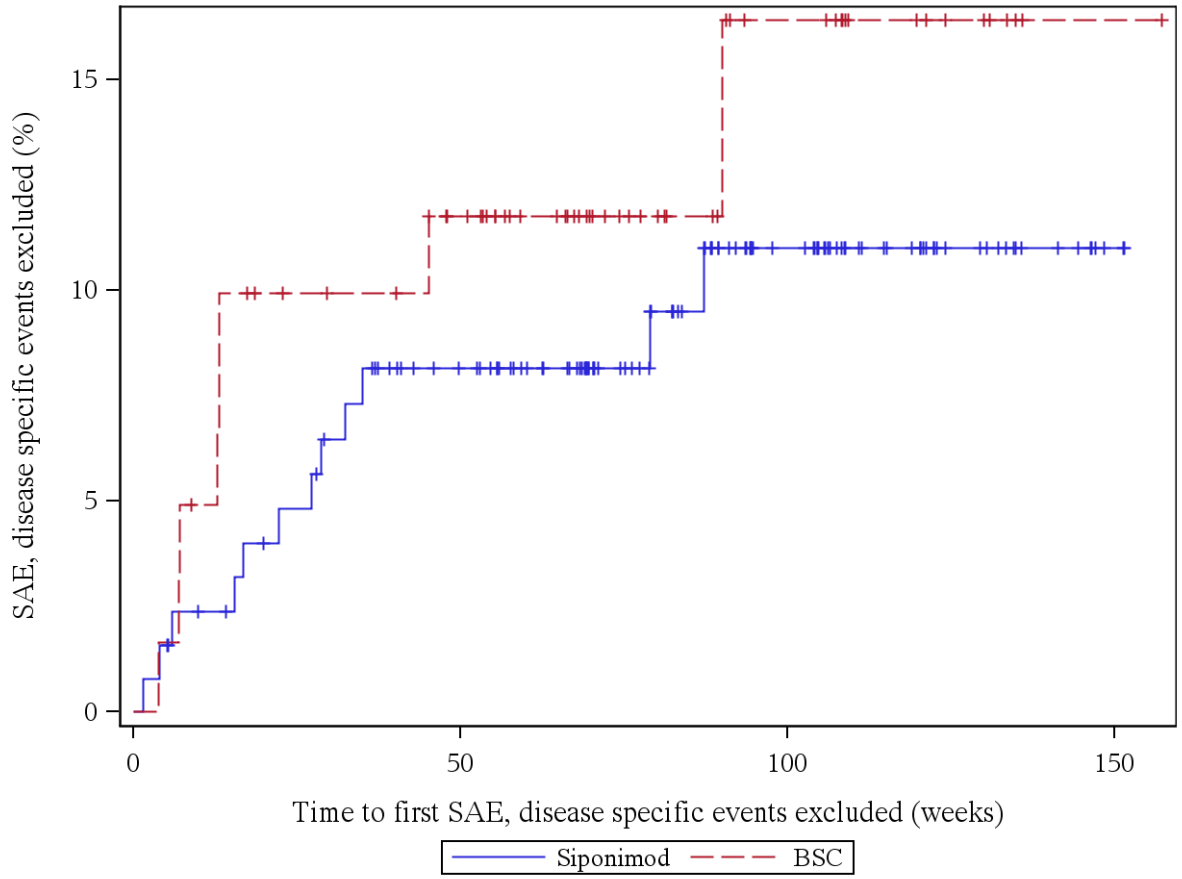
Patients at risk (Siponimod/BSC): Week 0: 127/61, Week 4: 81/43, Week 12: 58/32, Week 24: 39/23, Week 36: 28/17, Week 48: 24/16, Week 60: 14/11, Week 72: 9/9, Week 84: 6/5, Week 96: 4/2, Week 108: 3/2, Week 120: 1/1, Week 132: 0/0

Abbildung 29: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)



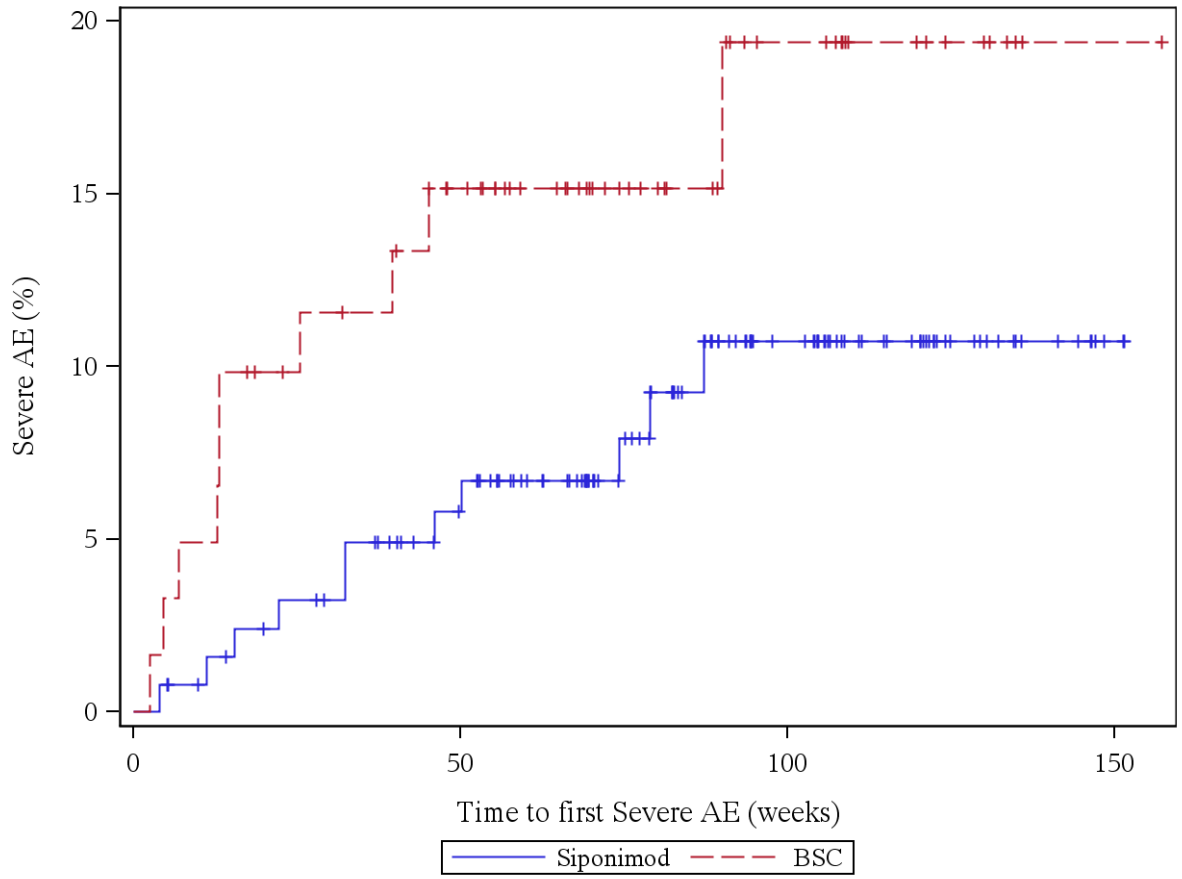
Patients at risk (Siponimod/BSC): Week 0: 127/61, Week 4: 125/59, Week 12: 120/56, Week 24: 115/50, Week 36: 108/49, Week 48: 101/44, Week 60: 91/35, Week 72: 71/27, Week 84: 58/21, Week 96: 40/15, Week 108: 28/13, Week 120: 21/8, Week 132: 11/4, Week 144: 7/1, Week 156: 0/1

Abbildung 30: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)



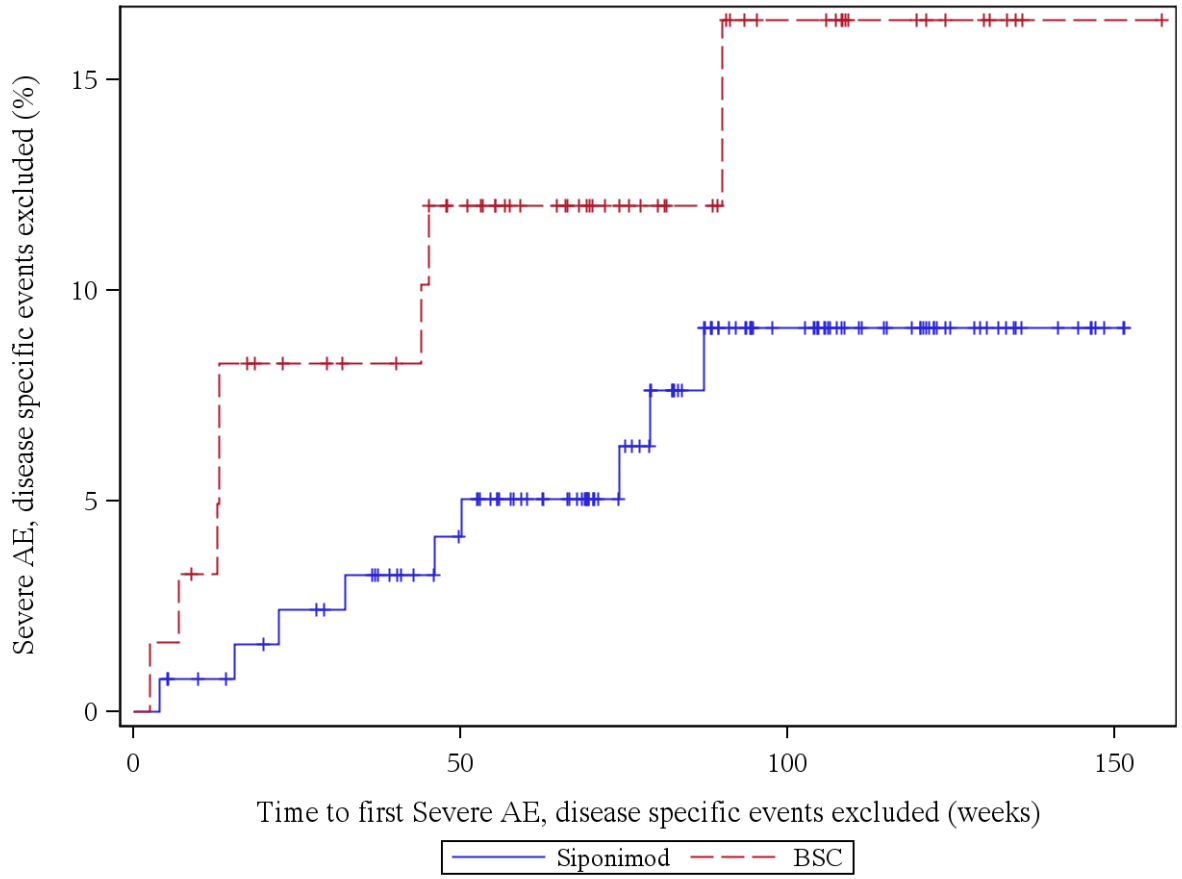
Patients at risk (Siponimod/BSC): Week 0: 127/61, Week 4: 125/60, Week 12: 121/57, Week 24: 116/51, Week 36: 110/50, Week 48: 102/45, Week 60: 92/36, Week 72: 73/27, Week 84: 60/21, Week 96: 43/15, Week 108: 31/13, Week 120: 23/8, Week 132: 13/4, Week 144: 7/1, Week 156: 0/1

Abbildung 31: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)



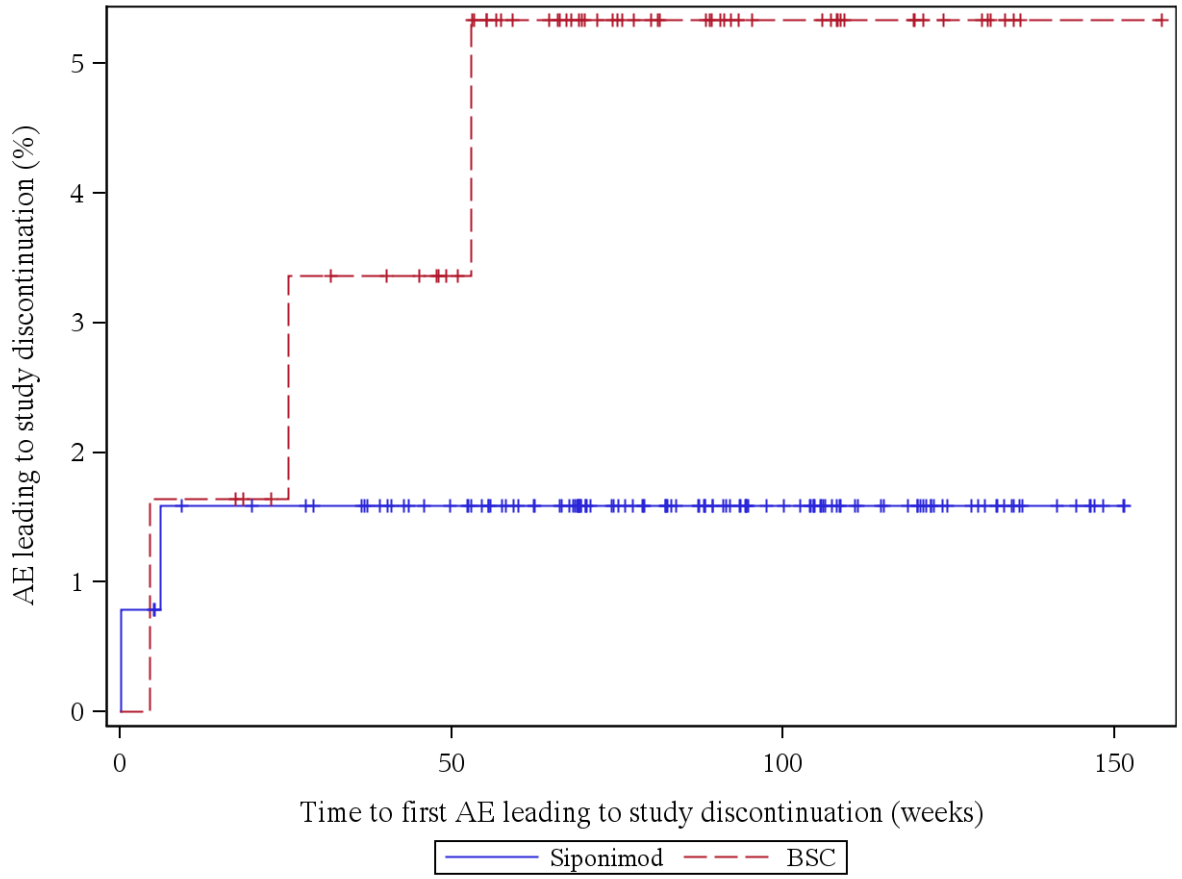
Patients at risk (Siponimod/BSC): Week 0: 127/61, Week 4: 126/60, Week 12: 122/58, Week 24: 118/52, Week 36: 114/50, Week 48: 106/44, Week 60: 94/36, Week 72: 76/28, Week 84: 61/22, Week 96: 44/15, Week 108: 32/13, Week 120: 25/8, Week 132: 12/4, Week 144: 7/1, Week 156: 0/1

Abbildung 32: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten schweren UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)



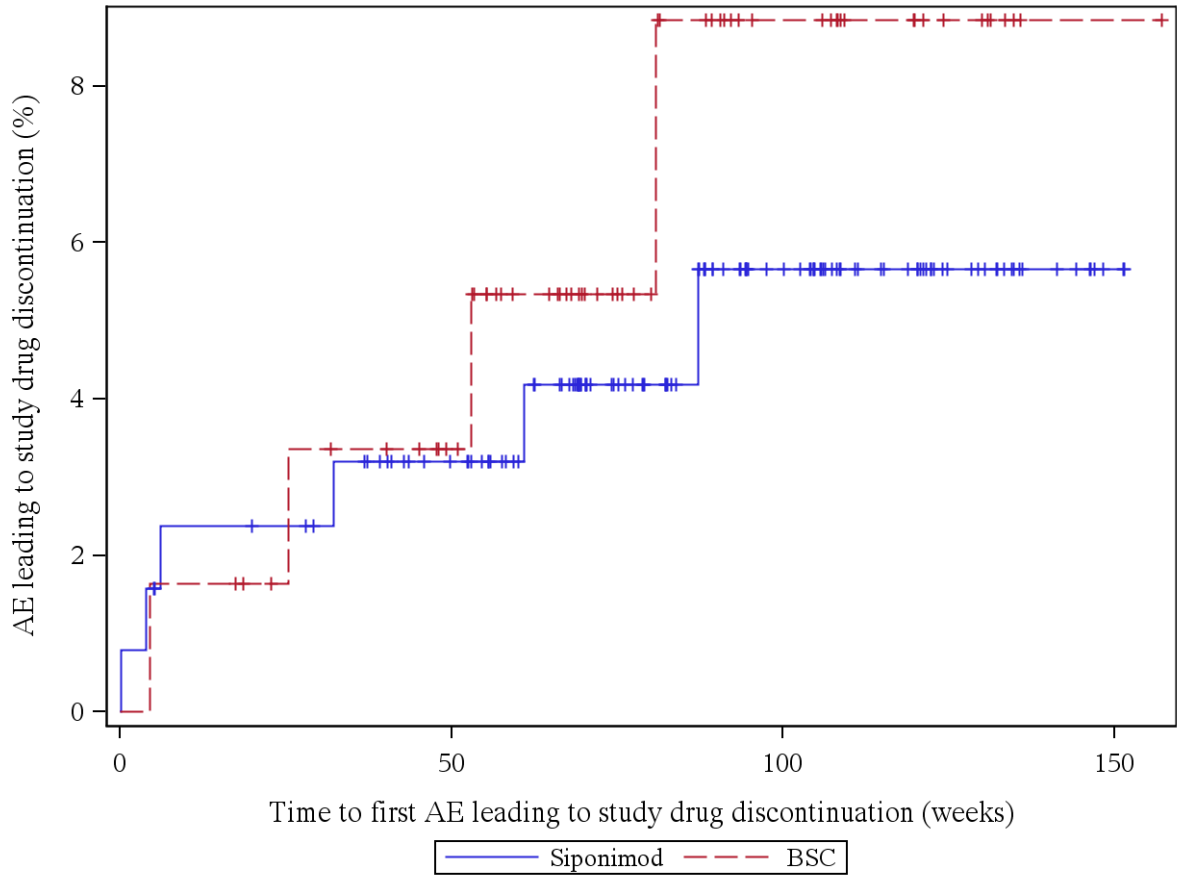
Patients at risk (Siponimod/BSC): Week 0: 127/61, Week 4: 126/60, Week 12: 123/58, Week 24: 119/52, Week 36: 116/50, Week 48: 107/44, Week 60: 95/36, Week 72: 77/28, Week 84: 62/22, Week 96: 45/15, Week 108: 33/13, Week 120: 26/8, Week 132: 13/4, Week 144: 7/1, Week 156: 0/1

Abbildung 33: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten schweren UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)



Patients at risk (Siponimod/BSC): Week 0: 127/61, Week 4: 126/61, Week 12: 122/60, Week 24: 121/57, Week 36: 119/55, Week 48: 110/51, Week 60: 99/41, Week 72: 80/32, Week 84: 66/25, Week 96: 49/17, Week 108: 36/15, Week 120: 28/9, Week 132: 15/4, Week 144: 7/1, Week 156: 0/1

Abbildung 34: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum Studienabbruch aufgrund von UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)



Patients at risk (Siponimod/BSC): Week 0: 127/61, Week 4: 125/61, Week 12: 122/60, Week 24: 121/57, Week 36: 118/55, Week 48: 110/51, Week 60: 99/41, Week 72: 79/32, Week 84: 65/24, Week 96: 49/17, Week 108: 36/15, Week 120: 28/9, Week 132: 15/4, Week 144: 7/1, Week 156: 0/1

Abbildung 35: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum Behandlungsabbruch aufgrund von UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

Die Ereigniszeitanalysen ergaben in der Gesamtpopulation ein statistisch signifikant höheres Risiko für UE (HR = 1,26; 95 %-KI [1,13; 1,41]; p < 0,001) und für UE ohne krankheitsspezifische Ereignisse (HR = 1,22; 95%-KI [1,09; 1,37]; p = 0,001) unter Siponimod.

In der Subpopulation B ergaben die Ereigniszeitanalysen ein statistisch signifikant höheres Risiko für UE (HR = 1,48; 95 %-KI [1,05; 2,08] p = 0,024) unter Siponimod.

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

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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.

4.3.1.3.12 Suizidalität (C-SSRS) – RCT

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

Tabelle 4-110: Operationalisierung des Endpunkts *Suizidalität (C-SSRS)*

Studie	Operationalisierung
EXPAND	Der Endpunkt Suizidalität wurde durch die Columbia-Beurteilungsskala (C-SSRS) zur Suizidalität erhoben. Die C-SSRS ist ein validierter Fragebogen zur Erfassung suizidaler Gedanken, suizidalen Verhaltens und von Selbstverletzung ohne suizidale Absicht (32). Der Bogen besteht aus fünf Fragenkategorien zu suizidalen Gedanken (Kategorien 1 bis 5) und sieben Fragenkategorien zu suizidalem Verhalten (Kategorien 6 bis 10) sowie einer Abfrage zum selbstverletzenden Verhalten ohne suizidale Absicht. Im Dossier berichtet werden: • Inzidenz von suizidalen Gedanken (Kategorien 1 bis 5 insgesamt und jede Kategorie separat); • Inzidenz von suizidalem Verhalten (Kategorien 6 bis 10 insgesamt und jede Kategorie separat); • Inzidenz von suizidalen Gedanken oder suizidalem Verhalten (Kategorien 1 bis 10 insgesamt); • Inzidenz von selbstverletzendem Verhalten ohne suizidale Absicht. Neben der jeweiligen Inzidenz werden auch jeweils die Risikoschätzer OR, RR und RD mit den jeweils dazugehörigen Konfidenzintervallen und p-Werten berichtet. Für den C-SSRS wurde nicht dokumentiert, zu welcher Visite der Bogen ausgefüllt wurde, sondern lediglich, ob die Auswertung vor oder nach der ersten Dosis der Studienmedikation erfolgte. Ausgewertet wurden in Übereinstimmung mit dem Studienbericht Bögen, die nach der ersten Dosis der Studienmedikation ausgefüllt wurden. Wenn für einen Patienten hier mehrere

	Fragebögen vorlagen, wurde in einem konservativen Ansatz jeweils die höchste Antwortkategorie verwendet. Die Anzahl der in die Analysen eingegangenen Patienten ist unter N' in den Tabellen aufgeführt.

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 4-111: Bewertung des Verzerrungspotenzials für den Endpunkt *Suizidalität (C-SSRS)* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Für die EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Suizidalität (C-SSRS)* als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkterheber (unabhängiger Bewerter) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde, eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war und sonstige das Verzerrungspotenzial beeinflussende Punkte nicht festgestellt wurden.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden und es ist nur mit einem niedrigen Verzerrungspotenzial zu rechnen.

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

Tabelle 4-112: Ergebnisse für den Endpunkt *Suizidalität (C-SSRS)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 1.099)	**BSC (N = 546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Suizidalität (C-SSRS)					
N'	1.079	542			
Suizidgedanken (Kat. 1 bis 5) – n (%)	121 (11,2)	58 (10,7)	1,05 [0,76; 1,47] 0,756	1,05 [0,78; 1,41] 0,756	0,01 [-0,03; 0,04] 0,754
Kat. 1: Wunsch, tot zu sein – n (%)	115 (10,7)	56 (10,3)	1,04 [0,74; 1,45] 0,840	1,03 [0,76; 1,40] 0,840	0,00 [-0,03; 0,03] 0,840
Kat. 2: Unspezifische, aktive Suizid-gedanken – n (%)	36 (3,3)	14 (2,6)	1,30 [0,70; 2,43] 0,408	1,29 [0,70; 2,37] 0,410	0,01 [-0,01; 0,02] 0,388
Kat. 3: Aktive Suizidgedanken (ohne Vorsatz) – n (%)	24 (2,2)	7 (1,3)	1,74 [0,74; 4,06] 0,196	1,72 [0,75; 3,97] 0,202	0,01 [0,00; 0,02] 0,158
Kat. 4: Aktive Suizidgedanken (mit Vorsatz) – n (%)	15 (1,4)	2 (0,4)	3,81 [0,87; 16,71] 0,057	3,77 [0,86; 16,42] 0,077	0,01 [0,00; 0,02] 0,021
Kat. 5: Aktive Suizidgedanken (Plan/Vorsatz) – n (%)	5 (0,5)	0 (0,0)	5,55 [0,31; 100,62] 0,112	5,53 [0,31; 99,84] 0,247	0,00 [0,00; 0,01] 0,025
Suizidales Verhalten (Kat. 6 bis 10) – n (%)	25 (2,3)	7 (1,3)	1,81 [0,78; 4,22] 0,161	1,79 [0,78; 4,12] 0,168	0,01 [0,00; 0,02] 0,124
Kat. 6: Vorbereitende Handlungen – n (%)	9 (0,8)	0 (0,0)	9,63 [0,56; 165,74] 0,033	9,55 [0,56; 163,83] 0,120	0,01 [0,00; 0,01] 0,003
Kat. 7: Abgebrochener Suizidversuch – n (%)	12 (1,1)	4 (0,7)	1,51 [0,49; 4,71] 0,472	1,51 [0,49; 4,65] 0,476	0,00 [-0,01; 0,01] 0,442
Kat. 8: Unterbrochener Suizidversuch – n (%)	11 (1,0)	2 (0,4)	2,78 [0,61; 12,59] 0,166	2,76 [0,61; 12,42] 0,185	0,01 [0,00; 0,01] 0,105
Kat. 9: Suizidversuch – n (%)	4 (0,4)	2 (0,4)	1,00 [0,18; 5,50] 0,996	1,00 [0,18; 5,47] 0,996	0,00 [-0,01; 0,01] 0,996
Kat. 10: Suizid – n (%)	1 (0,1)	0 (0,0)	1,51 [0,06; 37,10] 0,478	1,51 [0,06; 36,97] 0,801	0,00 [0,00; 0,00] 0,317
Suizidgedanken / suizidales Verhalten (Kat. 1 bis 10) – n (%)	128 (11,9)	61 (11,3)	1,06 [0,77; 1,47] 0,719	1,05 [0,79; 1,40] 0,719	0,01 [-0,03; 0,04] 0,717

Selbstverletzung ohne Suizidabsicht – n (%)	3 (0,3)	1 (0,2)	1,51 [0,16; 14,53] 0,720	1,51 [0,16; 14,45] 0,722	0,00 [0,00; 0,01] 0,702

C-SSRS: Columbia-Suicide Severity Rating Scale; Kat: Kategorie; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 1 zu Modul 4, Tabelle 13.1.

Tabelle 4-113: Ergebnisse für den Endpunkt *Suizidalität (C-SSRS)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N = 61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Suizidalität (C-SSRS)					
N'	125	61			
Suizidgedanken (Kat. 1 bis 5) – n (%)	20 (16,0)	6 (9,8)	1,75 [0,66; 4,60] 0,255	1,63 [0,69; 3,84] 0,267	0,06 [-0,04; 0,16] 0,220
Kat. 1: Wunsch, tot zu sein – n (%)	20 (16,0)	5 (8,2)	2,13 [0,76; 5,99] 0,143	1,95 [0,77; 4,95] 0,159	0,08 [-0,02; 0,17] 0,104
Kat. 2: Unspezifische, aktive Suizid-gedanken – n (%)	6 (4,8)	3 (4,9)	0,97 [0,24; 4,04] 0,972	0,98 [0,25; 3,77] 0,972	0,00 [-0,07; 0,06] 0,972
Kat. 3: Aktive Suizidgedanken (ohne Vorsatz) – n (%)	5 (4,0)	2 (3,3)	1,23 [0,23; 6,52] 0,808	1,22 [0,24; 6,11] 0,809	0,01 [-0,05; 0,06] 0,802
Kat. 4: Aktive Suizidgedanken (mit Vorsatz) – n (%)	1 (0,8)	1 (1,6)	0,48 [0,03; 7,87] 0,602	0,49 [0,03; 7,67] 0,610	-0,01 [-0,04; 0,03] 0,643
Kat. 5: Aktive Suizidgedanken (Plan/Vorsatz) – n (%)	0 (0,0)	0 (0,0)	-	-	-
Suizidales Verhalten (Kat. 6 bis 10) – n (%)	1 (0,8)	0 (0,0)	1,48 [0,06; 36,91] 0,484	1,48 [0,06; 35,72] 0,811	0,01 [-0,01; 0,02] 0,315
Kat. 6: Vorbereitende Handlungen – n (%)	1 (0,8)	0 (0,0)	1,48 [0,06; 36,91] 0,484	1,48 [0,06; 35,72] 0,811	0,01 [-0,01; 0,02] 0,315
Kat. 7: Abgebrochener Suizidversuch – n (%)	1 (0,8)	0 (0,0)	1,48 [0,06; 36,91] 0,484	1,48 [0,06; 35,72] 0,811	0,01 [-0,01; 0,02] 0,315

Kat. 8: Unterbrochener Suizidversuch – n (%)	1 (0,8)	0 (0,0)	1,48 [0,06; 36,91] 0,484	1,48 [0,06; 35,72] 0,811	0,01 [-0,01; 0,02] 0,315
Kat. 9: Suizidversuch – n (%)	0 (0,0)	0 (0,0)	-	-	-
Kat. 10: Suizid	0 (0,0)	0 (0,0)	-	-	-
Suizidgedanken / suizidales Verhalten (Kat. 1 bis 10) – n (%)	20 (16,0)	6 (9,8)	1,75 [0,66; 4,60] 0,255	1,63 [0,69; 3,84] 0,267	0,06 [-0,04; 0,16] 0,220
Selbstverletzung ohne Suizidabsicht – n (%)	0 (0,0)	0 (0,0)	-	-	-

C-SSRS: Columbia-Suicide Severity Rating Scale; Kat: Kategorie; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

Quelle: Anhang 2 zu Modul 4, Tabelle 13.1.

Keine der untersuchten Kategorien zur Suizidalität ergab einen statistisch signifikanten Unterschied zwischen den Behandlungsgruppen im relativen Risiko.

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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.

4.3.1.3.13 Läsionslast – RCT

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

Tabelle 4-114: Operationalisierung des Endpunkts *Läsionslast*

Studie	Operationalisierung
EXPAND	Die Läsionslast wird anhand der im MRT festgestellten T2-Läsionen sowie der Gd-anreichernden T1-Läsionen dargestellt. MRT-Untersuchungen erfolgten zu Baseline und dann einmal jährlich. Im Dossier berichtet werden: • die Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn im T2-Läsionsvolumen (dargestellt werden die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den zugehörigen Standardabweichungen sowie die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht); • die Anzahl neuer oder sich vergrößernder T2-Läsionen zu Monat 12 bzw. 24 mit Angabe des Rate Ratios mit den dazugehörigen Konfidenzintervallen und p-Werten, basierend auf einem negativ-binomialen Regressionsmodell; • die Anzahl Gd-anreichernder T1-Läsionen zu Monat 12 bzw. 24 mit Angabe des Rate Ratios mit den dazugehörigen Konfidenzintervallen und p-Werten, basierend auf einem negativ-binomialen Regressionsmodell.

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 4-115: Bewertung des Verzerrungspotenzials für den Endpunkt *Läsionslast* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Läsionslast* als niedrig bewertet, da das Verzerrungspotenzial der Studie insgesamt als niedrig bewertet wird,

die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen. Der Nichtberücksichtigungsanteil liegt für die Auswertung zu Monat 12 insgesamt unter 30 %, der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen unter 10 %. Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden und für die Auswertung zu Monat 12 ist nur mit einem niedrigen Verzerrungspotenzial zu rechnen.

Tabelle 4-116: Rücklaufquoten *Läsionslast* (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
T2-Läsionsvolumen			
Rückläufer Studienbeginn, n (%)	1.098 (99,9)	545 (99,8)	1.643 (99,9)
Rückläufer Monat 12, n (%)	1.024 (93,2)	509 (93,2)	1.533 (93,2)
Rückläufer Monat 24, n (%)	626 (57,0)	305 (55,9)	931 (56,6)
Rückläufer Monat 36, n (%)	133 (12,1)	53 (9,7)	186 (11,3)
Anzahl neuer oder sich vergrößernder T2-Läsionen			
Rückläufer Monat 12, n (%)	1.023 (93,1)	509 (93,2)	1.532 (93,1)
Rückläufer Monat 24, n (%)	626 (57,0)	304 (55,7)	930 (56,5)
Anzahl Gd-anreichernder T1-Läsionen			
Rückläufer Monat 12, n (%)	1.019 (92,7)	507 (92,9)	1.526 (92,8)
Rückläufer Monat 24, n (%)	622 (56,6)	304 (55,7)	926 (56,3)

N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 1 zu Modul 4, Tabellen 9.1.0, 9.2.0 und 9.3.0.

Tabelle 4-117: Rücklaufquoten *Läsionslast* (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
T2-Läsionsvolumen			
Rückläufer Studienbeginn, n (%)	127 (100,0)	61 (100,0)	188 (100,0)
Rückläufer Monat 12, n (%)	118 (92,9)	56 (91,8)	174 (92,6)
Rückläufer Monat 24, n (%)	81 (63,8)	33 (54,1)	114 (60,6)
Rückläufer Monat 36, n (%)	18 (14,2)	5 (8,2)	23 (12,2)
Anzahl neuer oder sich vergrößernder T2-Läsionen			
Rückläufer Monat 12, n (%)	118 (92,9)	56 (91,8)	174 (92,6)
Rückläufer Monat 24, n (%)	81 (63,8)	33 (54,1)	114 (60,6)
Anzahl Gd-anreichernder T1-Läsionen			
Rückläufer Monat 12, n (%)	118 (92,9)	56 (91,8)	174 (92,6)
Rückläufer Monat 24, n (%)	80 (63,0)	33 (54,1)	113 (60,1)

N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 2 zu Modul 4, Tabellen 9.1.0, 9.2.0 und 9.3.0.

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

Tabelle 4-118: Ergebnisse für den Endpunkt *Läsionslast (T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
Veränderung im T2-Läsionsvolumen seit Studienbeginn				
N'	995	495		
Mittelwert zu Studienbeginn (SD)	15.551,74 (16.319,94)	14.285,38 (15.438,57)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	204,94 (67,47)	818,03 (87,29)	-613,09 [-800,20; -425,98]	< 0,001
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	162,90 (73,90)	940,41 (97,20)	-777,52 [-990,63; -564,40]	< 0,001

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
a: Negative Werte zeigen eine Verbesserung an. Quelle: Anhang 1 zu Modul 4, Tabelle 9.1.1.				

Tabelle 4-119: Ergebnisse für den Endpunkt *Läsionslast (T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
Veränderung im T2-Läsionsvolumen seit Studienbeginn				
N'	119	56		
Mittelwert zu Studienbeginn (SD)	21.244,56 (17.083,27)	20.928,18 (18.321,47)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-24,02 (231,83)	953,81 (303,79)	-977,83 [-1.604,94; -350,73]	0,002
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-224,07 (240,63)	1.379,89 (324,06)	-1.603,96 [-2.284,99; -922,93]	< 0,001
KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error. a: Negative Werte zeigen eine Verbesserung an. Quelle: Anhang 2 zu Modul 4, Tabelle 9.1.1.				

Tabelle 4-120: Ergebnisse für den Endpunkt *Läsionslast (Anzahl neuer bzw. sich vergrößernder T2-Läsionen und Gd-anreichernder T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Rate Ratio [95%-KI]**	**p-Wert**
Anzahl neuer bzw. sich vergrößernder T2-Läsionen				
N'	997	496		
Monat 12: Anzahl Läsionen [95%-KI]	1,00 [0,86;1,17]	3,78 [3,15;4,53]	0,27 [0,21; 0,33]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,49 [0,37;0,64]	3,44 [2,80; 4,22]	0,14 [0,10; 0,20]	< 0,001
Anzahl Gd-anreichernder T1-Läsionen				
N'	996	496		
Monat 12: Anzahl Läsionen [95%-KI]	0,08 [0,06; 0,11]	0,64 [0,49; 0,84]	0,13 [0,08; 0,19]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,07 [0,04; 0,14]	0,42 [0,29; 0,61]	0,18 [0,09; 0,36]	< 0,001

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
Quelle: Anhang 1 zu Modul 4, Tabellen 9.2.1 und 9.3.1.

Tabelle 4-121: Ergebnisse für den Endpunkt *Läsionslast (Anzahl neuer bzw. sich vergrößernder T2-Läsionen und Gd-anreichernder T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Rate Ratio [95%-KI]**	**p-Wert**
Anzahl neuer bzw. sich vergrößernder T2-Läsionen				
N'	119	56		
Monat 12: Anzahl Läsionen [95%-KI]	3,33 [2,14; 5,18]	9,61 [7,04; 13,11]	0,35 [0,20; 0,60]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,69 [0,31; 1,51]	8,09 [5,18; 12,64]	0,08 [0,03; 0,21]	< 0,001
Anzahl Gd-anreichernder T1-Läsionen				
N'	119	56		
Monat 12: Anzahl Läsionen [95%-KI]	0,27 [0,12; 0,57]	1,76 [1,10; 2,81]	0,15 [0,06; 0,38]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,32 [0,10; 1,05]	1,05 [0,67;1,65]	0,30 [0,08; 1,10]	0,069

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
Quelle: Anhang 2 zu Modul 4, Tabellen 9.2.1 und 9.3.1.

Die Auswertung der Läsionslast zeigt einen statistisch signifikanten Vorteil von Siponimod in der Gesamtpopulation und in der Subpopulation B:

- Unter Siponimod nahm das T2-Läsionsvolumen weniger zu als unter BSC. Der Unterschied zwischen den Behandlungsarmen war in beiden Populationen und zu beiden Auswertungszeitpunkten statistisch signifikant. Die Auswertung zu Monat 24 ist wegen hoher Nichtberücksichtigungsanteile jedoch nicht zuverlässig interpretierbar (Tabelle 4-118 und Tabelle 4-119).

- Unter Siponimod bildeten sich weniger neue bzw. sich vergrößernde T2-Läsionen und Gd-anreichernde T1-Läsionen. Auch hier war der Unterschied zwischen den Behandlungsarmen in beiden Populationen und zu beiden Auswertungszeitpunkten statistisch signifikant. Einzige Ausnahme bildet hier die Anzahl Gd-anreichernder T1-Läsionen in der Subpopulation B zu Monat 24, die zwar numerisch zugunsten von Siponimod ausfällt, aber ohne die statistische Signifikanz zu erreichen. Die Auswertung zu Monat 24 ist allerdings wegen hoher Nichtberücksichtigungsanteile nicht zuverlässig interpretierbar (Tabelle 4-120 und Tabelle 4-121).

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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.

4.3.1.3.14 Hirnvolumen – RCT

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

Tabelle 4-122: Operationalisierung des Endpunkts *Hirnvolumen*

Studie	Operationalisierung
EXPAND	Die prozentuale Veränderung des Hirnvolumens wurde anhand von MRT-Aufnahmen festgestellt. Diese MRT-Untersuchungen erfolgten zu Baseline und dann einmal jährlich. Im Dossier berichtet wird: • die prozentuale Veränderung nach 12 bzw. 24 Monaten gegenüber Studienbeginn im Hirnvolumen (dargestellt werden die mittleren Werte zu Studienbeginn, die mittleren Veränderungen mit den zugehörigen Standardabweichungen sowie die Mittelwertdifferenz mit den dazugehörigen Konfidenzintervallen und p-Werten; Auswertung unter Verwendung von MMRM, eine darüber hinausgehende Imputation fehlender Werte erfolgte nicht).

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 4-123: Bewertung des Verzerrungspotenzials für den Endpunkt *Läsionslast* 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
EXPAND	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

In der EXPAND-Studie wird das Verzerrungspotenzial für den Endpunkt *Hirnvolumen* als niedrig bewertet, da das Verzerrungspotenzial der Studie insgesamt als niedrig bewertet wird, die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen. Der Nichtberücksichtigungsanteil liegt für die Auswertung zu Monat 12 insgesamt unter 30 %, der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen unter 10 %. Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %.

Es kann von einem objektivierbaren Endpunkt ausgegangen werden und für die Auswertung zu Monat 12 ist nur mit einem niedrigen Verzerrungspotenzial zu rechnen.

Tabelle 4-124: Rücklaufquoten *Hirnvolumen* (Gesamtpopulation)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Gesamt (N = 1.645)**
Prozentuale Veränderung des Hirnvolumens			
Rückläufer Monat 12, n (%)	903 (82,2)	439 (80,4)	1.342 (81,6)
Rückläufer Monat 24, n (%)	470 (42,8)	239 (43,8)	709 (43,1)
Rückläufer Monat 36, n (%)	80 (7,3)	31 (5,7)	111 (6,7)

N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 1 zu Modul 4, Tabelle 9.4.0.

Tabelle 4-125: Rücklaufquoten *Hirnvolumen* (Subpopulation B)

	Behandlungsgruppen		
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Gesamt (N = 188)**
Prozentuale Veränderung des Hirnvolumens			
Rückläufer Monat 12, n (%)	105 (82,7)	50 (82,0)	155 (82,4)
Rückläufer Monat 24, n (%)	59 (46,5)	26 (42,6)	85 (45,2)
Rückläufer Monat 36, n (%)	8 (6,3)	3 (4,9)	11 (5,9)

N: Patienten gesamt; n: Patienten mit Ereignis.
Quelle: Anhang 2 zu Modul 4, Tabelle 9.4.0.

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

Tabelle 4-126: Ergebnisse für den Endpunkt *Hirnvolumen* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
Prozentuale Veränderung im Hirnvolumen seit Studienbeginn				
N'	894	436		
Mittelwert zu Studienbeginn (SD)	1.425,03 (85,07)	1.424,58 (87,59)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,28 (0,03)	-0,46 (0,03)	0,17 [0,10; 0,25]	< 0,001
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,71 (0,04)	-0,84 (0,05)	0,13 [0,02; 0,24]	0,020

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.
a: Positive Werte zeigen eine Verbesserung an.
Quelle: Anhang 1 zu Modul 4, Tabelle 9.4.1.

Tabelle 4-127: Ergebnisse für den Endpunkt *Hirnvolumen* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Mittelwert-differenz [95%-KI]**	**p-Wert**
Prozentuale Veränderung im Hirnvolumen seit Studienbeginn				
N'	106	52		
Mittelwert zu Studienbeginn (SD)	1.417,21 (89,02)	1.418,54 (88,14)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,41 (0,07)	-0,68 (0,09)	0,27 [0,07; 0,47]	0,009
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,87 (0,09)	-1,24 (0,13)	0,37 [0,08; 0,66]	0,014

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Positive Werte zeigen eine Verbesserung an.

Quelle: Anhang 2 zu Modul 4, Tabelle 9.4.1.

Die Auswertung der prozentualen Veränderung des Hirnvolumens zeigt jeweils einen Vorteil von Siponimod in der Gesamtpopulation und in der Subpopulation B. Unter Siponimod verringerte sich das Hirnvolumen signifikant weniger als unter BSC. Der Unterschied zwischen den Behandlungsarmen war in beiden Populationen und zu beiden Auswertungszeitpunkten statistisch signifikant. Die Auswertung zu Monat 24 ist wegen hoher Nichtberücksichtigungsanteile jedoch nicht zuverlässig interpretierbar (Tabelle 4-126 und Tabelle 4-127).

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

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Die Ergebnisse sind auf den deutschen Versorgungskontext übertragbar, da die gesamte EXPAND-Studie den deutschen Versorgungsalltag widerspiegelt (siehe 4.3.1.2.1). Zudem

liegen auf Endpunktebene kein hohes Verzerrungspotenzial oder andere Gründe vor, die die Übertragbarkeit einschränken.

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.

4.3.1.3.15 Subgruppenanalysen – RCT

Für die Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 0. [16]

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

- Subgruppenanalysen sind nur für die Merkmale (z.B. Alter) durchzuführen, bei denen die resultierenden Subgruppen jeweils mindestens 10 Patienten umfassen.
- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.
- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ($p < 0,05$) dargestellt werden.
- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.
- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.
- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmodifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.
- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die

[16] unbesetzt

Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.

- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

Beschreiben Sie die Ergebnisse von Subgruppenanalysen. Stellen Sie dabei zunächst tabellarisch dar, zu welchen der in Abschnitt 4.2.5.5 genannten Effektmodifikatoren Subgruppenanalysen zu den relevanten Endpunkten vorliegen, und ob diese a priori geplant und im Studienprotokoll festgelegt waren oder posthoc durchgeführt wurden.

Orientieren Sie sich an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4 -128 Matrix der durchgeführten Subgruppenanalysen

Endpunkt Studie	**Alter**	**Geschlecht**	**DMT-Vorbehandlung**	**IFN-beta-1b-Vorbehandlung**	**Schnell fortschreitende Erkrankung**	**Moderater/ schwerer Erkrankungsverlauf**	**Region**
EXPAND							
Mortalität							
	n. d.	n. d.	n. d.	n. d.	n. d.	n. d.	n. d.
Bestätigte Krankheitsschübe							
	○	●	●	●	●	●	○
Behinderungsprogression (EDSS)							
	○	●	●	●	●	●	○
Behinderungsprogression (MSFC)							
	○	●	●	●	●	●	○
Schweregrad der kognitiven Dysfunktion (SDMT)							
	○	●	●	●	●	●	○
Schweregrad der kognitiven Dysfunktion (BVMT-R)							
	○	●	●	●	●	●	○
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)							
	○	●	●	●	●	●	○
Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)							
	○	●	●	●	●	●	○
Gesundheitsbezogene Lebensqualität (MSIS-29)							
	○	●	●	●	●	●	○

Gesundheitsbezogene Lebensqualität (EQ-5D)							
	○	●	●	●	●	●	○
Unerwünschte Ereignisse							
	○	●	●	●	●	●	○
Suizidalität							
	○	●	●	●	●	●	○
Läsionslast							
	○	●	●	●	●	●	○
Hirnvolumen							
	○	●	●	●	●	●	○
●: A priori geplante Subgruppenanalyse. ○: Posthoc durchgeführte Subgruppenanalyse. n.d.: Subgruppenanalyse nicht durchgeführt.							

Für das vorliegende Dossier wurden alle Endpunkte (außer Mortalität) in den folgenden Subgruppen analysiert.

Präspezifizierte Subgruppen:

- Geschlecht (männlich; weiblich),
- Vorherige krankheitsmodifizierende Therapie (ja; nein),
- Vorherige Behandlung mit IFN-beta 1b (ja; nein),
- Patienten mit schnell fortschreitender Erkrankung, d. h. EDSS-Anstieg um mindestens 1,5 Punkte in den zwei Jahren vor oder zu Studienbeginn (ja; nein),
- Moderater oder schwerer Erkrankungsverlauf (ja; nein), d. h. Wert von ≥ 4 auf der MS Status Scale (ja; nein).

Nachträglich definierte Subgruppen:

- Alter (18 – 30 Jahre; 31 – 40 Jahre bzw. 18 – 40 Jahre; 41 – 55 Jahre; > 55 Jahre),
- Region (Europa; andere Regionen).

Im Studienbericht definierte, aber im Dossier nicht aufgenommene Subgruppen:

- Vorhandensein aufgesetzter Schübe in den vergangenen zwei Jahren (ja; nein),
- Vorhandensein Gd-anreichernder Läsionen (ja; nein),
- Aufgesetzte Schübe während der Studie (ja; nein).

Aus der Vielzahl der Kombinationen aus Subgruppen und Endpunkten ergaben sich multiple statistische Tests. Trotz der Multiplizität wurden alle statistischen Tests ohne Adjustierung des Signifikanzniveaus durchgeführt. Die Aussagekraft der durchgeführten Subgruppenanalysen kann insbesondere in Subgruppen mit kleinen Fallzahlen in Verbindung mit einer geringen Zahl an Ereignissen zusätzlich limitiert sein.

Für die Subgruppenanalysen nach Alter in der Subpopulation B wurden die Altersgruppen 18 bis 30 Jahre und 31 bis 40 Jahre zusammengefasst, da die Fallzahl in der Gruppe der jüngsten Patienten zu klein für eine separate Analyse war.

Die Subgruppenanalyse zur Suizidalität anhand des C-SSRS erfolgte aufgrund zu geringer Ereigniszahlen in den einzelnen Kategorien nur für Zusammenfassungen der Kategorien. Für die Kategorie „Selbstverletzung“ erfolgte keine Auswertung nach Subgruppen, da hier weniger als zehn Ereignisse aufgetreten waren. Die Zusammenfassung für die Auswertung der Subgruppen war wie folgt:

- Gesamtpopulation
  - Suizidgedanken Kategorien 1 bis 5,
  - Suizidales Verhalten Kategorien 6 bis 10,
  - Suizidgedanken oder suizidales Verhalten Kategorien 1 bis 10.
- Subpopulation B
  - Suizidgedanken oder suizidales Verhalten Kategorien 1 bis 10.

Stellen Sie anschließend in Tabelle 4-129 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante (p < 0,05) Interaktionsterme.

Tabelle 4-129: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für die Studie EXPAND (Gesamtpopulation)

Endpunkt	**Subgruppenmerkmal**	**Interaktionstest (p-Wert)[a]**
Mortalität		
Todesfälle		k. A.
Bestätigte Krankheitsschübe		
Jährliche Schubrate	Alter	0,308
	Geschlecht	**0,037**
	DMT-Vorbehandlung	0,627
	IFN-beta-1b-Vorbehandlung	0,637
	Schnell fortschreitende Erkrankung	0,446
	Erkrankungsverlauf	0,209
	Region	0,336
Zeit bis zum ersten bestätigten Schub	Alter	0,347
	Geschlecht	0,093
	DMT-Vorbehandlung	0,857
	IFN-beta-1b-Vorbehandlung	0,572
	Schnell fortschreitende Erkrankung	0,289
	Erkrankungsverlauf	0,050
	Region	0,403
Behinderungsprogression (EDSS)		
Inzidenz der nach 3 Monaten bestätigten Progression	Alter	0,748
	Geschlecht	0,569
	DMT-Vorbehandlung	0,525
	IFN-beta-1b-Vorbehandlung	0,448

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	Schnell fortschreitende Erkrankung	0,280
	Erkrankungsverlauf	0,682
	Region	0,271
Zeit bis zur nach 3 Monaten bestätigten Progression	Alter	0,879
	Geschlecht	0,842
	DMT-Vorbehandlung	0,497
	IFN-beta-1b-Vorbehandlung	0,404
	Schnell fortschreitende Erkrankung	0,180
	Erkrankungsverlauf	0,703
	Region	0,269
Inzidenz der nach 6 Monaten bestätigten Progression	Alter	0,173
	Geschlecht	0,357
	DMT-Vorbehandlung	0,373
	IFN-beta-1b-Vorbehandlung	0,125
	Schnell fortschreitende Erkrankung	0,284
	Erkrankungsverlauf	0,465
	Region	0,586
Zeit bis zur nach 6 Monaten bestätigten Progression	Alter	0,302
	Geschlecht	0,435
	DMT-Vorbehandlung	0,288
	IFN-beta-1b-Vorbehandlung	0,098
	Schnell fortschreitende Erkrankung	0,210
	Erkrankungsverlauf	0,424
	Region	0,584
Behinderungsprogression (MSFC)		
MSFC-z – Veränderung seit Studienbeginn	Alter	0,124
	Geschlecht	**0,004**
	DMT-Vorbehandlung	0,639
	IFN-beta-1b-Vorbehandlung	0,277
	Schnell fortschreitende Erkrankung	0,240
	Erkrankungsverlauf	0,980
	Region	0,183
PASAT-3 – Veränderung seit Studienbeginn	Alter	0,932
	Geschlecht	**0,012**
	DMT-Vorbehandlung	0,928
	IFN-beta-1b-Vorbehandlung	0,495
	Schnell fortschreitende Erkrankung	0,079
	Erkrankungsverlauf	0,956
	Region	0,313
9-HPT – Veränderung seit Studienbeginn	Alter	0,671
	Geschlecht	0,571
	DMT-Vorbehandlung	0,730
	IFN-beta-1b-Vorbehandlung	0,262
	Schnell fortschreitende Erkrankung	0,910
	Erkrankungsverlauf	0,589
	Region	0,740
T25W – Veränderung seit Studienbeginn	Alter	0,727
	Geschlecht	**0,011**
	DMT-Vorbehandlung	0,886
	IFN-beta-1b-Vorbehandlung	0,223
	Schnell fortschreitende Erkrankung	0,156
	Erkrankungsverlauf	0,581
	Region	0,382

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
Inzidenz der Verschlechterung im T25W	Alter	0,756
	Geschlecht	0,239
	DMT-Vorbehandlung	0,206
	IFN-beta-1b-Vorbehandlung	0,287
	Schnell fortschreitende Erkrankung	0,388
	Erkrankungsverlauf	0,578
	Region	0,825
Zeit bis zur Verschlechterung im T25W	Alter	0,300
	Geschlecht	0,530
	DMT-Vorbehandlung	0,150
	IFN-beta-1b-Vorbehandlung	0,304
	Schnell fortschreitende Erkrankung	0,682
	Erkrankungsverlauf	0,282
	Region	0,695
Schweregrad der kognitiven Dysfunktion (SDMT)		
SDMT – Veränderung seit Studienbeginn	Alter	0,899
	Geschlecht	0,238
	DMT-Vorbehandlung	0,708
	IFN-beta-1b-Vorbehandlung	0,515
	Schnell fortschreitende Erkrankung	0,155
	Erkrankungsverlauf	0,983
	Region	n. v.[b]
Inzidenz der Verbesserung im SDMT	Alter	0,387
	Geschlecht	0,137
	DMT-Vorbehandlung	0,551
	IFN-beta-1b-Vorbehandlung	0,844
	Schnell fortschreitende Erkrankung	0,470
	Erkrankungsverlauf	0,986
	Region	0,187
Inzidenz der Verschlechterung im SDMT	Alter	0,557
	Geschlecht	**0,003**
	DMT-Vorbehandlung	0,343
	IFN-beta-1b-Vorbehandlung	0,436
	Schnell fortschreitende Erkrankung	0,705
	Erkrankungsverlauf	0,108
	Region	0,401
Schweregrad der kognitiven Dysfunktion (BVMT-R)		
BVMT-R Delayed Recall Score – Veränderung seit Studienbeginn	Alter	0,455
	Geschlecht	0,952
	DMT-Vorbehandlung	0,608
	IFN-beta-1b-Vorbehandlung	0,470
	Schnell fortschreitende Erkrankung	0,092
	Erkrankungsverlauf	0,319
	Region	0,229
BVMT-R Total Recall Score – Veränderung seit Studienbeginn	Alter	0,686
	Geschlecht	0,251
	DMT-Vorbehandlung	0,416
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	0,937
	Erkrankungsverlauf	0,217
	Region	0,816

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)		
LCVA – Veränderung seit Studienbeginn	Alter	0,661
	Geschlecht	**0,016**
	DMT-Vorbehandlung	**0,012**
	IFN-beta-1b-Vorbehandlung	0,968
	Schnell fortschreitende Erkrankung	0,091
	Erkrankungsverlauf	0,066
	Region	0,180
Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)		
MSWS-12 – Veränderung seit Studienbeginn	Alter	0,699
	Geschlecht	0,381
	DMT-Vorbehandlung	0,668
	IFN-beta-1b-Vorbehandlung	0,898
	Schnell fortschreitende Erkrankung	0,523
	Erkrankungsverlauf	0,123
	Region	0,729
Gesundheitsbezogene Lebensqualität (MSIS-29)		
MSIS-29 Physical Impact Scale – Veränderung seit Studienbeginn	Alter	**0,023**
	Geschlecht	0,616
	DMT-Vorbehandlung	0,465
	IFN-beta-1b-Vorbehandlung	0,490
	Schnell fortschreitende Erkrankung	0,543
	Erkrankungsverlauf	0,223
	Region	0,915
MSIS-29 Psychological Impact Scale – Veränderung seit Studienbeginn	Alter	**0,044**
	Geschlecht	0,104
	DMT-Vorbehandlung	0,720
	IFN-beta-1b-Vorbehandlung	0,773
	Schnell fortschreitende Erkrankung	0,614
	Erkrankungsverlauf	0,400
	Region	0,703
Gesundheitsbezogene Lebensqualität (EQ-5D)		
EQ-5D-Indexwert – Veränderung seit Studienbeginn	Alter	0,075
	Geschlecht	0,144
	DMT-Vorbehandlung	0,564
	IFN-beta-1b-Vorbehandlung	0,459
	Schnell fortschreitende Erkrankung	0,264
	Erkrankungsverlauf	0,763
	Region	n. v.[b]
EQ-5D-VAS – Veränderung seit Studienbeginn	Alter	n. v.[b]
	Geschlecht	0,843
	DMT-Vorbehandlung	0,952
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	0,918
	Erkrankungsverlauf	**0,036**
	Region	0,095
Unerwünschte Ereignisse[c]		
UE	Alter	0,760
	Geschlecht	0,274
	DMT-Vorbehandlung	0,144
	IFN-beta-1b-Vorbehandlung	0,804

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	Schnell fortschreitende Erkrankung	0,682
	Erkrankungsverlauf	0,148
	Region	0,440
UE ohne krankheitsspezifische UE	Alter	0,977
	Geschlecht	0,579
	DMT-Vorbehandlung	**0,036**
	IFN-beta-1b-Vorbehandlung	0,604
	Schnell fortschreitende Erkrankung	0,537
	Erkrankungsverlauf	0,079
	Region	0,899
SUE	Alter	0,062
	Geschlecht	0,089
	DMT-Vorbehandlung	0,986
	IFN-beta-1b-Vorbehandlung	0,640
	Schnell fortschreitende Erkrankung	0,050
	Erkrankungsverlauf	0,527
	Region	0,772
SUE ohne krankheitsspezifische UE	Alter	0,349
	Geschlecht	0,134
	DMT-Vorbehandlung	0,952
	IFN-beta-1b-Vorbehandlung	0,446
	Schnell fortschreitende Erkrankung	0,421
	Erkrankungsverlauf	0,238
	Region	0,917
Schwere UE	Alter	0,236
	Geschlecht	0,093
	DMT-Vorbehandlung	0,234
	IFN-beta-1b-Vorbehandlung	0,079
	Schnell fortschreitende Erkrankung	0,534
	Erkrankungsverlauf	0,616
	Region	0,570
Schwere UE ohne krankheitsspezifische Ereignisse	Alter	0,243
	Geschlecht	0,065
	DMT-Vorbehandlung	0,077
	IFN-beta-1b-Vorbehandlung	0,089
	Schnell fortschreitende Erkrankung	0,678
	Erkrankungsverlauf	0,176
	Region	0,841
Studienabbruch aufgrund von UE	Alter	n. v.[b]
	Geschlecht	0,909
	DMT-Vorbehandlung	0,612
	IFN-beta-1b-Vorbehandlung	0,228
	Schnell fortschreitende Erkrankung	0,840
	Erkrankungsverlauf	0,765
	Region	0,828
Behandlungsabbruch aufgrund von UE	Alter	n. v.[b]
	Geschlecht	0,703
	DMT-Vorbehandlung	0,514
	IFN-beta-1b-Vorbehandlung	0,560
	Schnell fortschreitende Erkrankung	0,225
	Erkrankungsverlauf	0,590
	Region	0,592
Zeit bis zum ersten UE	Alter	0,978
	Geschlecht	0,321

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	DMT-Vorbehandlung	0,313
	IFN-beta-1b-Vorbehandlung	0,736
	Schnell fortschreitende Erkrankung	0,259
	Erkrankungsverlauf	0,377
	Region	0,792
Zeit bis zum ersten UE ohne krankheitsspezifische UE	Alter	0,991
	Geschlecht	0,650
	DMT-Vorbehandlung	0,204
	IFN-beta-1b-Vorbehandlung	0,960
	Schnell fortschreitende Erkrankung	0,191
	Erkrankungsverlauf	0,201
	Region	0,578
Zeit bis zum ersten SUE	Alter	**0,040**
	Geschlecht	0,098
	DMT-Vorbehandlung	0,917
	IFN-beta-1b-Vorbehandlung	0,698
	Schnell fortschreitende Erkrankung	**0,050**
	Erkrankungsverlauf	0,459
	Region	0,675
Zeit bis zum ersten SUE ohne krankheitsspezifische UE	Alter	0,317
	Geschlecht	0,133
	DMT-Vorbehandlung	0,877
	IFN-beta-1b-Vorbehandlung	0,525
	Schnell fortschreitende Erkrankung	0,418
	Erkrankungsverlauf	0,208
	Region	0,967
Zeit bis zum ersten schweren UE	Alter	0,201
	Geschlecht	0,084
	DMT-Vorbehandlung	0,242
	IFN-beta-1b-Vorbehandlung	0,069
	Schnell fortschreitende Erkrankung	0,536
	Erkrankungsverlauf	0,554
	Region	0,521
Zeit bis zum ersten schweren UE ohne krankheitsspezifische UE	Alter	0,269
	Geschlecht	0,055
	DMT-Vorbehandlung	0,085
	IFN-beta-1b-Vorbehandlung	0,076
	Schnell fortschreitende Erkrankung	0,694
	Erkrankungsverlauf	0,157
	Region	0,811
Zeit bis zum Studienabbruch aufgrund von UE	Alter	0,392
	Geschlecht	0,905
	DMT-Vorbehandlung	0,624
	IFN-beta-1b-Vorbehandlung	0,243
	Schnell fortschreitende Erkrankung	0,839
	Erkrankungsverlauf	0,784
	Region	0,834
Zeit bis zum Behandlungsabbruch aufgrund von UE	Alter	0,792
	Geschlecht	0,705
	DMT-Vorbehandlung	0,538
	IFN-beta-1b-Vorbehandlung	0,602
	Schnell fortschreitende Erkrankung	0,231
	Erkrankungsverlauf	0,631
	Region	0,599

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
Unerwünschte Ereignisse von besonderem Interesse		
Infektionen und parasitäre Erkrankungen	Alter	0,684
	Geschlecht	0,915
	DMT-Vorbehandlung	0,502
	IFN-beta-1b-Vorbehandlung	0,754
	Schnell fortschreitende Erkrankung	0,339
	Erkrankungsverlauf	0,422
	Region	0,574
Herpesvirusinfektionen	Alter	n. v.[b]
	Geschlecht	0,913
	DMT-Vorbehandlung	**0,014**
	IFN-beta-1b-Vorbehandlung	0,276
	Schnell fortschreitende Erkrankung	0,269
	Erkrankungsverlauf	**0,046**
	Region	0,380
Varizella-Zoster-Infektionen	Alter	n. v.[b]
	Geschlecht	0,897
	DMT-Vorbehandlung	0,059
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	0,822
	Erkrankungsverlauf	0,136
	Region	n. v.[b]
Lymphopenie	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Makulaödem	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Erhöhte Leberenzymwerte	Alter	0,991
	Geschlecht	0,597
	DMT-Vorbehandlung	0,417
	IFN-beta-1b-Vorbehandlung	0,758
	Schnell fortschreitende Erkrankung	0,428
	Erkrankungsverlauf	0,057
	Region	0,852
Malignitäten	Alter	n. v.[b]
	Geschlecht	0,420
	DMT-Vorbehandlung	0,856
	IFN-beta-1b-Vorbehandlung	**0,044**
	Schnell fortschreitende Erkrankung	0,802
	Erkrankungsverlauf	0,363
	Region	0,962
Hypertonie	Alter	0,502
	Geschlecht	0,138
	DMT-Vorbehandlung	0,327

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	IFN-beta-1b-Vorbehandlung	0,310
	Schnell fortschreitende Erkrankung	0,535
	Erkrankungsverlauf	0,943
	Region	0,956
Krampfanfälle	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	0,553
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Thromboembolische Ereignisse	Alter	n. v.[b]
	Geschlecht	0,409
	DMT-Vorbehandlung	0,434
	IFN-beta-1b-Vorbehandlung	0,458
	Schnell fortschreitende Erkrankung	0,752
	Erkrankungsverlauf	0,616
	Region	0,497
Atemwegserkrankungen	Alter	0,470
	Geschlecht	0,904
	DMT-Vorbehandlung	**0,019**
	IFN-beta-1b-Vorbehandlung	0,055
	Schnell fortschreitende Erkrankung	0,858
	Erkrankungsverlauf	0,675
	Region	0,633
Periphere Ödeme / Schwellungen	Alter	n. v.[b]
	Geschlecht	**0,006**
	DMT-Vorbehandlung	0,490
	IFN-beta-1b-Vorbehandlung	0,734
	Schnell fortschreitende Erkrankung	0,085
	Erkrankungsverlauf	0,320
	Region	0,065
Asthma/Bronchospasmen	Alter	n. v.[b]
	Geschlecht	0,773
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	0,444
	Schnell fortschreitende Erkrankung	0,929
	Erkrankungsverlauf	0,241
	Region	n. v.[b]
Schwangerschaften / neonatale Ereignisse	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Bradyarrhythmien/Bradykardien	Alter	0,750
	Geschlecht	**0,031**
	DMT-Vorbehandlung	0,561
	IFN-beta-1b-Vorbehandlung	**0,049**
	Schnell fortschreitende Erkrankung	0,605
	Erkrankungsverlauf	0,235
	Region	0,549

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
Suizidalität (C-SSRS)		
Suizidgedanken (Kat. 1 bis 5)	Alter	0,884
	Geschlecht	0,958
	DMT-Vorbehandlung	0,608
	IFN-beta-1b-Vorbehandlung	0,887
	Schnell fortschreitende Erkrankung	0,933
	Erkrankungsverlauf	0,569
	Region	0,119
Suizidales Verhalten (Kat. 6 bis 10)	Alter	0,824
	Geschlecht	0,335
	DMT-Vorbehandlung	0,962
	IFN-beta-1b-Vorbehandlung	0,674
	Schnell fortschreitende Erkrankung	0,413
	Erkrankungsverlauf	0,379
	Region	0,818
Suizidgedanken oder suizidales Verhalten (Kat. 1 bis 10)	Alter	0,975
	Geschlecht	0,794
	DMT-Vorbehandlung	0,461
	IFN-beta-1b-Vorbehandlung	0,840
	Schnell fortschreitende Erkrankung	0,750
	Erkrankungsverlauf	0,670
	Region	0,199
Läsionslast		
T2-Läsionsvolumen – Veränderung seit Studienbeginn	Alter	**< 0,001**
	Geschlecht	0,587
	DMT-Vorbehandlung	**0,009**
	IFN-beta-1b-Vorbehandlung	0,054
	Schnell fortschreitende Erkrankung	0,304
	Erkrankungsverlauf	**0,035**
	Region	**0,009**
Anzahl neuer oder sich vergrößernder T2-Läsionen	Alter	0,869
	Geschlecht	0,604
	DMT-Vorbehandlung	**0,030**
	IFN-beta-1b-Vorbehandlung	**0,010**
	Schnell fortschreitende Erkrankung	0,377
	Erkrankungsverlauf	0,549
	Region	**0,045**
Anzahl Gd-anreichernder T1-Läsionen	Alter	0,710
	Geschlecht	0,095
	DMT-Vorbehandlung	0,355
	IFN-beta-1b-Vorbehandlung	0,142
	Schnell fortschreitende Erkrankung	0,189
	Erkrankungsverlauf	0,724
	Region	**0,001**
Hirnvolumen		
Prozentuale Veränderung im Hirnvolumen	Alter	0,138
	Geschlecht	0,461
	DMT-Vorbehandlung	0,346
	IFN-beta-1b-Vorbehandlung	0,790
	Schnell fortschreitende Erkrankung	0,901
	Erkrankungsverlauf	0,637
	Region	0,483

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]

BVMT-R: Brief Visuospatial Memory Test – Revised; C-SSRS: Columbia-Suicide Severity Rating Scale; DMT: Disease Modifying Therapie; EQ-5D: EuroQol – 5 Dimensionen; IFN: Interferon; k. A.: keine Angabe; Kat.: Kategorie; LCVA: Low Contrast Visual Acuity; MSFC: Multiple Sclerosis Functional Composite; MSIS: Multiple Sclerosis Impact Scale; MSWS: Multiple Sclerosis Walking Scale; n. v.: nicht verfügbar; PASAT: Paced Auditory Serial Addition Test; PT: Preferred Term; QoL: Quality of Life; SDMT: Symbol Digit Modalities Test; T25W: Timed 25 Foot Walking Test; SOC: System Organ Class; SUE: Schwere Unerwünschte Ereignisse; UE: Unerwünschte Ereignisse; 9-HPT: 9 Hole Peg Test.

a: Beleg für eine mögliche Interaktion bei $p < 0,05$ (gefettet dargestellt); es werden im Folgenden die Subgruppenanalysen im Detail dargestellt, für die der Interaktionstest einen Beleg für eine mögliche Effektmodifikation ergab.

b: keine Bestimmung möglich, da keine Konvergenz im Modell.

c: Subgruppenanalysen für UE nach SOC/PT; SUE nach SOC/PT und schwere UE nach SOC/PT wurden nur durchgeführt, wenn der Effekt in der Hauptanalyse statistisch signifikant war ($p < 0,05$). Die Subgruppenauswertungen nach SOC/PT sind im Anhang 1 zu finden.

Tabelle 4-130: Ergebnis des Interaktionsterms der Subgruppenanalysen je Endpunkt für die Studie EXPAND (Subpopulation B)

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
Mortalität		
Todesfälle		k. A.
Bestätigte Krankheitsschübe		
Jährliche Schubrate	Alter	0,358
	Geschlecht	0,441
	DMT-Vorbehandlung	0,380
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	0,690
	Erkrankungsverlauf	0,446
	Region	0,386
Zeit bis zum ersten bestätigten Schub	Alter	0,651
	Geschlecht	0,897
	DMT-Vorbehandlung	0,995
	IFN-beta-1b-Vorbehandlung	0,526
	Schnell fortschreitende Erkrankung	0,328
	Erkrankungsverlauf	0,837
	Region	0,897
Behinderungsprogression (EDSS)		
Inzidenz der nach 3 Monaten bestätigten Progression	Alter	0,272
	Geschlecht	0,246
	DMT-Vorbehandlung	0,557
	IFN-beta-1b-Vorbehandlung	0,144
	Schnell fortschreitende Erkrankung	0,948
	Erkrankungsverlauf	0,750
	Region	0,595
Zeit bis zur nach 3 Monaten bestätigten Progression	Alter	0,122
	Geschlecht	0,622
	DMT-Vorbehandlung	0,734
	IFN-beta-1b-Vorbehandlung	0,291
	Schnell fortschreitende Erkrankung	0,699
	Erkrankungsverlauf	0,627

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	Region	0,264
Inzidenz der nach 6 Monaten bestätigten Progression	Alter	0,406
	Geschlecht	**0,047**
	DMT-Vorbehandlung	0,870
	IFN-beta-1b-Vorbehandlung	0,229
	Schnell fortschreitende Erkrankung	0,682
	Erkrankungsverlauf	0,574
	Region	0,898
Zeit bis zur nach 6 Monaten bestätigten Progression	Alter	0,386
	Geschlecht	0,253
	DMT-Vorbehandlung	0,959
	IFN-beta-1b-Vorbehandlung	0,399
	Schnell fortschreitende Erkrankung	0,646
	Erkrankungsverlauf	0,953
	Region	0,688
Behinderungsprogression (MSFC)		
MSFC-z – Veränderung seit Studienbeginn	Alter	0,800
	Geschlecht	0,927
	DMT-Vorbehandlung	0,905
	IFN-beta-1b-Vorbehandlung	0,543
	Schnell fortschreitende Erkrankung	0,834
	Erkrankungsverlauf	0,506
	Region	0,902
PASAT-3 – Veränderung seit Studienbeginn	Alter	0,265
	Geschlecht	0,804
	DMT-Vorbehandlung	0,853
	IFN-beta-1b-Vorbehandlung	0,335
	Schnell fortschreitende Erkrankung	0,771
	Erkrankungsverlauf	0,957
	Region	0,176
9-HPT – Veränderung seit Studienbeginn	Alter	0,719
	Geschlecht	0,390
	DMT-Vorbehandlung	0,403
	IFN-beta-1b-Vorbehandlung	0,115
	Schnell fortschreitende Erkrankung	0,857
	Erkrankungsverlauf	0,572
	Region	0,212
T25W – Veränderung seit Studienbeginn	Alter	0,217
	Geschlecht	0,205
	DMT-Vorbehandlung	0,767
	IFN-beta-1b-Vorbehandlung	0,709
	Schnell fortschreitende Erkrankung	**0,024**
	Erkrankungsverlauf	0,583
	Region	0,229
Inzidenz der Verschlechterung im T25W	Alter	0,879
	Geschlecht	0,384
	DMT-Vorbehandlung	0,196
	IFN-beta-1b-Vorbehandlung	0,353
	Schnell fortschreitende Erkrankung	0,717
	Erkrankungsverlauf	0,247
	Region	0,152
Zeit bis zur Verschlechterung im T25W	Alter	0,604
	Geschlecht	0,774

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	DMT-Vorbehandlung	0,148
	IFN-beta-1b-Vorbehandlung	0,315
	Schnell fortschreitende Erkrankung	0,700
	Erkrankungsverlauf	0,125
	Region	0,694
Schweregrad der kognitiven Dysfunktion (SDMT)		
SDMT – Veränderung seit Studienbeginn	Alter	0,664
	Geschlecht	0,866
	DMT-Vorbehandlung	0,670
	IFN-beta-1b-Vorbehandlung	0,771
	Schnell fortschreitende Erkrankung	0,619
	Erkrankungsverlauf	0,981
	Region	0,083
Inzidenz der Verbesserung im SDMT	Alter	0,828
	Geschlecht	0,758
	DMT-Vorbehandlung	0,072
	IFN-beta-1b-Vorbehandlung	0,862
	Schnell fortschreitende Erkrankung	0,093
	Erkrankungsverlauf	0,666
	Region	0,443
Inzidenz der Verschlechterung im SDMT	Alter	**0,034**
	Geschlecht	0,610
	DMT-Vorbehandlung	0,057
	IFN-beta-1b-Vorbehandlung	0,677
	Schnell fortschreitende Erkrankung	0,601
	Erkrankungsverlauf	0,430
	Region	0,559
Schweregrad der kognitiven Dysfunktion (BVMT-R)		
BVMT-R Delayed Recall Score – Veränderung seit Studienbeginn	Alter	0,393
	Geschlecht	**0,036**
	DMT-Vorbehandlung	0,910
	IFN-beta-1b-Vorbehandlung	0,079
	Schnell fortschreitende Erkrankung	0,566
	Erkrankungsverlauf	0,377
	Region	0,901
BVMT-R Total Recall Score – Veränderung seit Studienbeginn	Alter	**0,016**
	Geschlecht	**0,014**
	DMT-Vorbehandlung	0,177
	IFN-beta-1b-Vorbehandlung	**0,034**
	Schnell fortschreitende Erkrankung	0,123
	Erkrankungsverlauf	0,059
	Region	0,297
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)		
LCVA – Veränderung seit Studienbeginn	Alter	0,303
	Geschlecht	0,117
	DMT-Vorbehandlung	0,845
	IFN-beta-1b-Vorbehandlung	**0,038**
	Schnell fortschreitende Erkrankung	0,394
	Erkrankungsverlauf	0,131
	Region	0,162

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)		
MSWS-12 – Veränderung seit Studienbeginn	Alter	0,515
	Geschlecht	0,138
	DMT-Vorbehandlung	0,786
	IFN-beta-1b-Vorbehandlung	0,437
	Schnell fortschreitende Erkrankung	0,775
	Erkrankungsverlauf	0,292
	Region	0,814
Gesundheitsbezogene Lebensqualität (MSIS-29)		
MSIS-29 Physical Impact Scale – Veränderung seit Studienbeginn	Alter	0,490
	Geschlecht	0,423
	DMT-Vorbehandlung	0,565
	IFN-beta-1b-Vorbehandlung	0,692
	Schnell fortschreitende Erkrankung	0,168
	Erkrankungsverlauf	0,866
	Region	0,861
MSIS-29 Psychological Impact Scale – Veränderung seit Studienbeginn	Alter	0,777
	Geschlecht	0,388
	DMT-Vorbehandlung	0,869
	IFN-beta-1b-Vorbehandlung	0,390
	Schnell fortschreitende Erkrankung	0,143
	Erkrankungsverlauf	0,522
	Region	0,412
Gesundheitsbezogene Lebensqualität (EQ-5D)		
EQ-5D-Indexwert – Veränderung seit Studienbeginn	Alter	0,743
	Geschlecht	0,425
	DMT-Vorbehandlung	0,947
	IFN-beta-1b-Vorbehandlung	0,207
	Schnell fortschreitende Erkrankung	**0,037**
	Erkrankungsverlauf	0,441
	Region	0,423
EQ-5D-VAS – Veränderung seit Studienbeginn	Alter	0,789
	Geschlecht	0,845
	DMT-Vorbehandlung	0,827
	IFN-beta-1b-Vorbehandlung	0,874
	Schnell fortschreitende Erkrankung	0,082
	Erkrankungsverlauf	0,465
	Region	0,631
Unerwünschte Ereignisse[c]		
UE	Alter	0,155
	Geschlecht	**0,028**
	DMT-Vorbehandlung	0,705
	IFN-beta-1b-Vorbehandlung	0,681
	Schnell fortschreitende Erkrankung	0,605
	Erkrankungsverlauf	0,873
	Region	0,138
UE ohne krankheitsspezifische UE	Alter	0,491
	Geschlecht	0,103
	DMT-Vorbehandlung	0,978
	IFN-beta-1b-Vorbehandlung	0,528
	Schnell fortschreitende Erkrankung	0,654
	Erkrankungsverlauf	0,607

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	Region	0,243
SUE	Alter	0,851
	Geschlecht	0,088
	DMT-Vorbehandlung	0,096
	IFN-beta-1b-Vorbehandlung	0,738
	Schnell fortschreitende Erkrankung	0,257
	Erkrankungsverlauf	n. v. [b]
	Region	0,115
SUE ohne krankheitsspezifische UE	Alter	0,716
	Geschlecht	0,162
	DMT-Vorbehandlung	0,111
	IFN-beta-1b-Vorbehandlung	0,753
	Schnell fortschreitende Erkrankung	0,268
	Erkrankungsverlauf	n. v. [b]
	Region	0,192
Schwere UE	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	0,906
	IFN-beta-1b-Vorbehandlung	0,246
	Schnell fortschreitende Erkrankung	0,428
	Erkrankungsverlauf	0,874
	Region	0,190
Schwere UE ohne krankheitsspezifische Ereignisse	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	0,852
	IFN-beta-1b-Vorbehandlung	0,223
	Schnell fortschreitende Erkrankung	0,885
	Erkrankungsverlauf	0,904
	Region	0,363
Studienabbruch aufgrund von UE	Alter	n. v.[b]
	Geschlecht	0,721
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	0,769
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Behandlungsabbruch aufgrund von UE	Alter	n. v.[b]
	Geschlecht	0,257
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	0,843
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Zeit bis zum ersten UE	Alter	0,738
	Geschlecht	**0,028**
	DMT-Vorbehandlung	0,753
	IFN-beta-1b-Vorbehandlung	0,536
	Schnell fortschreitende Erkrankung	0,907
	Erkrankungsverlauf	0,716
	Region	0,955
Zeit bis zum ersten UE ohne krankheitsspezifische UE	Alter	0,848
	Geschlecht	**0,033**
	DMT-Vorbehandlung	0,915
	IFN-beta-1b-Vorbehandlung	0,354

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	Schnell fortschreitende Erkrankung	0,439
	Erkrankungsverlauf	0,339
	Region	0,973
Zeit bis zum ersten SUE	Alter	0,852
	Geschlecht	0,072
	DMT-Vorbehandlung	0,093
	IFN-beta-1b-Vorbehandlung	0,811
	Schnell fortschreitende Erkrankung	0,228
	Erkrankungsverlauf	0,987
	Region	0,071
Zeit bis zum ersten SUE ohne krankheitsspezifische UE	Alter	0,771
	Geschlecht	0,140
	DMT-Vorbehandlung	0,108
	IFN-beta-1b-Vorbehandlung	0,717
	Schnell fortschreitende Erkrankung	0,247
	Erkrankungsverlauf	0,989
	Region	0,136
Zeit bis zum ersten schweren UE	Alter	0,964
	Geschlecht	0,990
	DMT-Vorbehandlung	0,890
	IFN-beta-1b-Vorbehandlung	0,289
	Schnell fortschreitende Erkrankung	0,381
	Erkrankungsverlauf	0,855
	Region	0,122
Zeit bis zum ersten schweren UE ohne krankheitsspezifische UE	Alter	0,985
	Geschlecht	0,991
	DMT-Vorbehandlung	0,857
	IFN-beta-1b-Vorbehandlung	0,256
	Schnell fortschreitende Erkrankung	0,797
	Erkrankungsverlauf	0,866
	Region	0,245
Zeit bis zum Studienabbruch aufgrund von UE	Alter	1,000
	Geschlecht	0,762
	DMT-Vorbehandlung	1,000
	IFN-beta-1b-Vorbehandlung	1,000
	Schnell fortschreitende Erkrankung	0,786
	Erkrankungsverlauf	1,000
	Region	0,995
Zeit bis zum Behandlungsabbruch aufgrund von UE	Alter	1,000
	Geschlecht	0,232
	DMT-Vorbehandlung	0,994
	IFN-beta-1b-Vorbehandlung	0,993
	Schnell fortschreitende Erkrankung	0,930
	Erkrankungsverlauf	1,000
	Region	0,993
Unerwünschte Ereignisse von besonderem Interesse		
Infektionen und parasitäre Erkrankungen	Alter	0,156
	Geschlecht	0,510
	DMT-Vorbehandlung	0,551
	IFN-beta-1b-Vorbehandlung	0,299
	Schnell fortschreitende Erkrankung	0,191
	Erkrankungsverlauf	0,285
	Region	0,524

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
Herpesvirusinfektionen	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Varizella-Zoster-Infektionen	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Lymphopenie	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Makulaödem	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Erhöhte Leberenzymwerte	Alter	0,994
	Geschlecht	0,316
	DMT-Vorbehandlung	0,329
	IFN-beta-1b-Vorbehandlung	0,746
	Schnell fortschreitende Erkrankung	0,846
	Erkrankungsverlauf	0,229
	Region	n. v.[b]
Malignitäten	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Hypertonie	Alter	n. v.[b]
	Geschlecht	**0,050**
	DMT-Vorbehandlung	0,709
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	0,810
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Krampfanfälle	Alter	n. v.[c]
	Geschlecht	n. v.[c]
	DMT-Vorbehandlung	n. v.[c]
	IFN-beta-1b-Vorbehandlung	n. v.[c]
	Schnell fortschreitende Erkrankung	n. v.[c]

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
	Erkrankungsverlauf	n. v.[c]
	Region	n. v.[c]
Thromboembolische Ereignisse	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Atemwegserkrankungen	Alter	n. v.[b]
	Geschlecht	0,175
	DMT-Vorbehandlung	**0,038**
	IFN-beta-1b-Vorbehandlung	0,694
	Schnell fortschreitende Erkrankung	0,448
	Erkrankungsverlauf	0,808
	Region	0,538
Periphere Ödeme / Schwellungen	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Asthma/Bronchospasmen	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Schwangerschaften / neonatale Ereignisse	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Bradyarrhythmien/Bradykardien	Alter	n. v.[b]
	Geschlecht	n. v.[b]
	DMT-Vorbehandlung	n. v.[b]
	IFN-beta-1b-Vorbehandlung	n. v.[b]
	Schnell fortschreitende Erkrankung	n. v.[b]
	Erkrankungsverlauf	n. v.[b]
	Region	n. v.[b]
Suizidalität (C-SSRS)		
Suizidgedanken oder suizidales Verhalten (Kat. 1 bis 10)	Alter	0,696
	Geschlecht	0,796
	DMT-Vorbehandlung	0,218
	IFN-beta-1b-Vorbehandlung	0,952
	Schnell fortschreitende Erkrankung	0,234
	Erkrankungsverlauf	0,959
	Region	0,852

Endpunkt	Subgruppenmerkmal	Interaktionstest (p-Wert)[a]
Läsionslast		
T2-Läsionsvolumen – Veränderung seit Studienbeginn	Alter	0,659
	Geschlecht	0,337
	DMT-Vorbehandlung	0,676
	IFN-beta-1b-Vorbehandlung	0,896
	Schnell fortschreitende Erkrankung	0,505
	Erkrankungsverlauf	0,051
	Region	0,625
Anzahl neuer oder sich vergrößernder T2-Läsionen	Alter	0,356
	Geschlecht	0,706
	DMT-Vorbehandlung	**< 0,001**
	IFN-beta-1b-Vorbehandlung	0,233
	Schnell fortschreitende Erkrankung	0,822
	Erkrankungsverlauf	0,786
	Region	**< 0,001**
Anzahl Gd-anreichernder T1-Läsionen	Alter	n. v.[b]
	Geschlecht	0,508
	DMT-Vorbehandlung	**0,006**
	IFN-beta-1b-Vorbehandlung	**0,027**
	Schnell fortschreitende Erkrankung	0,895
	Erkrankungsverlauf	0,723
	Region	**< 0,001**
Hirnvolumen		
Prozentuale Veränderung im Hirnvolumen	Alter	0,101
	Geschlecht	0,852
	DMT-Vorbehandlung	0,149
	IFN-beta-1b-Vorbehandlung	0,619
	Schnell fortschreitende Erkrankung	0,345
	Erkrankungsverlauf	0,125
	Region	0,761

BVMT-R: Brief Visuospatial Memory Test – Revised; C-SSRS: Columbia-Suicide Severity Rating Scale; DMT: Disease Modifying Therapie; EQ-5D: EuroQol – 5 Dimensionen; IFN: Interferon; k. A.: keine Angabe; Kat.: Kategorie; LCVA: Low Contrast Visual Acuity; MSFC: Multiple Sclerosis Functional Composite; MSIS: Multiple Sclerosis Impact Scale; MSWS: Multiple Sclerosis Walking Scale; n. v.: nicht verfügbar; PASAT: Paced Auditory Serial Addition Test; PT: Preferred Term; QoL: Quality of Life; SDMT: Symbol Digit Modalities Test; T25W: Timed 25 Foot Walking Test; SOC: System Organ Class; SUE: Schwere Unerwünschte Ereignisse; UE: Unerwünschte Ereignisse; 9-HPT: 9 Hole Peg Test.

a: Beleg für eine mögliche Interaktion bei $p < 0,05$ (gefettet dargestellt); es werden im Folgenden die Subgruppenanalysen im Detail dargestellt, für die der Interaktionstest einen Beleg für eine mögliche Effektmodifikation ergab.

b: Keine Bestimmung möglich, da keine Konvergenz im Modell.

c: Subgruppenanalysen für UE nach SOC/PT; SUE nach SOC/PT und schwere UE nach SOC/PT wurden nur durchgeführt, wenn der Effekt in der Hauptanalyse statistisch signifikant war ($p < 0,05$). Die Subgruppenauswertungen nach SOC/PT sind im Anhang 2 zu finden.

Stellen Sie schließlich alle Subgruppenergebnisse dar.

Sofern eine Effektmodifikation für mehr als ein Subgruppenmerkmal vorliegt, kann eine Untersuchung auf eine Wechselwirkung höherer Ordnung sinnvoll sein. Dies gilt insbesondere dann, wenn diese Effektmodifikation konsistent über mehrere Endpunkte besteht. Zur

Interpretation der Ergebnisse sollte dann für diese Endpunkte zusätzlich eine Subgruppenanalyse durchgeführt werden, die die Merkmale mit Effektmodifikation kombiniert. Beispiel: Für die Endpunkte Mortalität, gesundheitsbezogene Lebensqualität und schwere unerwünschte Ereignisse liegt sowohl für das Merkmal Geschlecht (mit den Ausprägungen „weiblich" und „männlich") als auch für das Merkmal Schweregrad (mit den Ausprägungen „niedrig" und „hoch") eine Effektmodifikation vor. Die zusätzliche Subgruppenanalyse erfolgt dann für die 3 genannten Endpunkte für das kombinierte Merkmal Geschlecht/Schweregrad mit den 4 Ausprägungen weiblich/niedrig, weiblich/hoch, männlich/niedrig und männlich/hoch.

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.

Da nur die EXPAND-Studie in die Nutzenbewertung eingeht, ist zur quantitativen Zusammenfassung keine Meta-Analyse erforderlich.

Im Folgenden werden lediglich die Subgruppenanalysen dargestellt, für die der Interaktionstest einen Beleg für eine mögliche Effektmodifikation ergab. Eine Übersicht hierzu liefern Tabelle 4-131 und Tabelle 4-132.

Tabelle 4-131: Zusammenfassung der Ergebnisse aus den Subgruppenanalysen (Gesamtpopulation)

Endpunkt/ Subgruppenmerkmal	**Interaktionstest**	**Subgruppen**	
Jährliche Schubrate			
Geschlecht		**Weiblich**	**Männlich**
	p = 0,037	+	+
Behinderungsprogression (MSFC)			
Geschlecht		**Weiblich**	**Männlich**
Veränderung im MSFC-z	p = 0,004	+ / = [a]	= / = [a]
Veränderung im PASAT-3	p = 0,012	+ / = [a]	= / = [a]
Veränderung im T25W	p = 0,011	= / = [a]	– / = [a]
Schweregrad der kognitiven Dysfunktion (SDMT)			
Geschlecht		**Weiblich**	**Männlich**
Inzidenz einer Verschlechterung	p = 0,003	+	=

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen			
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)					
Geschlecht		**Weiblich**		**Männlich**	
Veränderung im LCVA	p = 0,016	= / = [a]		= / = [a]	
DMT-Vorbehandlung		**Nein**		**Ja**	
Veränderung im LCVA	p = 0,012	= / = [a]		= / = [a]	
Gesundheitsbezogene Lebensqualität (MSIS-29)					
Alter		**18 – 30 J.**	**31 – 40 J.**	**41 – 55 J.**	**> 55 J.**
Physical Impact Scale	p = 0,023	= / + [a]	= / = [a]	+ / = [a]	= / = [a]
Psychological Impact Scale	p = 0,044	= / = [a]	= / = [a]	= / = [a]	= / = [a]
Gesundheitsbezogene Lebensqualität (EQ-5D)					
Moderater/schwerer Erkrankungsverlauf		**Nein**		**Ja**	
EQ-5D-VAS	p = 0,036	= / = [a]		= / = [a]	
Unerwünschte Ereignisse					
Alter		**18 – 30 J.**	**31 – 40 J.**	**41 – 55 J.**	**> 55 J.**
SUE (Zeit bis zum Ereignis)	p = 0,040	=	=	=	–
Geschlecht		**Weiblich**		**Männlich**	
Periphere Ödeme	p = 0,006	–		=	
Bradyarrhythmien/ Bradykardien	p = 0,031	=		–	
DMT-Vorbehandlung		**Nein**		**Ja**	
UE (ohne krankheitsspezifische UE)	p = 0,036	–		=	
Herpes-simplex-Infektionen	p = 0,014	+		=	
Atemwegserkrankungen	p = 0,019	+		=	
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**	
Malignitäten	p = 0,044	=		+	
Bradyarrhythmien/ Bradykardien	p = 0,049	=		–	
Schnell fortschreitende Erkrankung		**Nein**		**Ja**	
UE ohne krankheitsspezifische UE (Zeit bis zum Ereignis)	p = 0,050	–		=	
Moderater/schwerer Erkrankungsverlauf		**Nein**		**Ja**	
Herpes-simplex-Infektionen	p = 0,046	=		=	

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen			
Läsionslast					
Alter		**18 – 30 J.**	**31 – 40 J.**	**41 – 55 J.**	**> 55 J.**
T2-Läsionsvolumen	p < 0,001	+ / + [a]	+ / + [a]	+ / + [a]	= / + [a]
DMT-Vorbehandlung		**Nein**		**Ja**	
T2-Läsionsvolumen	p = 0,009	= / + [a]		+ / + [a]	
Neue/größere T2-Läsionen	p = 0,030	+ / + [a]		+ / + [a]	
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**	
Neue/größere T2-Läsionen	p = 0,010	+ / + [a]		+ / + [a]	
Moderater/schwerer Erkrankungsverlauf		**Nein**		**Ja**	
T2-Läsionsvolumen	p = 0,035	= / = [a]		+ / + [a]	
Region		**Europa**		**Andere Regionen**	
T2-Läsionsvolumen	p = 0,009	+/ + [a]		= / + [a]	
Neue/größere T2-Läsionen	p = 0,045	+ / + [a]		+ / + [a]	
Gd-anreichernde T1-Läsionen	p = 0,001	+ / + [a]		+ / = [a]	

Dargestellt sind die Ergebnisse von Subgruppenanalysen nur für diejenigen Subgruppenmerkmale, bei denen der Interaktionstest einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Subgruppenmerkmal ergab.

a: Monat 12 / Monat 24

\+ Siponimod signifikant überlegen ($p < 0,05$).

= Kein Unterschied zwischen den Behandlungen ($p \geq 0,05$).

– Siponimod signifikant unterlegen ($p < 0,05$).

Tabelle 4-132: Zusammenfassung der Ergebnisse aus den Subgruppenanalysen (Subpopulation B)

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen		
Behinderungsprogression (EDSS)				
Geschlecht		**Weiblich**		**Männlich**
Nach 6 Monaten bestätigte Behinderungsprogression	p = 0,047	+		=
Behinderungsprogression (MSFC)				
Schnell fortschreitende Erkrankung		**Nein**		**Ja**
Veränderung im T25W	p = 0,024	= / = [a]		= / = [a]
Schweregrad der kognitiven Dysfunktion (SDMT)				
Alter		**18– 40 J.**	**41 – 55 J.**	**> 55**
Inzidenz der Verschlechterung	p = 0,034	=	=	=

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen		
Schweregrad der kognitiven Dysfunktion (BVMT-R)				
Geschlecht		**Weiblich**		**Männlich**
Delayed Recall Score	p = 0,036	= / = [a]		– / = [a]
Total Recall Score	p = 0,014	= / = [a]		– / = [a]
Alter		**18– 40 J.**	**41 – 55 J.**	**> 55 J.**
Total Recall Score	p = 0,016	= / – [a]	= / = [a]	= / = [a]
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**
Total Recall Score	p = 0,034	= / = [a]		= / + [a]
Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)				
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**
Veränderung im LCVA	p = 0,038	= / = [a]		= / – [a]
Gesundheitsbezogene Lebensqualität (EQ-5D)				
Schnell fortschreitende Erkrankung		**Nein**		**Ja**
EQ-5D-Index	p = 0,037	= / = [a]		= / + [a]
Unerwünschte Ereignisse				
Geschlecht		**Weiblich**		**Männlich**
UE	p = 0,028	–		=
UE (Zeit bis zum Ereignis)	p = 0,028	–		=
UE ohne krankheitsspezifische UE (Zeit bis zum Ereignis)	p = 0,033	–		=
Hypertonie	p = 0,050	=		=
DMT-Vorbehandlung		**Nein**		**Ja**
Atemwegserkrankungen	p = 0,038	=		=
Läsionslast				
DMT-Vorbehandlung		**Nein**		**Ja**
Neue/größere T2-Läsionen	p < 0,001	+ / + [a]		= / + [a]
Gd-anreichernde T1-Läsionen	p = 0,006	+ / + [a]		+ / = [a]
IFN-beta-1b-Vorbehandlung		**Nein**		**Ja**
Gd-anreichernde T1-Läsionen	p = 0,027	+ / + [a]		= / = [a]
Region		**Europa**		**Andere Regionen**
Neue/größere T2-Läsionen	p < 0,001	+ / + [a]		= / = [a]
Gd-anreichernde T1-Läsionen	p < 0,001	+ / + [a]		= / = [a]

Dargestellt sind die Ergebnisse von Subgruppenanalysen nur für diejenigen Subgruppenmerkmale, bei denen der Interaktionstest einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Subgruppenmerkmal ergab.

a: Monat 12 / Monat 24

\+ Siponimod signifikant überlegen (p < 0,05).

Endpunkt/ Subgruppenmerkmal	Interaktionstest	Subgruppen
=	Kein Unterschied zwischen den Behandlungen (p > 0,05).	
–	Siponimod signifikant unterlegen (p < 0,05).	

Die Ergebnisse aller Subgruppenanalysen, unabhängig vom jeweiligen Ergebnis des Interaktionstests, können im Anhang 1 bzw. Anhang 2 zum Modul 4 eingesehen werden.

4.3.1.3.15.1 Endpunkt *Bestätigte Krankheitsschübe* – RCT (Gesamtpopulation)

In der Gesamtpopulation ergab der Interaktionstest für den Endpunkt *Jährliche Rate bestätigter Krankheitsschübe* einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Geschlecht. Weibliche wie männliche Patienten zeigten unter Siponimod jeweils eine statistisch signifikante Reduktion der jährlichen Schubrate (Tabelle 4-133). Von einer Modifikation des Behandlungseffekts wird daher nicht ausgegangen.

Tabelle 4-133: Ergebnisse für den Endpunkt *Bestätigte Krankheitsschübe (jährliche Rate bestätigter Schübe)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Rate Ratio [95%-KI]**	**p-Wert**
Jährliche Rate bestätigter Schübe				
Interaktionstest[a]				0,037
Weiblich				
N'	643	315		
Anzahl der Patienten mit Schüben	83	63		
Anzahl der Schübe	99	85		
Jährliche Schubrate [95%-KI]	0,09 [0,07; 0,12]	0,16 [0,12; 0,22]	0,55 [0,39; 0,77]	< 0,001
Männlich				
N'	418	213		
Anzahl der Patienten mit Schüben	30	37		
Anzahl der Schübe	35	56		
Jährliche Schubrate [95%-KI]	0,05 [0,03; 0,07]	0,16 [0,11; 0,23]	0,29 [0,18; 0,47]	< 0,001

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
a: Aus dem negativen Binomialmodell; Beleg einer Interaktion p < 0,05.
Quelle: Anhang 1 zu Modul 4, Tabelle 3.1.3.

4.3.1.3.15.2Endpunkt *Behinderungsprogression (MSFC)* – RCT (Gesamtpopulation)

In der Gesamtpopulation ergab der Interaktionstest für den Endpunkt *Behinderungsprogression (MSFC)* einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Geschlecht in der jeweiligen Veränderung im MSFC-z, im PASAT-3 und im T25W.

Im MSFC-z und im PASAT-3 zeigten nur weibliche Patienten zu Monat 12 einen statistisch signifikanten Vorteil zugunsten von Siponimod. Bei männlichen Patienten war dies nicht der Fall. Zu Monat 24 wurde in beiden Subgruppen keine statistische Signifikanz erreicht. Im T25W zeigten männliche Patienten zu Monat 12 einen statistisch signifikanten Vorteil zugunsten von BSC. Bei weiblichen Patienten war dies nicht der Fall. Zu Monat 24 wurde in beiden Subgruppen keine statistische Signifikanz erreicht (Tabelle 4-134).

Da kein konsistenter Beleg beobachtet werden konnte, wird im Endpunkt *Behinderungsprogression (MSFC)* nicht von einer Modifikation des Behandlungseffekts durch das Geschlecht ausgegangen.

Tabelle 4-134: Ergebnisse für den Endpunkt *Behinderungsprogression (MSFC)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung im MSFC-z				
Interaktionstest[c]				0,004
Weiblich				
N'	578	295		
Mittelwert zu Studienbeginn (SD)	0,06 (0,71)	0,05 (0,66)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,03 (0,02)	-0,06 (0,03)	0,09 [0,03; 0,15]	0,006
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,05 (0,03)	-0,07 (0,04)	0,03 [-0,07; 0,12]	0,574
Männlich				
N'	394	201		
Mittelwert zu Studienbeginn (SD)	0,01 (0,61)	0,04 (0,62)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,05 (0,02)	0,03 (0,03)	-0,07 [-0,15; 0,00]	0,052
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,12 (0,03)	-0,06 (0,05)	-0,05 [-0,17; 0,06]	0,354

Veränderung im PASAT-3				
Interaktionstest[b]				0,012
Weiblich				
N'	603	303		
Mittelwert zu Studienbeginn (SD)	38,80 (14,72)	38,98 (13,68)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	2,55 (0,39)	0,99 (0,55)	1,56 [0,23; 2,88]	0,021
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	3,48 (0,54)	2,13 (0,74)	1,35 [-0,45; 3,15]	0,141
Männlich				
N'	408	206		
Mittelwert zu Studienbeginn (SD)	40,48 (13,84)	40,16 (14,14)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	1,82 (0,47)	2,99 (0,67)	-1,17 [-2,77; 0,44]	0,153
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	2,30 (0,61)	2,62 (0,93)	-0,32 [-2,51; 1,86]	0,771
Veränderung im T25W				
Interaktionstest[b]				0,011
Weiblich				
N'	622	315		
Mittelwert zu Studienbeginn (SD)	17,77 (19,93)	16,76 (18,96)		
Monat 12: Adjustierte mittlere Veränderung (SE)[b]	3,11 (0,86)	5,77 (1,22)	-2,66 [-5,60; 0,28]	0,076
Monat 24: Adjustierte mittlere Veränderung (SE)[b]	8,15 (1,39)	7,90 (1,97)	0,25 [-4,48; 4,98]	0,918
Männlich				
N'	423	217		
Mittelwert zu Studienbeginn (SD)	14,37 (17,24)	13,04 (15,76)		
Monat 12: Adjustierte mittlere Veränderung (SE)[b]	6,00 (1,05)	2,01 (1,48)	3,99 [0,43; 7,55]	0,028
Monat 24: Adjustierte mittlere Veränderung (SE)[b]	11,29 (1,62)	5,74 (2,39)	5,56 [-0,11; 11,22]	0,055

KI: Konfidenzintervall; N: Patienten gesamt; MMRM: Mixed Model for Repeated Measures; MSFC: Multiple Sclerosis Functional Composite; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; PASAT: Paced Auditory Serial Addition Test; SD: Standard Deviation ;SE: Standard Error ; T25W: Timed 25 Foot Walking Test.

a: Positive Werte zeigen eine Verbesserung an.

b: Negative Werte zeigen eine Verbesserung an.

c: Aus dem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabellen 6.1.3, 6.2.3 und 6.4.3.

4.3.1.3.15.3 Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* – RCT (Gesamtpopulation)

In der Gesamtpopulation ergab der Interaktionstest für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* in der Inzidenz der Verschlechterung des SDMT um mindestens vier Punkte einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Geschlecht.

Patientinnen zeigten statistisch signifikant weniger häufig eine relevante Verschlechterung unter Siponimod verglichen mit BSC, bei männlichen Patienten war kein Unterschied zwischen den Behandlungsgruppen festgestellt worden (Tabelle 4-135). Im Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* wird daher nicht von einer Modifikation des Behandlungseffekts ausgegangen.

Tabelle 4-135: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 1.099)	**BSC (N =546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Inzidenz der Verschlechterung im SDMT nach 12 Monaten					
Interaktionstest[b]				0,003	
Weiblich					
N'	664	323			
Patienten mit Verschlechterung um ≥ 4 Punkte – n (%)	166 (25,0)	124 (38,4)	0,53 [0,40; 0,71] < 0,001	0,65 [0,54; 0,79] < 0,001	-0,13 [-0,20; -0,07] < 0,001
Männlich					
N'	435	223			
Patienten mit Verschlechterung um ≥ 4 Punkte – n (%)	125 (28,7)	61 (27,4)	1,07 [0,75; 1,54] 0,710	1,05 [0,81; 1,36] 0,710	0,01 [-0,06; 0,09] 0,708

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 7.3.3.

4.3.1.3.15.4 Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* – RCT (Gesamtpopulation)

In der Gesamtpopulation ergab der Interaktionstest für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* jeweils einen Beleg für eine mögliche Modifikation des Behandlungseffekts in der Veränderung der LCVA durch das Geschlecht sowie durch eine DMT-Vorbehandlung.

In der Veränderung der LCVA wurde weder bei weiblichen noch bei männlichen Patienten ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen festgestellt (gilt jeweils für beide Auswertungszeitpunkte) (Tabelle 4-136).

Auch in der Auswertung nach DMT-Vorbehandlung konnte in keiner der Subgruppen ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen festgestellt werden (gilt jeweils für beide Auswertungszeitpunkte) (Tabelle 4-137).

Im Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* wird daher weder von einer Modifikation des Behandlungseffekts durch das Geschlecht noch durch eine DMT-Vorbehandlung ausgegangen.

Tabelle 4-136: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung der LCVA				
Interaktionstest[b]				0,016
Weiblich				
N'	569	287		
Mittelwert zu Studienbeginn (SD)	0,38 (0,26)	0,37 (0,27)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,02 (0,01)	0,02 (0,01)	0,01 [-0,02; 0,04]	0,609
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,02 (0,01)	0,04 (0,02)	-0,02 [-0,07; 0,03]	0,458
Männlich				
N'	371	199		
Mittelwert zu Studienbeginn (SD)	0,35 (0,26)	0,34 (0,25)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,02 (0,01)	0,01 (0,02)	0,01 [-0,03; 0,05]	0,484

Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,02 (0,02)	-0,02 (0,03)	0,04 [-0,02; 0,10]	0,162

KI: Konfidenzintervall; LCVA: Low Contrast Visual Acuity; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation ; SE: Standard Error .

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 8.3.3.

Tabelle 4-137: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung der LCVA				
Interaktionstest[b]				0,012
DMT-Vorbehandlung: Nein				
N'	206	100		
Mittelwert zu Studienbeginn (SD)	0,36 (0,26)	0,38 (0,28)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,01 (0,02)	0,01 (0,02)	0,00 [-0,06; 0,05]	0,880
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,00 (0,02)	0,00 (0,03)	0,00 [-0,07; 0,08]	0,950
DMT-Vorbehandlung: Ja				
N'	734	386		
Mittelwert zu Studienbeginn (SD)	0,37 (0,26)	0,36 (0,26)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,02 (0,01)	0,00 (0,01)	0,02 [-0,01; 0,04]	0,269
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,01 (0,01)	0,01 (0,02)	0,00 [-0,04; 0,05]	0,861

DMT: Disease Modifying Therapy ; KI: Konfidenzintervall; LCVA: Low Contrast Visual Acuity; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation ; SE: Standard Error .

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 8.3.4.

4.3.1.3.15.5 Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* – RCT (Gesamtpopulation)

In der Gesamtpopulation ergab der Interaktionstest für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Alter.

In der Veränderung des MSIS-29 wurden in der Auswertung der Physical Impact Scale nur in der Altersgruppe 18 bis 30 Jahre zu Monat 24 und in der Altersgruppe 41 bis 55 Jahre zu Monat 12 ein statistisch signifikanter Behandlungseffekt beobachtet, jeweils zugunsten von Siponimod. In allen anderen Subgruppen wurde zu keinem Zeitpunkt ein statistisch signifikanter Behandlungsunterschied festgestellt. In der Auswertung der Psychological Impact Scale wurde in keiner Subgruppe und zu keinem Zeitpunkt ein statistisch signifikanter Behandlungsunterschied festgestellt (Tabelle 4-138).

Im Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* wird daher nicht von einer Modifikation des Behandlungseffekts durch das Alter ausgegangen.

Tabelle 4-138: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29)* aus RCT mit dem zu bewertenden Arzneimittel nach Alter (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung in der MSIS-29 – Physical Impact Scale				
Interaktionstest[b]				0,023
Alter 18 – 30 Jahre				
N'	22	10		
Mittelwert zu Studienbeginn (SD)	53,94 (27,03)	52,50 (17,24)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-3,31 (3,59)	9,37 (5,28)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-2,60 (3,95)	11,08 (6,58)	-13,68 [-28,64; 1,28]	0,073
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-11,34 (4,63)	12,42 (6,92)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-11,43 (6,17)	16,06 (8,39)	-27,50 [-47,86; -7,13]	0,008
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-10,86 (7,19)	23,48 (11,87)		
Alter 31 – 40 Jahre				
N'	147	86		
Mittelwert zu Studienbeginn (SD)	53,24 (22,88)	51,57 (21,13)		

Monat 6: Adjustierte mittlere Veränderung (SE)[a]	2,51 (1,42)	2,24 (1,83)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	2,01 (1,51)	4,73 (1,99)	-2,72 [-7,54; 2,10]	0,268
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	4,07 (1,69)	6,34 (2,14)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	3,77 (2,02)	5,57 (2,67)	-1,80 [-8,31; 4,70]	0,587
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	1,66 (2,82)	10,71 (3,63)		
Alter 41 – 55 Jahre				
N'	653	307		
Mittelwert zu Studienbeginn (SD)	50,71 (21,31)	50,11 (21,72)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	0,13 (0,74)	1,47 (1,00)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	1,46 (0,78)	4,43 (1,08)	-2,97 [-5,43; -0,51]	0,018
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	3,01 (0,84)	4,69 (1,20)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	4,11 (0,98)	5,42 (1,40)	-1,31 [-4,52; 1,91]	0,426
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	5,77 (1,29)	4,52 (2,05)		
Alter > 55 Jahre				
N'	179	105		
Mittelwert zu Studienbeginn (SD)	50,99 (20,54)	50,73 (21,03)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-0,90 (1,30)	0,38 (1,68)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,71 (1,41)	2,54 (1,83)	-1,83 [-6,23; 2,57]	0,414
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	2,28 (1,55)	4,47 (2,01)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	1,05 (1,85)	0,41 (2,45)	0,64 [-5,30; 6,58]	0,833
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-0,43 (2,30)	5,40 (3,18)		
Veränderung in der MSIS-29 – Psychological Impact Scale				
Interaktionstest[b]				0,044
Alter 18 – 30 Jahre				
N'	22	10		
Mittelwert zu Studienbeginn (SD)	50,17 (29,91)	43,33 (20,59)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-8,33 (3,83)	5,23 (5,63)		

Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-4,81 (4,32)	5,39 (7,28)	-10,20 [-26,70; 6,31]	0,226
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-13,68 (5,31)	6,41 (7,92)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-4,95 (7,53)	11,64 (10,04)	-16,59 [-41,17; 8,00]	0,186
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-18,41 (8,10)	16,31 (13,60)		
Alter 31 – 40 Jahre				
N'	147	85		
Mittelwert zu Studienbeginn (SD)	40,19 (25,34)	39,69 (23,86)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	0,19 (1,52)	1,84 (1,96)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,52 (1,64)	4,54 (2,18)	-4,02 [-9,29; 1,25]	0,135
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	0,24 (1,93)	5,65 (2,43)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,04 (2,40)	1,47 (3,19)	-1,43 [-9,20; 6,34]	0,718
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-2,83 (3,23)	8,57 (4,15)		
Alter 41 – 55 Jahre				
N'	650	306		
Mittelwert zu Studienbeginn (SD)	37,18 (22,96)	37,61 (23,48)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-0,65 (0,79)	1,23 (1,07)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,53 (0,85)	2,02 (1,18)	-1,48 [-4,18; 1,21]	0,280
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	2,02 (0,95)	1,68 (1,37)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	2,18 (1,15)	2,10 (1,65)	0,08 [-3,75; 3,90]	0,968
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	0,51 (1,46)	0,14 (2,35)		
Alter > 55 Jahre				
N'	179	105		
Mittelwert zu Studienbeginn (SD)	37,97 (21,62)	39,47 (25,42)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-0,80 (1,38)	-0,94 (1,79)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-1,14 (1,54)	0,32 (2,00)	-1,46 [-6,28; 3,35]	0,550
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-1,42 (1,77)	1,78 (2,29)		

Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-3,19 (2,19)	-1,47 (2,93)	-1,72 [-8,81; 5,38]	0,635
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-3,49 (2,59)	3,44 (3,61)		

KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MSIS: Multiple Sclerosis Impact Scale; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation ; SE: Standard Error .

a: Negative Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabellen 10.1.2 und 10.2.2.

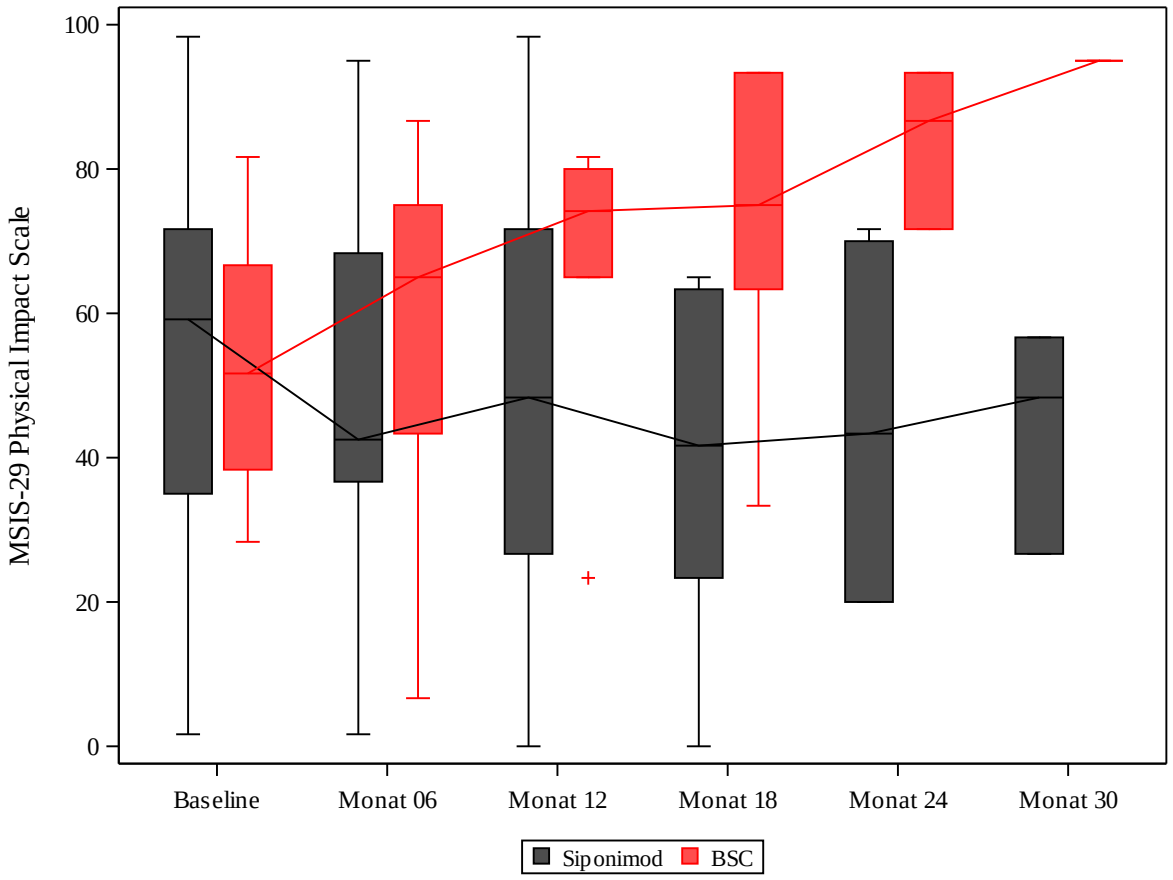


Abbildung 36: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 18 – 30 Jahre)

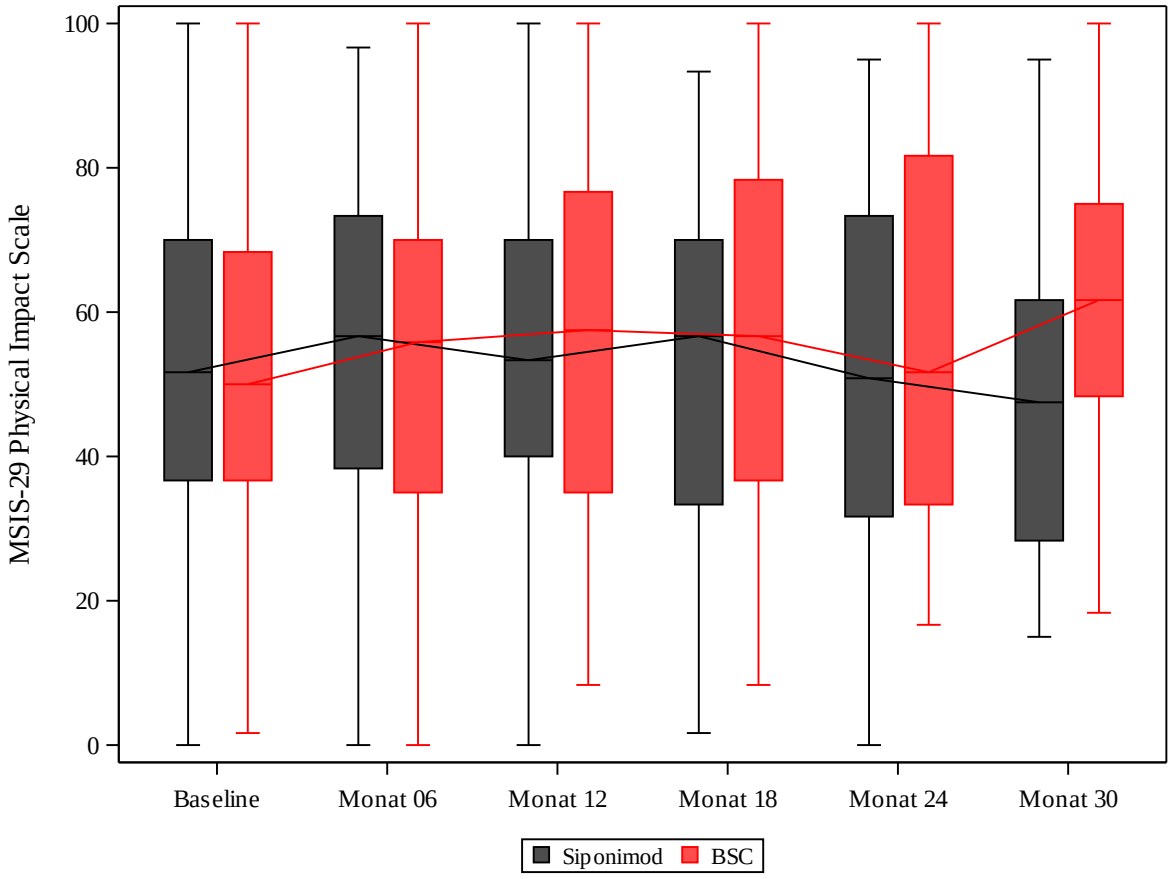


Abbildung 37: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 31 – 40 Jahre)

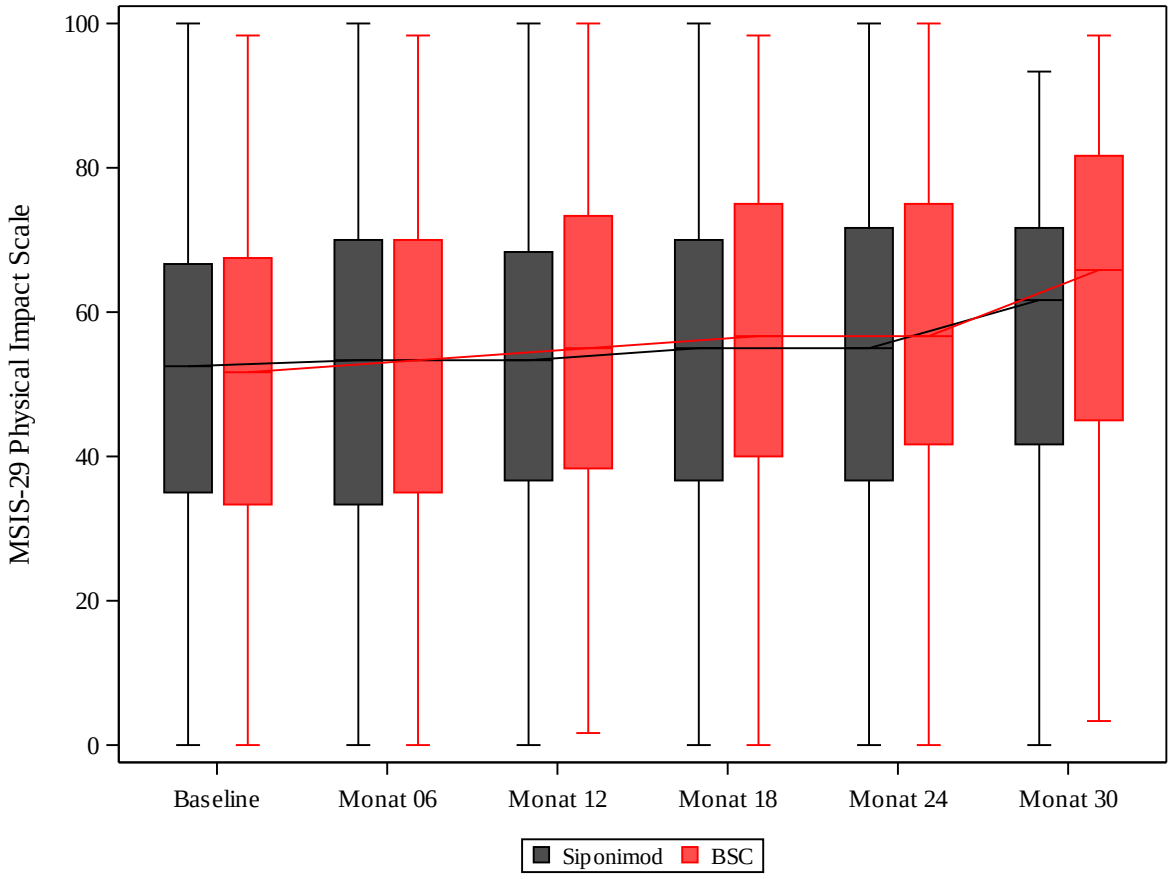


Abbildung 38: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 41 – 55 Jahre)

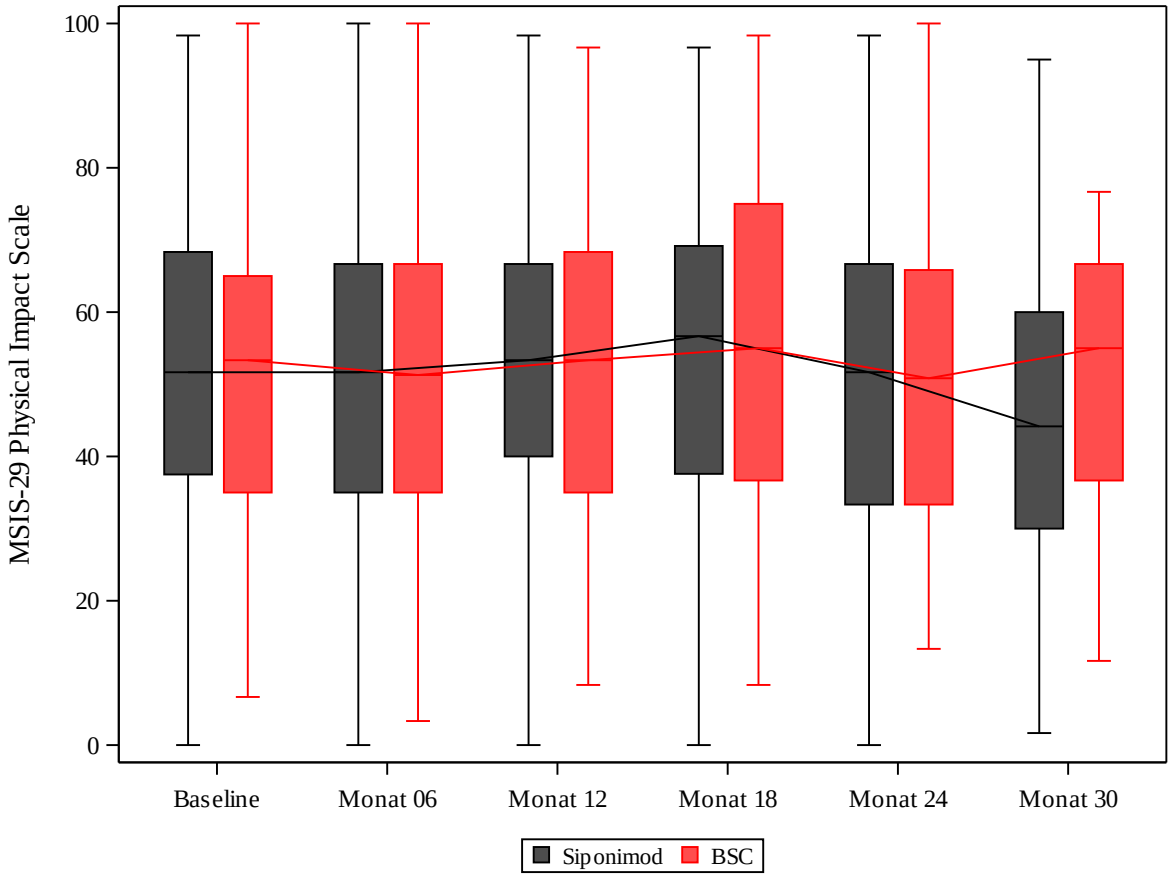


Abbildung 39: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Physical Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter > 55 Jahre)

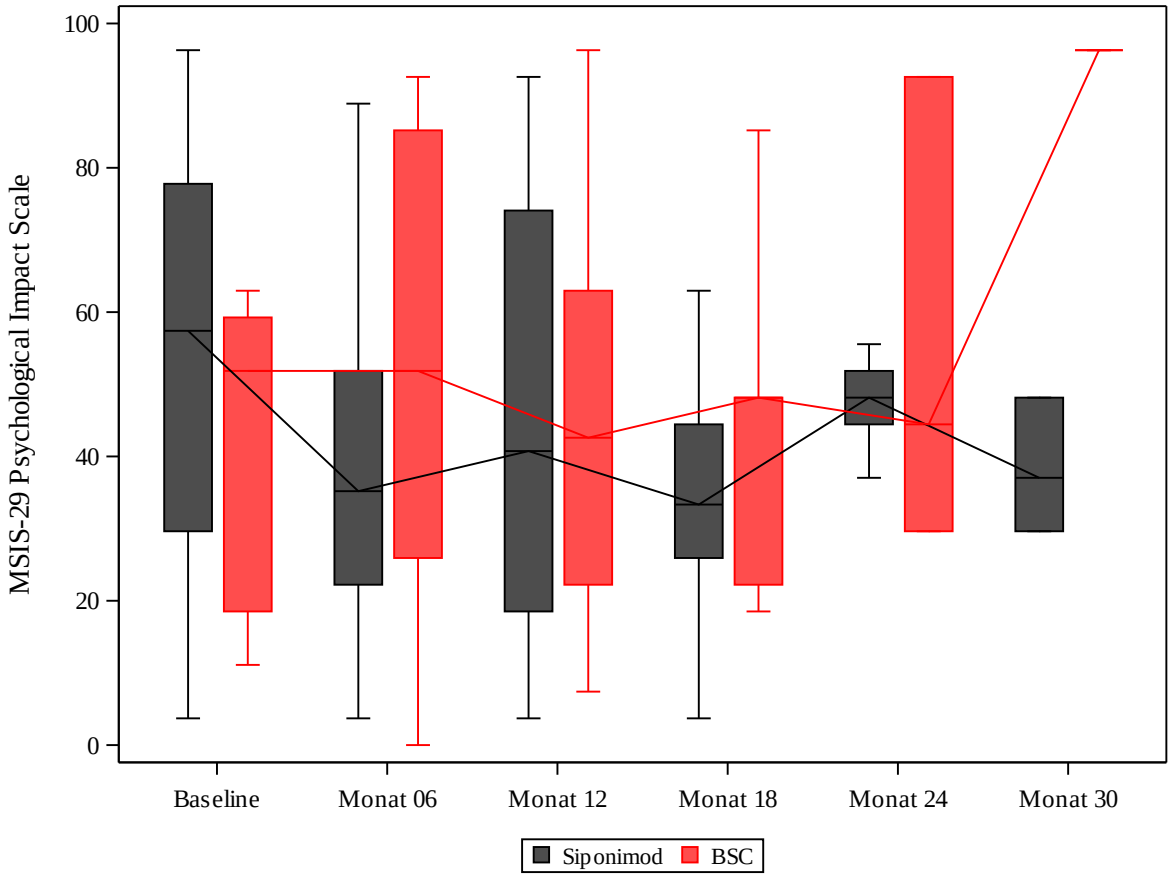


Abbildung 40: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 18 – 30 Jahre)

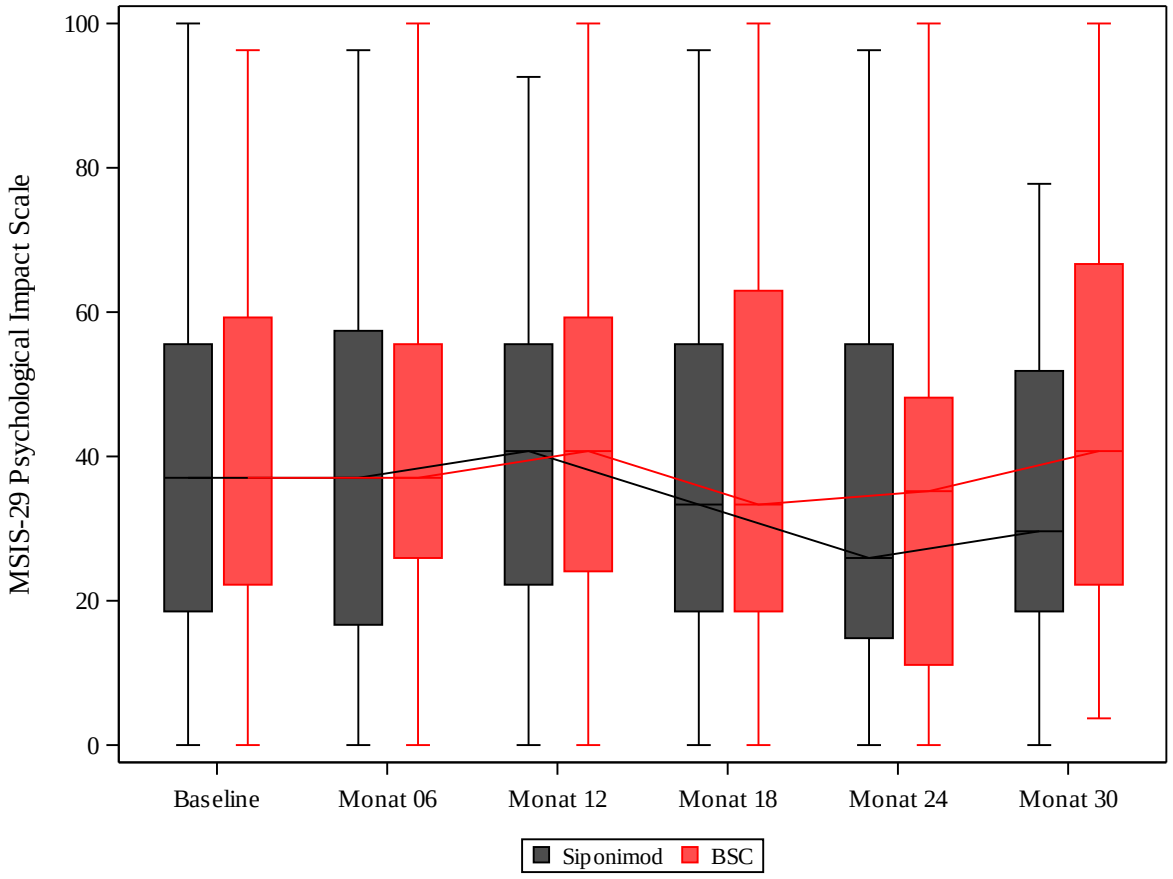


Abbildung 41: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 31 – 40 Jahre)

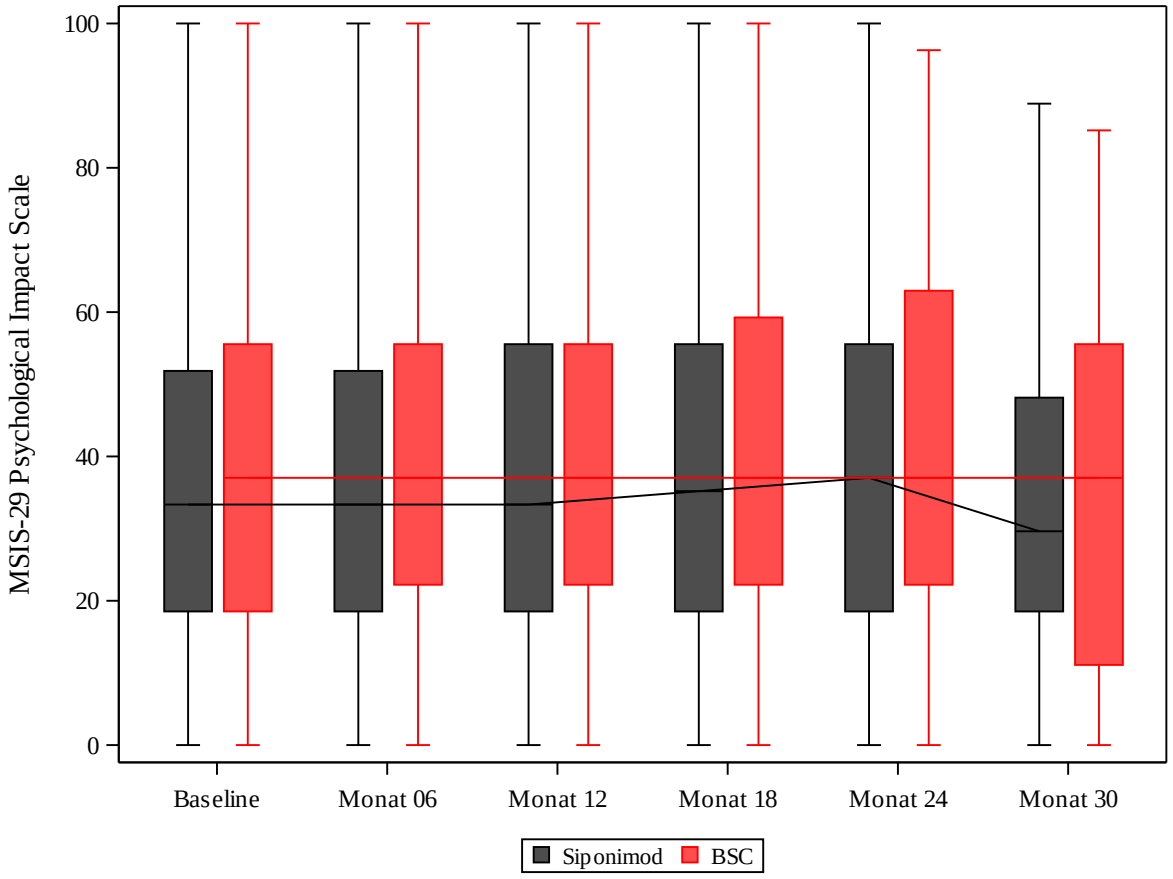


Abbildung 42: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 41 – 55 Jahre)

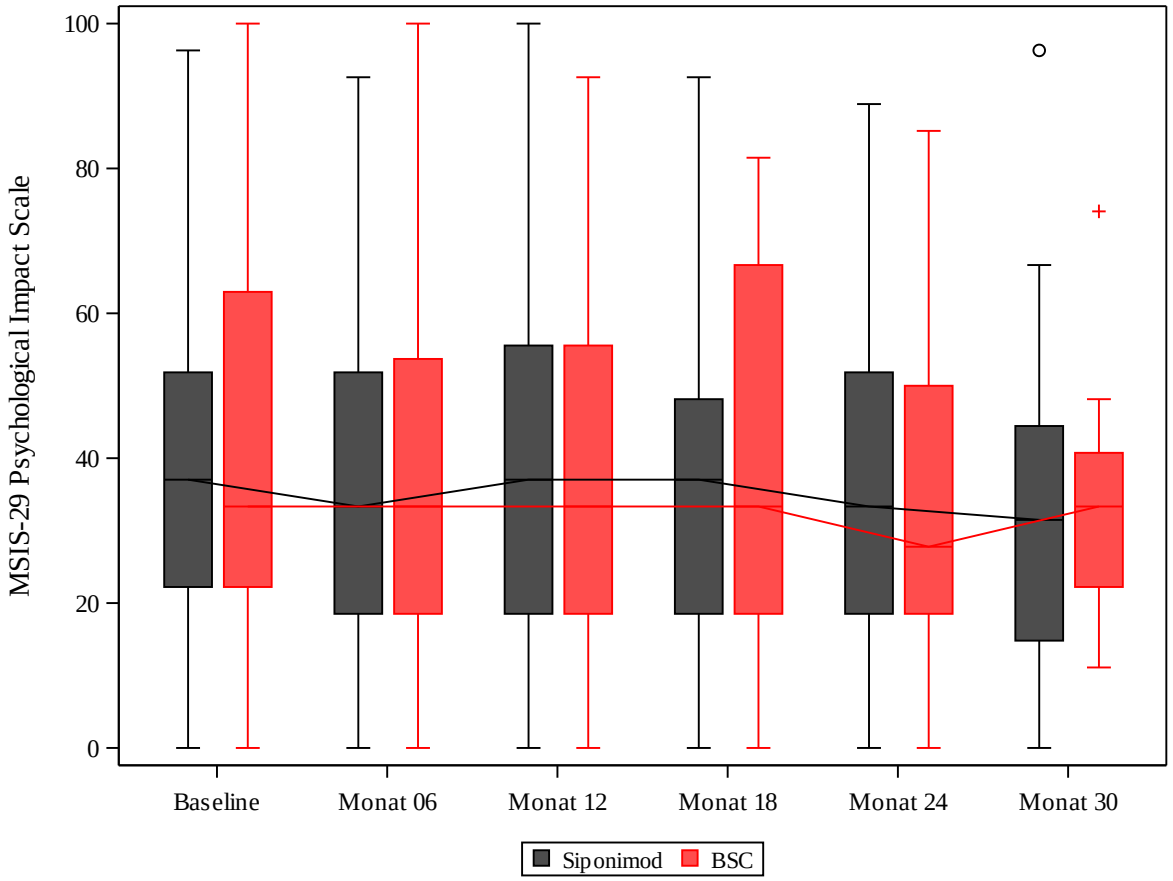


Abbildung 43: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (MSIS-29 Psychological Impact Scale)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter > 55 Jahre)

4.3.1.3.15.6 Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* – RCT (Gesamtpopulation)

In der Gesamtpopulation ergab der Interaktionstest für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* in der Auswertung der VAS einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch den Erkrankungsverlauf.

In der Auswertung der EQ-5D-VAS wurde in keiner Subgruppe und zu keinem Zeitpunkt ein statistisch signifikanter Behandlungsunterschied festgestellt (Tabelle 4-139).

Im Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* wird daher nicht von einer Modifikation des Behandlungseffekts durch den Erkrankungsverlauf ausgegangen.

Tabelle 4-139: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* aus RCT mit dem zu bewertenden Arzneimittel nach Erkrankungsverlauf (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung in der EQ-5D-VAS				
Interaktionstest[b]				0,036
Moderater oder schwerer Erkrankungssverlauf: Nein				
N'	181	83		
Mittelwert zu Studienbeginn (SD)	64,10 (18,02)	61,61 (19,00)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	3,15 (1,26)	0,42 (1,85)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,36 (1,37)	-2,21 (1,99)	2,57 [-2,15; 7,30]	0,286
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	1,13 (1,63)	-3,02 (2,30)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	3,23 (2,10)	-2,71 (3,29)	5,94 [-1,72; 13,60]	0,128
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	4,59 (4,28)	-10,90 (7,93)		
Moderater oder schwerer Erkrankungssverlauf: Ja				
N'	835	431		
Mittelwert zu Studienbeginn (SD)	56,65 (18,88)	55,25 (19,15)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-1,92 (0,58)	-1,17 (0,81)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-2,60 (0,63)	-4,22 (0,89)	1,62 [-0,53; 3,76]	0,139
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-2,86 (0,74)	-3,32 (1,06)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-3,22 (0,99)	-3,06 (1,42)	-0,16 [-3,57; 3,25]	0,927
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-5,37 (1,90)	-5,23 (2,83)		

EQ-5D: EuroQol – 5 Dimensionen; KI: Konfidenzintervall; N: Patienten gesamt; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error; VAS: Visuelle Analogskala.

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 10.4.7.

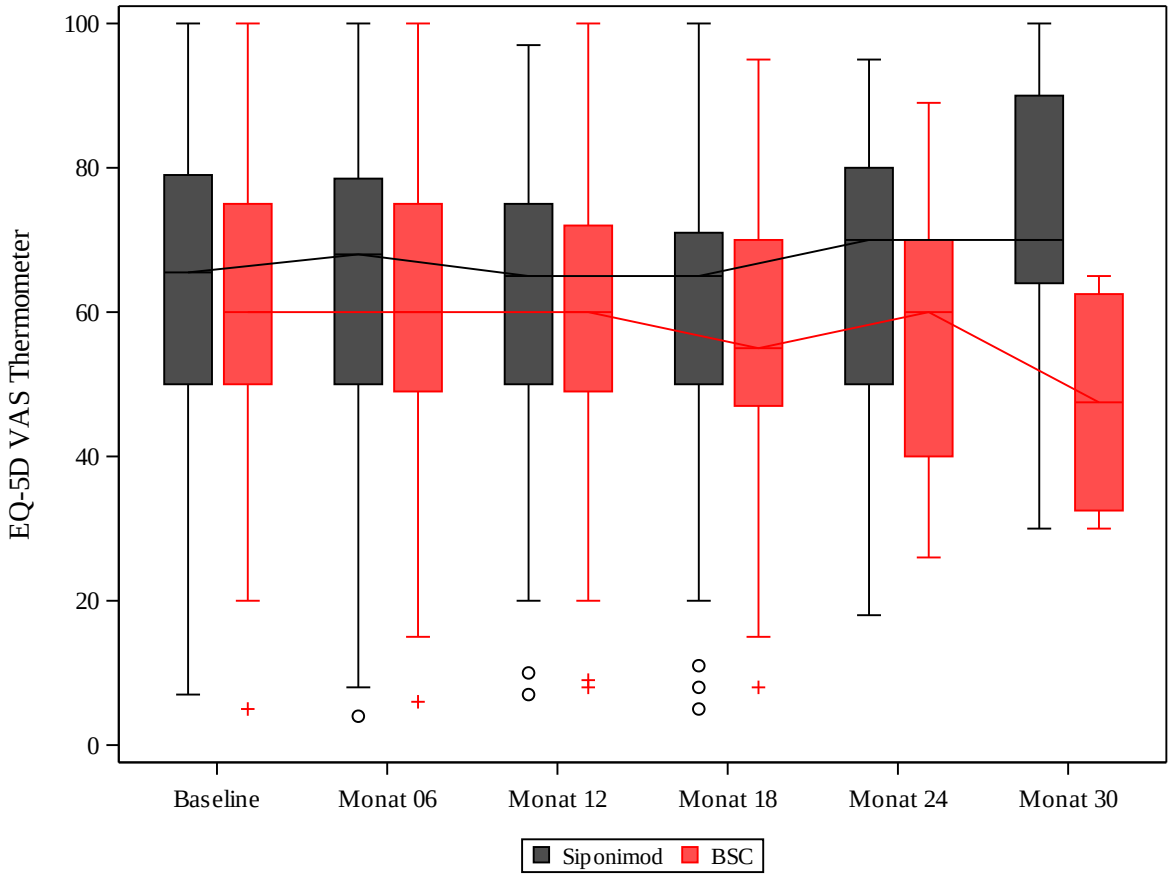


Abbildung 44: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-VAS)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – moderater/schwerer Erkrankungsverlauf: Nein)

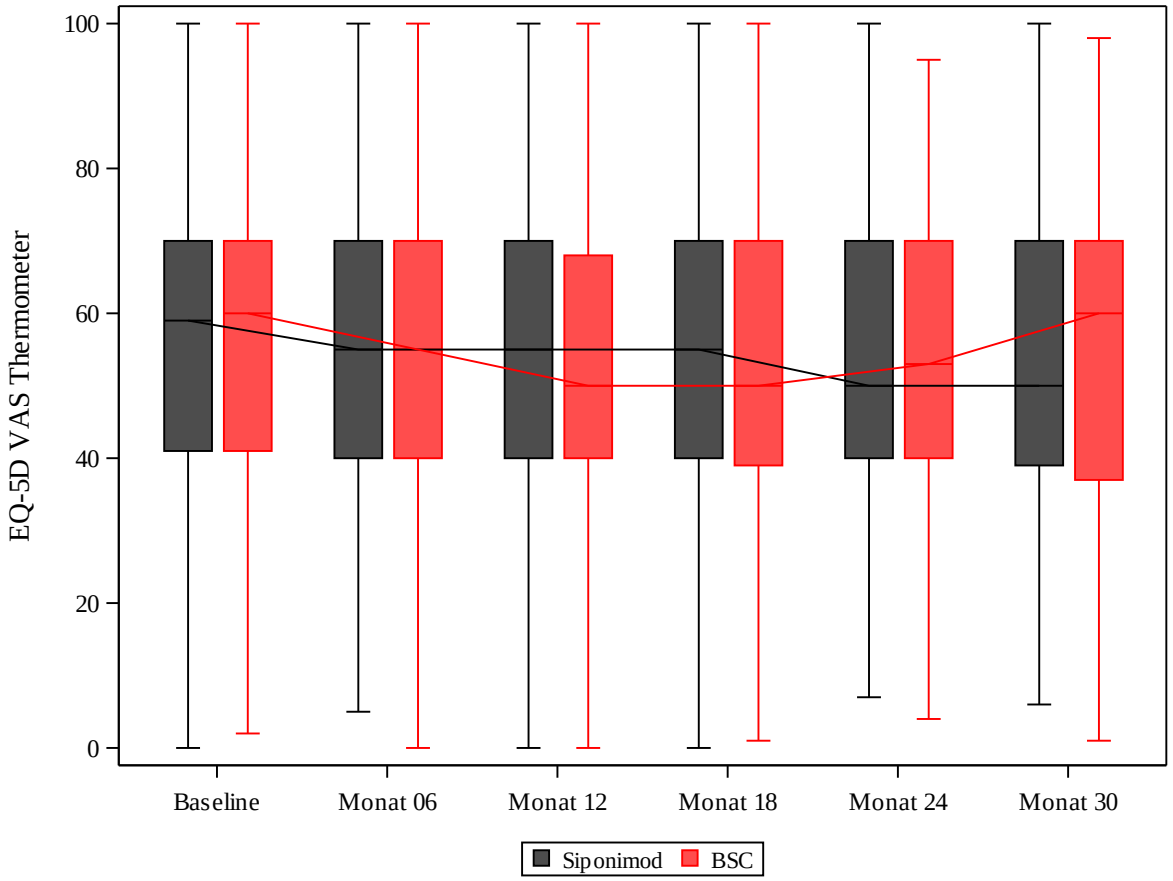


Abbildung 45: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-VAS)* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – moderater/schwerer Erkrankungsverlauf: Ja)

4.3.1.3.15.7Endpunkt *Unerwünschte Ereignisse* – RCT (Gesamtpopulation)

In der Gesamtpopulation ergaben die Interaktionstests für den Endpunkt *Unerwünschte Ereignisse* Belege für eine mögliche Modifikation des Behandlungseffekts durch eine DMT-Vorbehandlung in den Gesamtraten der UE ohne krankheitsspezifische Ereignisse sowie durch das Alter und eine unterschiedlich schnell fortschreitende Erkrankung in der Zeit bis zum ersten SUE. In der Gesamtpopulation ergaben die Interaktionstests weiterhin Belege für eine mögliche Modifikation des Behandlungseffekts durch das Geschlecht, eine DMT-Vorbehandlung, eine IFN-beta-1b-Vorbehandlung und den Erkrankungsverlauf für Unerwünschte Ereignisse von besonderem Interesse.

UE ohne Berücksichtigung krankheitsspezifischer UE (binäre Analyse)

- Mögliche Effektmodifikation durch eine DMT-Vorbehandlung: Patienten ohne DMT-Vorbehandlung zeigten unter Siponimod statistisch signifikant mehr Ereignisse. Bei Patienten mit DMT-Vorbehandlung war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-140). Von einer Modifikation des Behandlungseffekts wird nicht ausgegangen.

Tabelle 4-140: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 1.099)	**BSC (N = 546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
UE ohne Berücksichtigung krankheitsspezifischer UE					
Interaktionstest[b]				0,036	
DMT-Vorbehandlung: Nein					
N'	242	114			
UE ohne krankheits-spezifische UE – n (%)	205 (84,7)	78 (68,4)	2,56 [1,51; 4,33] <0,001	1,24 [1,08; 1,42] 0,002	0,16 [0,07; 0,26] 0,001
DMT-Vorbehandlung: Ja					
N'	857	432			
UE ohne krankheits-spezifische UE – n (%)	713 (83,2)	340 (78,7)	1,34 [1,00; 1,79] 0,049	1,06 [1,00; 1,12] 0,059	0,04 [-0,00; 0,09] 0,056

DMT: Disease Modifying Therapy ; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschte Ereignisse.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.4.1.

SUE (Zeit bis zum Ereignis)

- Mögliche Effektmodifikation durch das Alter: Lediglich in der Altersgruppe > 55 Jahre wurde ein statistisch signifikanter Effekt zugunsten von BSC festgestellt. In den anderen Altersgruppen waren jeweils keine statistisch signifikanten Behandlungseffekte zu beobachten (Tabelle 4-141).

- Mögliche Effektmodifikation durch eine unterschiedlich schnell fortschreitende Erkrankung: Patienten ohne schnell fortschreitende Erkrankung zeigten unter Siponimod statistisch signifikant mehr Ereignisse. Bei Patienten mit schnell fortschreitender Erkrankung war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-142).

Von einer Modifikation des Behandlungseffekts wird nicht ausgegangen.

Tabelle 4-141: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Alter (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**HR [95%-KI][c]**	**p-Wert**
SUE				
Interaktionstest[a]				0,040
Alter 18 – 30 Jahre				
N'	26	12		
Anzahl der Ereignisse	2	2		
Median (in Wochen) [95%-KI][b]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	91,27 [79,69; 100,0]	83,33 [62,25; 100,0]		
			0,45 [0,06; 3,16]	0,419
Alter 31 – 40 Jahre				
N'	161	91		
Anzahl der Ereignisse	12	11		
Median (in Wochen) [95%-KI][b]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	89,51 [83,60; 95,42]	83,55 [73,77; 93,32]		
			0,58 [0,26; 1,31]	0,191
Alter 41 – 55 Jahre				
N'	712	331		
Anzahl der Ereignisse	118	45		
Median (in Wochen) [95%-KI][b]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	75,52 [68,81; 82,23]	77,18 [62,26; 92,10]		
			1,20 [0,85; 1,70]	0,291
Alter > 55 Jahre				
N'	200	112		
Anzahl der Ereignisse	54	16		
Median (in Wochen) [95%-KI][b]	n. b.	n. b.		

Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	63,93 [53,48; 74,38]	79,40 [69,29; 89,50]		
			2,18 [1,25; 3,80]	0,006

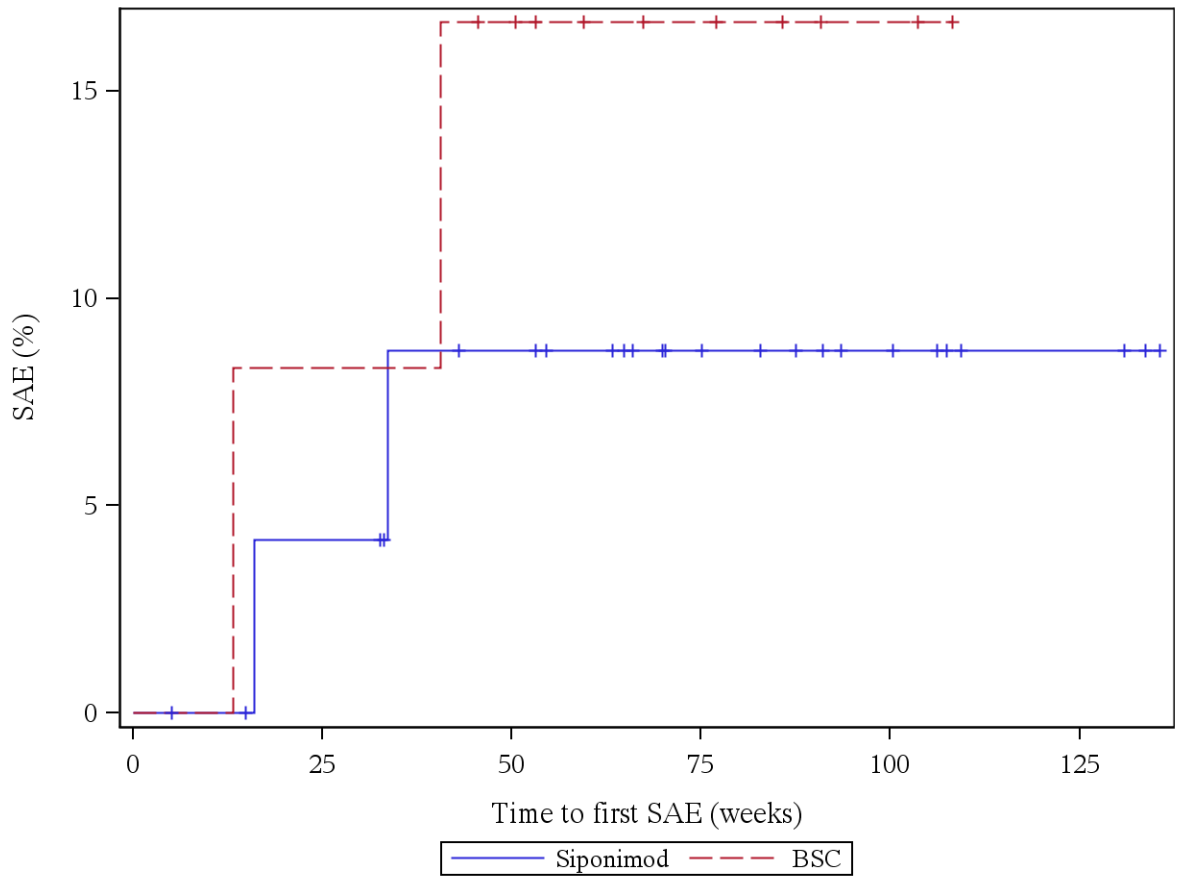
HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar; SUE: Schwerwiegendes Unerwünschtes Ereignis.

a: Aus einem Cox proportional hazards model.

b: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

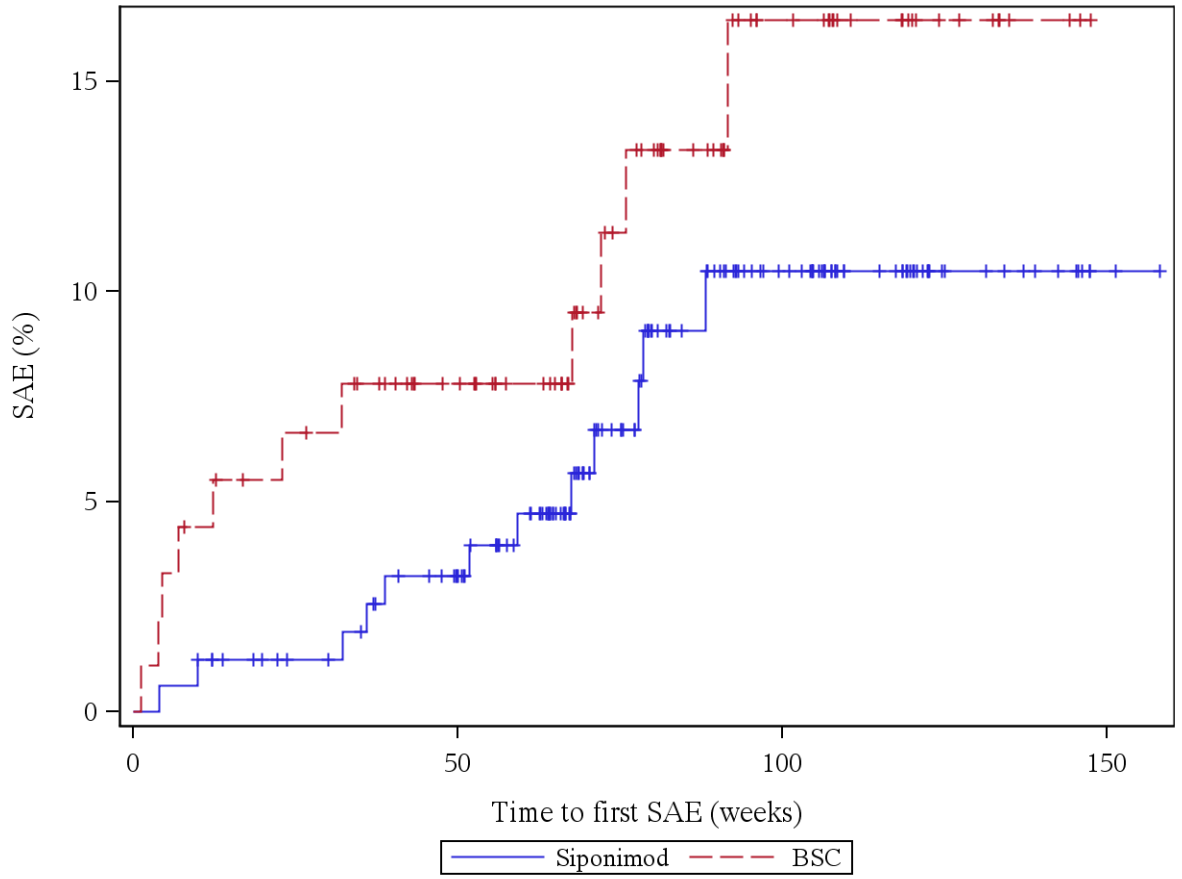
c: Cox proportional hazards model.

Quelle: Anhang 1 zu Modul 4, Tabelle 12.2.1.



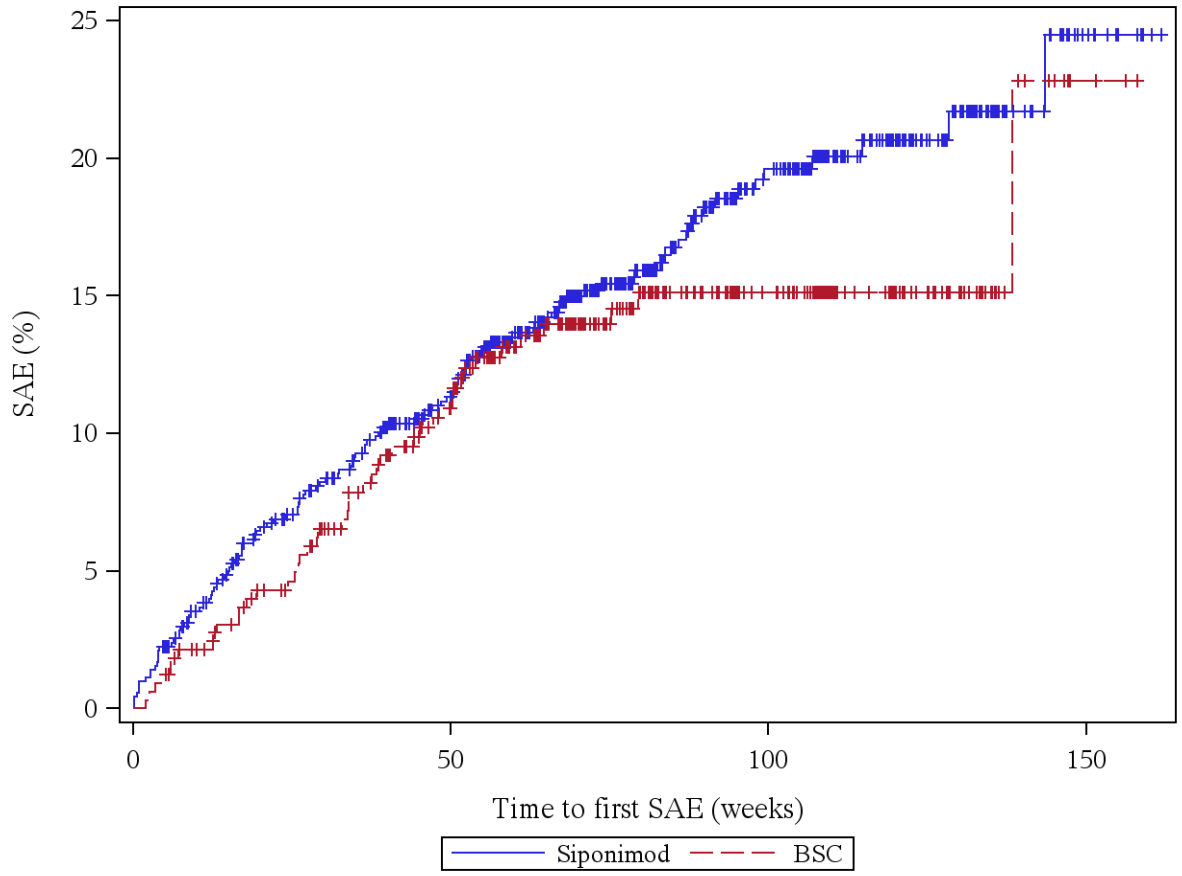
Patients at risk (Siponimod/BSC): Week 0: 26/12, Week 4: 26/12, Week 12: 25/12, Week 24: 23/11, Week 36: 20/11, Week 48: 19/9, Week 60: 17/6, Week 72: 12/5, Week 84: 10/4, Week 96: 7/2, Week 108: 4/1, Week 120: 3/0, Week 132: 2/0, Week 144: 0/0

Abbildung 46: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 18 bis 30 Jahre)



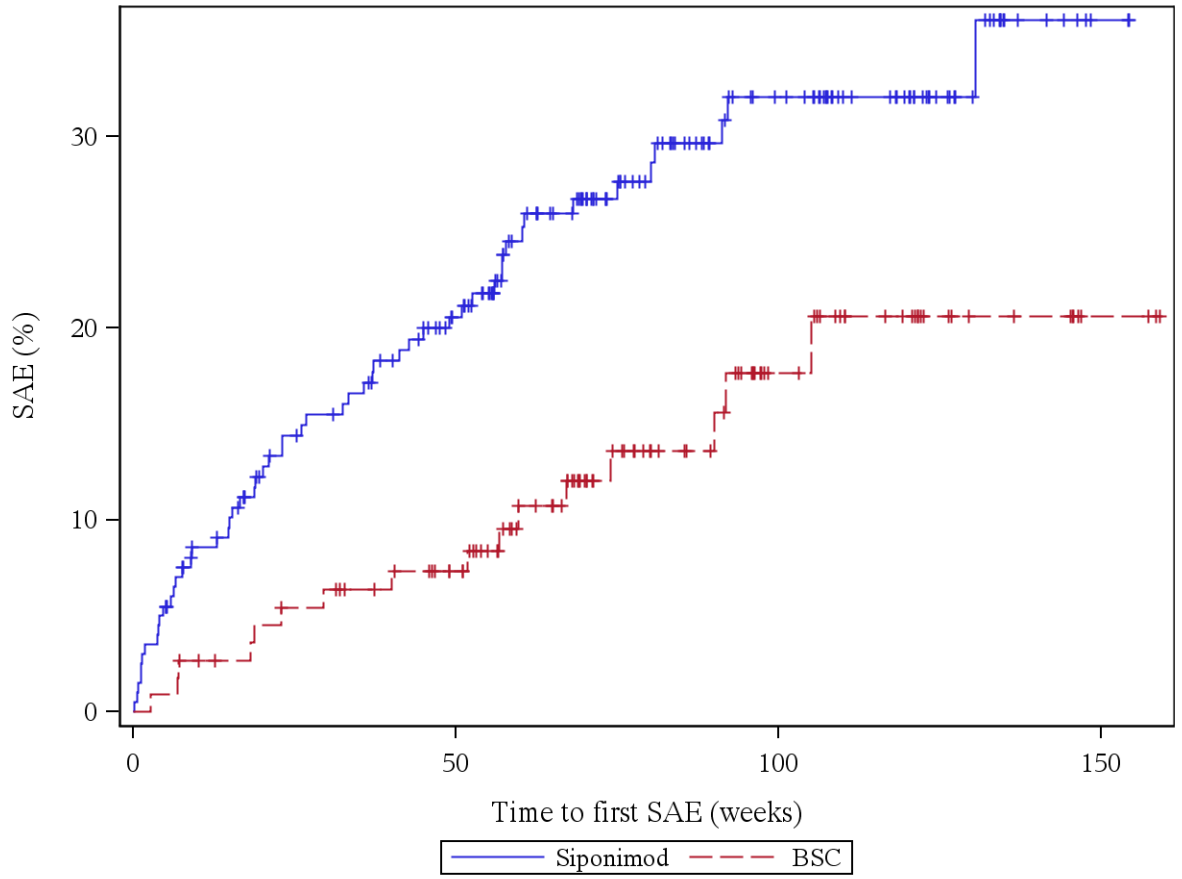
Patients at risk (Siponimod/BSC): Week 0: 161/91, Week 4: 160/89, Week 12: 158/86, Week 24: 151/82, Week 36: 147/78, Week 48: 141/70, Week 60: 124/62, Week 72: 87/47, Week 84: 65/35, Week 96: 51/22, Week 108: 34/16, Week 120: 22/10, Week 132: 11/7, Week 144: 7/3, Week 156: 1/0, Week 168: 0/0

Abbildung 47: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 31 bis 40 Jahre)



Patients at risk (Siponimod/BSC): Week 0: 712/331, Week 4: 697/328, Week 12: 665/317, Week 24: 628/301, Week 36: 595/278, Week 48: 549/252, Week 60: 483/213, Week 72: 381/169, Week 84: 297/126, Week 96: 228/98, Week 108: 164/71, Week 120: 108/44, Week 132: 60/22, Week 144: 27/8, Week 156: 6/2, Week 168: 0/0

Abbildung 48: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter 41 bis 55 Jahre)



Patients at risk (Siponimod/BSC): Week 0: 200/112, Week 4: 190/111, Week 12: 176/107, Week 24: 157/102, Week 36: 150/98, Week 48: 136/92, Week 60: 106/73, Week 72: 84/56, Week 84: 65/46, Week 96: 51/35, Week 108: 40/24, Week 120: 31/18, Week 132: 15/9, Week 144: 6/8, Week 156: 0/3

Abbildung 49: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Alter > 55 Jahre)

Tabelle 4-142: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach schnell fortschreitender Erkrankung (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**HR [95 %-KI][c]**	**p-Wert**
SUE				
Interaktionstest[a]				0,050
Schnell fortschreitende Erkrankung: Nein				
N'	835	401		
Anzahl der Ereignisse	152	51		
Median (in Wochen) [95%-KI][b]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	72,63 [65,85; 79,41]	79,47 [69,65; 89,30]		
			1,46 [1,06; 2,00]	0,020
Schnell fortschreitende Erkrankung: Ja				
N'	264	145		
Anzahl der Ereignisse	34	23		
Median (in Wochen) [95%-KI][b]	n. b.	n. b.		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	83,85 [78,60; 89,10]	81,85 [74,94; 88,77]		
			0,79 [0,46; 1,34]	0,374

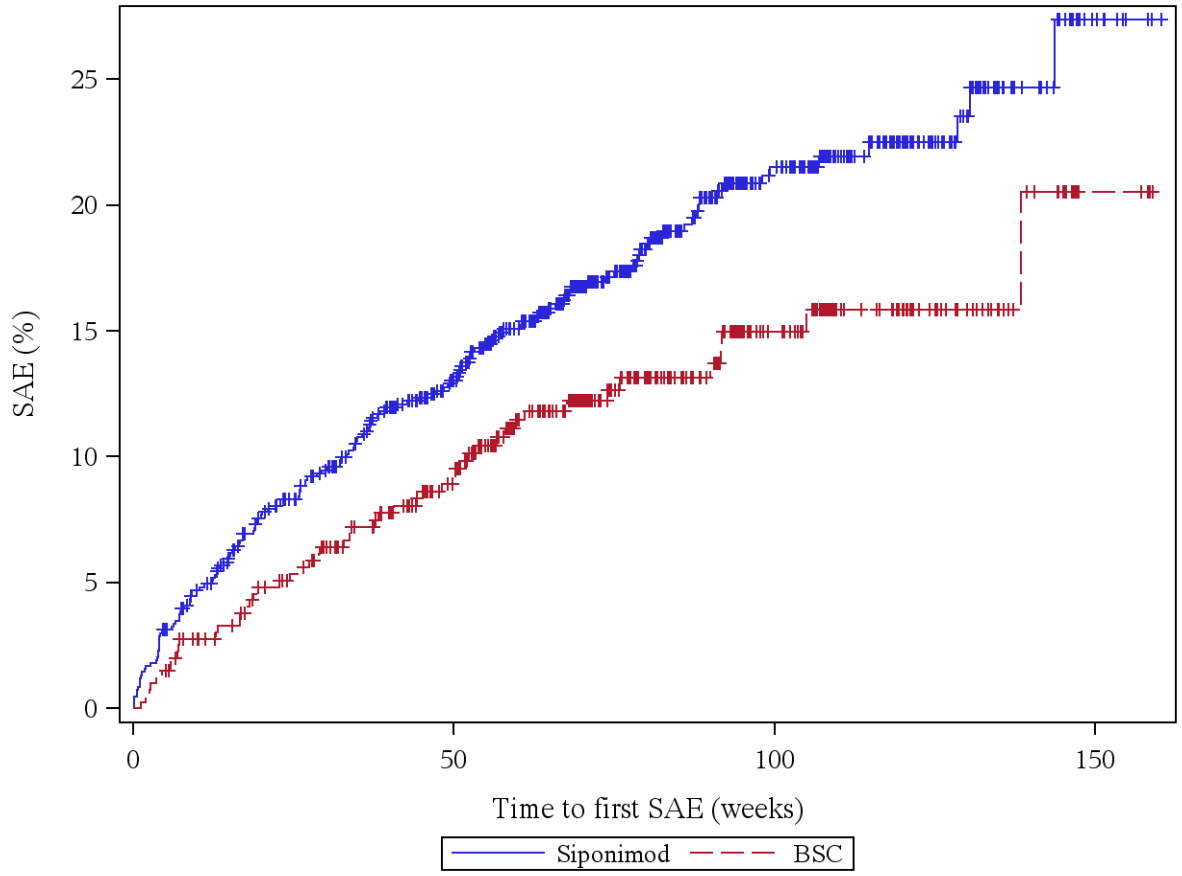
HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar; SUE: Schwerwiegendes Unerwünschtes Ereignis.

a: Aus einem Cox proportional hazards model; Beleg einer Interaktion $p < 0,05$.

b: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

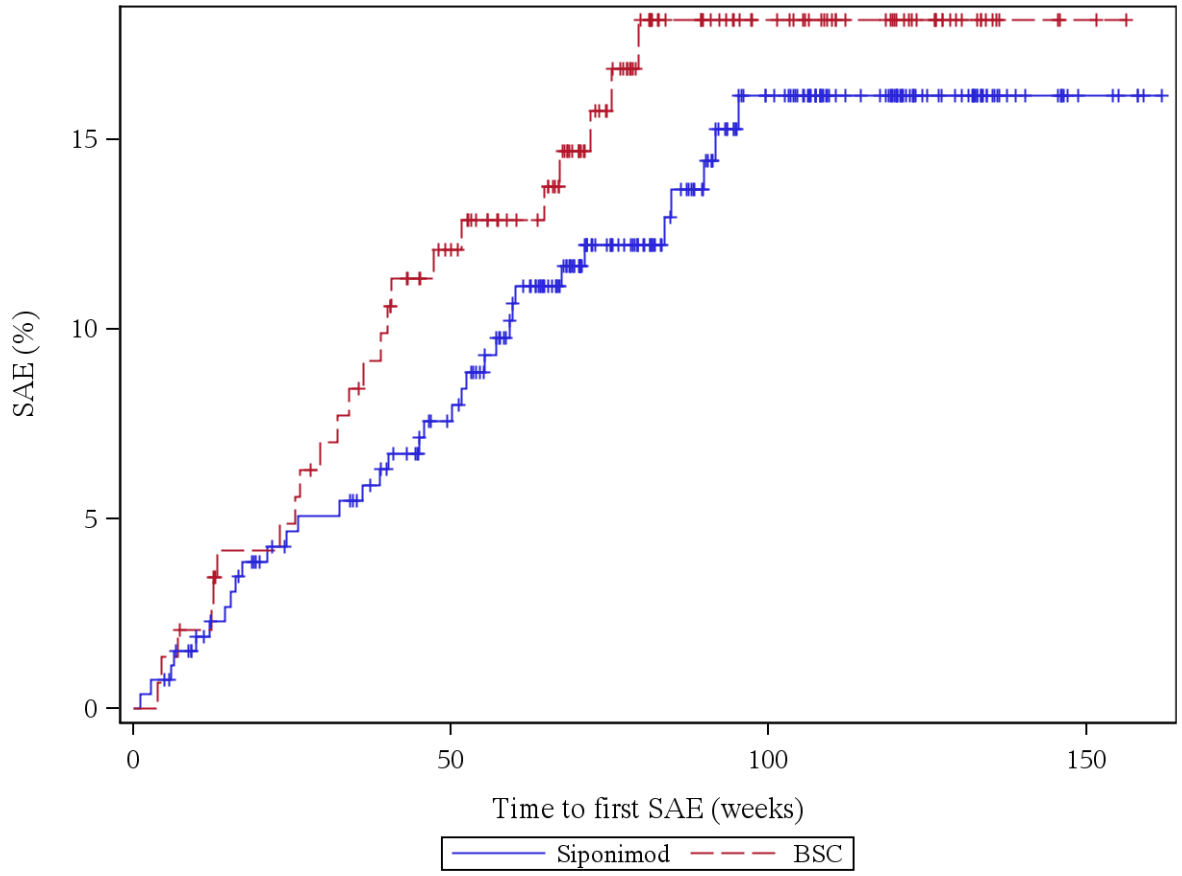
c: Cox proportional hazards model.

Quelle: Anhang 1 zu Modul 4, Tabelle 12.6.1.



Patients at risk (Siponimod/BSC): Week 0: 835/401, Week 4: 811/396, Week 12: 774/381, Week 24: 722/362, Week 36: 682/338, Week 48: 630/308, Week 60: 537/252, Week 72: 416/198, Week 84: 316/157, Week 96: 246/113, Week 108: 170/76, Week 120: 111/48, Week 132: 56/27, Week 144: 27/15, Week 156: 3/4, Week 168: 0/0

Abbildung 50: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Schnell fortschreitende Erkrankung: Nein)



Patients at risk (Siponimod/BSC): Week 0: 264/145, Week 4: 262/144, Week 12: 250/141, Week 24: 237/134, Week 36: 230/127, Week 48: 215/115, Week 60: 193/102, Week 72: 148/79, Week 84: 121/54, Week 96: 91/44, Week 108: 72/36, Week 120: 53/24, Week 132: 32/11, Week 144: 13/4, Week 156: 4/1, Week 168: 0/0

Abbildung 51: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten SUE* aus RCT mit dem zu bewertenden Arzneimittel (Gesamtpopulation – Schnell fortschreitende Erkrankung: Ja)

UE von besonderem Interesse (binäre Analyse)

- Mögliche Effektmodifikation durch das Geschlecht für ‚Periphere Ödeme': Bei Frauen wurden unter Siponimod statistisch signifikant mehr Ereignisse beobachtet. Bei Männern war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-143).

- Mögliche Effektmodifikation durch das Geschlecht für ‚Bradyarrhythmien/ Bradykardien': Bei Männern wurden unter Siponimod statistisch signifikant mehr Ereignisse beobachtet. Bei Frauen war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-143).

- Mögliche Effektmodifikation durch eine DMT-Vorbehandlung für ‚Herpes-simplex-Infektionen': In der Gruppe ohne DMT-Vorbehandlung wurde ein statistisch

signifikanter Effekt zugunsten von Siponimod festgestellt. Bei Patienten mit DMT-Vorbehandlung war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-144).

- Mögliche Effektmodifikation durch eine DMT-Vorbehandlung für ‚Atemwegserkrankungen': Bei Patienten ohne DMT-Vorbehandlung wurde ein statistisch signifikanter Effekt zugunsten von Siponimod festgestellt. Bei Patienten mit DMT-Vorbehandlung war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-144).

- Mögliche Effektmodifikation durch eine IFN-beta-1b-Vorbehandlung für ‚Malignitäten': Bei Patienten mit IFN-beta-1b-Vorbehandlung wurde ein statistisch signifikanter Effekt zugunsten von Siponimod festgestellt. Bei Patienten ohne IFN-beta-1b-Vorbehandlung war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-146).

- Mögliche Effektmodifikation durch eine IFN-beta-1b-Vorbehandlung für ‚Bradyarrhythmien/Bradykardien': Bei Patienten mit IFN-beta-1b-Vorbehandlung wurde ein statistisch signifikanter Effekt zugunsten von BSC festgestellt. Bei Patienten ohne IFN-beta-1b Vorbehandlung war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-146).

- Mögliche Effektmodifikation durch den Erkrankungsverlauf für ‚Herpes-simplex-Infektionen': In keiner der beiden Subgruppen wurde ein statistisch signifikanter Behandlungseffekt festgestellt (Tabelle 4-146).

Da kein konsistenter Effekt beobachtet wurde, wird nicht von einer Modifikation des Behandlungseffekts ausgegangen.

Tabelle 4-143: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Gesamtpopulation)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 1.099)	**BSC (N = 546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
UE von besonderem Interesse (Periphere Ödeme)					
Interaktionstest[b]				0,006	
Männlich					
N'	435	223			
Periphere Ödeme – n (%)	19 (4,4)	11 (4,9)	0,88 [0,41; 1,88] 0,742	0,89 [0,43; 1,83] 0,742	-0,01 [-0,04; 0,03] 0,747
Weiblich					
N'	664	323			
Periphere Ödeme – n (%)	55 (8,3)	7 (2,2)	4,08 [1,84; 9,06] <0,001	3,82 [1,76; 8,30] 0,001	0,06 [0,03; 0,09] < 0,001
UE von besonderem Interesse (Bradyarrhythmien/Bradykardien)					
Interaktionstest[b]				0,031	
Männlich					
N'	435	223			
Bradyarrhythmien/ Bradykardien – n (%)	57 (13,1)	11 (4,9)	2,91 [1,49; 5,66] 0,001	2,66 [1,42; 4,96] 0,002	0,08 [0,04; 0,12] <0,001
Weiblich					
N'	664	323			
Bradyarrhythmien/ Bradykardien – n (%)	46 (6,9)	20 (6,2)	1,13 [0,66; 1,94] 0,664	1,12 [0,67; 1,86] 0,665	0,01 [-0,03; 0,04] 0,658

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.3.5.

Tabelle 4-144: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 1.099)	**BSC (N = 546)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
UE von besonderem Interesse (Herpes-simplex-Infektionen)					
Interaktionstest[b]				0,014	
DMT-Vorbehandlung: Nein					
N'	242	114			
Herpes-simplex-Infektionen – n (%)	2 (0,8)	5 (4,4)	0,18 [0,03; 0,95] 0,024	0,19 [0,04; 0,96] 0,044	-0,04 [-0,07; 0,00] 0,076
DMT-Vorbehandlung: Ja					
N'	857	432			
Herpes-simplex-Infektionen – n (%)	23 (2,7)	6 (1,4)	1,96 [0,79; 4,85] 0,139	1,93 [0,79; 4,71] 0,147	0,01 [-0,00; 0,03] 0,101
UE von besonderem Interesse (Atemwegsinfektionen)					
Interaktionstest[b]				0,019	
DMT-Vorbehandlung: Nein					
N'	242	114			
Atemwegsinfektionen – n (%)	17 (7,0)	19 (16,7)	0,38 [0,19; 0,76] 0,005	0,42 [0,23; 0,78] 0,006	-0,10 [-0,17; -0,02] 0,012
DMT-Vorbehandlung: Ja					
N'	857	432			
Atemwegsinfektionen – n (%)	100 (11,7)	52 (12,0)	0,97 [0,68; 1,38] 0,846	0,97 [0,71; 1,33] 0,846	-0,00 [-0,04; 0,03] 0,847

DMT: Disease Modifying Therapy ; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.4.5.

Tabelle 4-145: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Gesamtpopulation)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 1.099)	BSC (N = 546)	OR [95%-KI][a] p-Wert	RR [95%-KI][a] p-Wert	RD [95%-KI][a] p-Wert
UE von besonderem Interesse (Malignitäten)					
Interaktionstest[b]				0,044	
IFN-beta-1b-Vorbehandlung: Nein					
N'	755	392			
Malignitäten – n (%)	16 (2,1)	7 (1,8)	1,19 [0,49; 2,92] 0,702	1,19 [0,49; 2,86] 0,703	0,00 [-0,01; 0,02] 0,695
IFN-beta-1b-Vorbehandlung: Ja					
N'	344	154			
Malignitäten – n (%)	4 (1,2)	7 (4,5)	0,25 [0,07; 0,86] 0,018	0,26 [0,08; 0,86] 0,028	-0,03 [-0,07; 0,00] 0,057
UE von besonderem Interesse (Bradyarrhythmien/Bradykardien)					
Interaktionstest[b]				0,049	
IFN-beta-1b-Vorbehandlung: Nein					
N'	755	392			
Bradyarrhythmien/ Bradykardien – n (%)	72 (9,5)	28 (7,1)	1,37 [0,87; 2,16] 0,173	1,34 [0,88; 2,03] 0,177	0,02 [-0,01; 0,06] 0,155
IFN-beta-1b-Vorbehandlung: Ja					
N'	344	154			
Bradyarrhythmien/ Bradykardien – n (%)	31 (9,0)	3 (1,9)	4,99 [1,50; 16,57] 0,004	4,63 [1,44; 14,90] 0,010	0,07 [0,03; 0,11] <0,001

IFN: Interferon; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.5.5.

Tabelle 4-146: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Erkrankungsverlauf (Gesamtpopulation)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 1.099)	BSC (N = 546)	OR [95%-KI][a] p-Wert	RR [95%-KI][a] p-Wert	RD [95%-KI][a] p-Wert
UE von besonderem Interesse (Herpes-simplex-Infektionen)					
Interaktionstest[b]				0,046	
Moderater/schwerer Erkrankungsverlauf: Nein					
N'	195	87			
Herpes-simplex-Infektionen – n (%)	4 (2,1)	5 (5,7)	0,34 [0,09; 1,31] 0,103	0,36 [0,10; 1,30] 0,118	-0,04 [-0,09; 0,02] 0,170
Moderater/schwerer Erkrankungsverlauf: Ja					
N'	904	459			
Herpes-simplex-Infektionen – n (%)	21 (2,3)	6 (1,3)	1,80 [0,72; 4,48] 0,203	1,78 [0,72; 4,37] 0,211	0,01 [-0,00; 0,02] 0,164

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 11.7.5.

4.3.1.3.15.8Endpunkt *Läsionslast* – RCT (Gesamtpopulation)

In der Gesamtpopulation ergab der Interaktionstest für den Endpunkt *Läsionslast* Belege für eine mögliche Modifikation des Behandlungseffekts durch das Alter, eine DMT-Vorbehandlung, den Erkrankungsverlauf und die Region im T2-Läsionsvolumen, durch eine DMT-Vorbehandlung, eine IFN-beta-1b-Vorbehandlung und die Region in der Anzahl neuer bzw. sich vergrößernder T2-Läsionen sowie durch die Region in der Anzahl Gd-anreichernder T1-Läsionen.

Veränderung im T2-Läsionsvolumen

- Mögliche Modifikation durch das Alter: Mit Ausnahme der Auswertung zu Monat 12 in der Alterskategorie > 55 Jahre wurde in allen Subgruppen und zu allen Zeitpunkten jeweils ein statistisch signifikanter Behandlungsvorteil von Siponimod festgestellt (Tabelle 4-147).

- Mögliche Modifikation durch eine DMT-Vorbehandlung: Mit Ausnahme der Auswertung zu Monat 12 bei nicht vorbehandelten Patienten wurde in allen Subgruppen

und zu allen Zeitpunkten jeweils ein statistisch signifikanter Behandlungsvorteil von Siponimod festgestellt (Tabelle 4-148).

- Mögliche Modifikation durch den Erkrankungsverlauf: Bei Patienten mit moderatem bzw. schwerem Erkrankungsverlauf bestand jeweils ein statistisch signifikanter Behandlungsvorteil unter Siponimod zu beiden Auswertungszeitpunkten. Bei Patienten ohne moderaten bzw. schweren Erkrankungsverlauf wurde keine statistische Signifikanz erreicht (Tabelle 4-149).

- Mögliche Modifikation durch die Region: Bei Patienten aus europäischen Zentren bestand ein statistisch signifikanter Behandlungsvorteil unter Siponimod zu beiden Auswertungszeitpunkten. Bei Patienten aus nicht-europäischen Zentren wurde nur zu Monat 24 eine statistische Signifikanz erreicht. Zu Monat 12 war der Behandlungseffekt hier nicht statistisch signifikant unterschiedlich (Tabelle 4-150).

Da kein konsistenter und eindeutig interpretierbarer Beleg beobachtet werden konnte, wird im Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* von keiner Modifikation des Behandlungseffekts durch eines der untersuchten Merkmale ausgegangen.

Tabelle 4-147: Ergebnisse für den Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel nach Alter (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung im T2-Läsionsvolumen				
Interaktionstest[b]				< 0,001
Alter 18 – 30 Jahre				
N'	21	10		
Mittelwert zu Studienbeginn (SD)	20.949,52 (18.147,16)	20.924,60 (17.225,68)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-89,36 (388,76)	2.646,82 (553,66)	-2.736,17 [-4.043,82; -1.428,53]	< 0,001
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	174,91 (447,26)	2.930,22 (602,91)	-2.755,31 [-4.214,34; -1.296,28]	< 0,001
Alter 31 – 40 Jahre				
N'	149	83		
Mittelwert zu Studienbeginn (SD)	17.911,99 (19.081,43)	16.064,82 (15.955,19)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	301,72 (147,74)	1.335,29 (194,10)	-1.033,57 [-1.501,05; -566,10]	< 0,001

Monat 24: Adjustierte mittlere Veränderung (SE)[a]	233,40 (169,13)	993,73 (213,80)	-760,33 [-1.284,56; -236,10]	0,005
Alter 41 – 55Jahre				
N'	642	297		
Mittelwert zu Studienbeginn (SD)	14.683,47 (15248,04)	13.383,41 (14.396,17)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	233,89 (80,00)	814,31 (107,57)	-580,41 [-818,48; -342,35]	< 0,001
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	161,16 (87,67)	974,99 (121,11)	-813,82 [-1.084,79; -542,85]	< 0,001
Alter > 55 Jahre				
N'	183	105		
Mittelwert zu Studienbeginn (SD)	16.056,64 (17.107,01)	14.797,77 (17.508,46)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	63,09 (135,96)	234,62 (176,40)	-171,53 [-587,58; 244,52]	0,419
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	125,70 (153,76)	638,49 (202,93)	-512,79 [-995,80; -29,78]	0,037

KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation ; SE: Standard Error .

a: Negative Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 9.1.2.

Tabelle 4-148: Ergebnisse für den Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung im T2-Läsionsvolumen				
Interaktionstest[b]				0,009
DMT-Vorbehandlung: Nein				
N'	215	103		
Mittelwert zu Studienbeginn (SD)	15.412,61 (17.623,09)	14.770,50 (15.546,45)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	395,17 (127,15)	672,61 (179,22)	-277,44 [-686,35; 131,46]	0,183
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	269,66 (141,59)	894,32 (202,49)	-624,66 [-1.088,97; -160,35]	0,008

DMT-Vorbehandlung: Ja				
N'	780	392		
Mittelwert zu Studienbeginn (SD)	15.590,09 (15.953,82)	14.157,92 (15.427,53)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	151,16 (74,28)	855,76 (96,41)	-704,60 [-915,41; -493,80]	< 0,001
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	133,24 (81,82)	951,96 (107,70)	-818,73 [-1.059,32; -578,13]	< 0,001

DMT: Disease Modifying Therapy ; KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation ; SE: Standard Error .

a: Negative Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 9.1.4.

Tabelle 4-149: Ergebnisse für den Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel nach Erkrankungsverlauf (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung im T2-Läsionsvolumen				
Interaktionstest[b]				0,035
Moderater oder schwerer Erkrankungsverlauf: Nein				
N'	179	79		
Mittelwert zu Studienbeginn (SD)	15.095,97 (15.674,31)	14.457,58 (13.335,12)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	335,93 (138,78)	508,90 (201,34)	-172,97 [-634,31; 288,36]	0,462
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	131,69 (154,81)	501,52 (221,76)	-369,82 [-884,17; 144,52]	0,159
Moderater oder schwerer Erkrankungsverlauf: Ja				
N'	816	416		
Mittelwert zu Studienbeginn (SD)	15.651,72 (16.465,82)	14.252,68 (15.820,64)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	175,55 (71,74)	876,30 (93,33)	-700,76 [-905,85; -495,66]	< 0,001
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	169,47 (79,13)	1028,42 (104,84)	-858,95 [-1.093,38; -624,53]	< 0,001

KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Negative Werte zeigen eine Verbesserung an.
b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.
Quelle: Anhang 1 zu Modul 4, Tabellen 9.1.7.

Tabelle 4-150: Ergebnisse für den Endpunkt *Läsionslast (Veränderung im T2-Läsionsvolumen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung im T2-Läsionsvolumen				
Interaktionstest[b]				0,009
Europa				
N'	790	393		
Mittelwert zu Studienbeginn (SD)	15.716,80 (16.777,73)	14.134,29 (14.611,25)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	187,13 (62,75)	887,36 (88,02)	-700,23 [-910,42; -490,04]	< 0,001
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	140,52 (69,11)	971,24 (97,66)	-830,71 [-1.064,05; -597,38]	< 0,001
Andere Regionen				
N'	205	102		
Mittelwert zu Studienbeginn (SD)	14.915,65 (14.440,66)	14.867,52 (18.351,24)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	331,01 (122,97)	582,27 (172,41)	-251,26 [-663,64; 161,13]	0,232
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	327,16 (159,34)	902,14 (224,78)	-574,98 [-1.112,63; -37,33]	0,036

KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.
a: Negative Werte zeigen eine Verbesserung an.
b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.
Quelle: Anhang 1 zu Modul 4, Tabellen 9.1.8.

<u>Anzahl neuer bzw. sich vergrößernder T2-Läsionen</u>

- Mögliche Modifikation durch eine DMT-Vorbehandlung: In allen Subgruppen und zu allen Zeitpunkten wurde jeweils ein statistisch signifikanter Behandlungsvorteil unter Siponimod festgestellt (Tabelle 4-151).

- Mögliche Modifikation durch eine IFN-beta-1b-Vorbehandlung: In allen Subgruppen und zu allen Zeitpunkten wurde jeweils ein statistisch signifikanter Behandlungsvorteil unter Siponimod festgestellt (Tabelle 4-152).

- Mögliche Modifikation durch die Region: In allen Subgruppen und zu allen Zeitpunkten wurde jeweils ein statistisch signifikanter Behandlungsvorteil von Siponimod festgestellt (Tabelle 4-153).

Da kein konsistenter und eindeutig interpretierbarer Beleg für eine Effektmodifikation beobachtet werden konnte, wird im Endpunkt *Läsionslast (Anzahl neuer bzw. sich vergrößernder T2-Läsionen)* von keiner Modifikation des Behandlungseffekts durch eines der untersuchten Merkmale ausgegangen.

Tabelle 4-151: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Rate Ratio [95%-KI]**	**p-Wert**
Neue / sich vergrößernde T2-Läsionen				
Interaktionstest[a]				0,030
DMT-Vorbehandlung: Nein				
N'	215	103		
Monat 12: Anzahl Läsionen [95%-KI]	0,81 [0,60; 1,08]	4,89 [3,39; 7,04]	0,16 [0,11; 0,26]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,40 [0,24; 0,65]	4,48 [3,08; 6,54]	0,09 [0,05; 0,16]	< 0,001
DMT-Vorbehandlung: Ja				
N'	782	393		
Monat 12: Anzahl Läsionen [95%-KI]	1,05 [0,88; 1,26]	3,50 [2,89; 4,24]	0,30 [0,24; 0,38]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,51 [0,37; 0,69]	3,17 [2,51; 3,99]	0,16 [0,11; 0,23]	< 0,001

DMT: Disease Modifying Therapy; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.

a: Aus dem negativen Binomialmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 9.2.4.

Tabelle 4-152: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Rate Ratio [95%-KI]**	**p-Wert**
Neue / sich vergrößernde T2-Läsionen				
Interaktionstest[a]				0,010
IFN-beta-1b-Vorbehandlung: Nein				
N'	679	358		
Monat 12: Anzahl Läsionen [95%-KI]	0,90 [0,75;1,07]	3,77 [3,04;4,66]	0,24 [0,19;0,31]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,38 [0,27;0,53]	3,68 [2,90;4,67]	0,10 [0,07;0,15]	< 0,001
IFN-beta-1b-Vorbehandlung: Ja				
N'	318	138		
Monat 12: Anzahl Läsionen [95%-KI]	1,22 [0,95;1,58]	3,71 [2,78;4,95]	0,33 [0,23;0,48]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,72 [0,48;1,09]	2,80 [2,01;3,92]	0,26 [0,15;0,43]	< 0,001

IFN: Interferon; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.

a: Aus dem negativen Binomialmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 1 zu Modul 4, Tabelle 9.2.5.

Tabelle 4-153: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Rate Ratio [95%-KI]**	**p-Wert**
Neue / sich vergrößernde T2-Läsionen				
Interaktionstest[a]				0,045
Europa				
N'	792	394		
Monat 12: Anzahl Läsionen [95%-KI]	1,07 [0,92; 1,24]	4,09 [3,43; 4,89]	0,26 [0,21; 0,33]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,44 [0,34; 0,57]	3,85 [3,09; 4,79]	0,11 [0,08; 0,16]	< 0,001

Andere Regionen				
N'	205	102		
Monat 12: Anzahl Läsionen [95%-KI]	0,96 [0,56; 1,64]	3,40 [2,21; 5,21]	0,28 [0,14; 0,56]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	1,11 [0,59; 2,10]	2,60 [1,68; 4,04]	0,43 [0,20; 0,93]	0,031

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
a: Aus dem negativen Binomialmodell; Beleg einer Interaktion p < 0,05.
Quelle: Anhang 1 zu Modul 4, Tabelle 9.2.8.

Gd-anreichernde T1-Läsionen:

- Mögliche Modifikation durch die Region: Bei Patienten aus europäischen Zentren bestand jeweils ein statistisch signifikanter Behandlungsvorteil unter Siponimod zu beiden Auswertungszeitpunkten. Bei Patienten aus nicht-europäischen Zentren wurde nur zu Monat 12 eine statistische Signifikanz zugunsten von Siponimod erreicht. Zu Monat 24 war der Behandlungseffekt nicht statistisch signifikant unterschiedlich (Tabelle 4-154).

Da kein konsistenter und eindeutig interpretierbarer Beleg beobachtet werden konnte, wird im Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen*) von keiner Modifikation des Behandlungseffekts durch eines der untersuchten Merkmale ausgegangen.

Tabelle 4-154: Ergebnisse für den Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Gesamtpopulation)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 1.099)**	**BSC (N = 546)**	**Rate Ratio [95%-KI]**	**p-Wert**
Gd-anreichernde T1-Läsionen				
Interaktionstest[a]				0,001
Europa				
N'	791	394		
Monat 12: Anzahl Läsionen [95%-KI]	0,07 [0,05; 0,10]	0,64 [0,48; 0,85]	0,11 [0,07; 0,17]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,04 [0,03; 0,07]	0,45 [0,30; 0,68]	0,10 [0,05; 0,18]	< 0,001
Andere Regionen				
N'	205	102		
Monat 12: Anzahl Läsionen [95%-KI]	0,11 [0,05; 0,24]	0,62 [0,30; 1,28]	0,18 [0,06; 0,52]	0,002

Monat 24: Anzahl Läsionen [95%-KI]	0,32 [0,11; 0,93]	0,18 [0,08; 0,40]	1,80 [0,47; 6,89]	0,392

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
a: Aus dem negativen Binomialmodell; Beleg einer Interaktion $p < 0,05$.
Quelle: Anhang 1 zu Modul 4, Tabelle 9.3.8.

4.3.1.3.15.9Endpunkt *Behinderungsprogression (EDSS)* – RCT (Subpopulation B)

In der Subpopulation B ergab der Interaktionstest für den Endpunkt *Behinderungsprogression (EDSS)* in der binären Analyse der nach sechs Monaten bestätigten Behinderungsprogression einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Geschlecht.

Weibliche Patientinnen zeigten unter Siponimod eine statistisch signifikante Reduktion des Progressionsrisikos. Bei Männern war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-155). Im Endpunkt *Behinderungsprogression (EDSS)* wird daher nicht von einer Modifikation des Behandlungseffekts durch das Geschlecht ausgegangen.

Tabelle 4-155: Ergebnisse für den Endpunkt *Behinderungsprogression (EDSS)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N =61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Inzidenz der nach 6 Monaten bestätigten Behinderungsprogression					
Interaktionstest[b]				0,047	
Weiblich					
N'	76	37			
Patienten mit Behinderungs-progression – n (%)	11 (14,5)	12 (32,4)	0,35 [0,14; 0,90] 0,030	0,45 [0,22; 0,91] 0,028	-0,18 [-0,35; -0,01] 0,039
Männlich					
N'	51	24			
Patienten mit Behinderungs-progression – n (%)	13 (25,5)	4 (16,7)	1,71 [0,49; 5,94] 0,398	1,53 [0,56; 4,20] 0,410	0,09 [-0,10; 0,28] 0,366

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.
a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.
b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.
Quelle: Anhang 2 zu Modul 4, Tabelle 5.1.3.

4.3.1.3.15.10 Endpunkt *Behinderungsprogression (MSFC)* – RCT (Subpopulation B)

In der Subpopulation B ergab der Interaktionstest für den Endpunkt *Behinderungsprogression (MSFC)* in der Veränderung im T25W einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch eine unterschiedlich schnell fortschreitende Erkrankung.

Im T25W wurde in beiden Subgruppen zu keinem Zeitpunkt eine statistische Signifikanz erreicht (Tabelle 4-156). Im Endpunkt *Behinderungsprogression (MSFC)* wird daher nicht von einer Modifikation des Behandlungseffekts ausgegangen.

Tabelle 4-156: Ergebnisse für den Endpunkt *Behinderungsprogression (MSFC)* aus RCT mit dem zu bewertenden Arzneimittel bei schnell fortschreitender Erkrankung (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung im T25W				
Interaktionstest[b]				0,024
Schnell fortschreitende Erkrankung: Nein				
N'	89	46		
Mittelwert zu Studienbeginn (SD)	17,79 (14,94)	15,16 (13,00)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	7,94 (2,06)	6,49 (2,92)	1,45 [-5,69; 8,59]	0,685
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	15,31 (3,44)	9,55 (5,08)	5,76 [-6,61; 18,13]	0,353
Schnell fortschreitende Erkrankung: Ja				
N'	32	14		
Mittelwert zu Studienbeginn (SD)	12,72 (17,72)	12,07 (9,01)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	1,15 (3,39)	12,56 (5,74)	-11,41 [-24,73; 1,91]	0,092
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	3,98 (5,72)	25,53 (9,71)	-21,55 [-44,31; 1,21]	0,063

KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error; T25W: Timed 25 Foot Walking Test.

a: Negative Werte zeigen eine Verbesserung an.

b: Aus dem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 6.4.6.

4.3.1.3.15.11 Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* – RCT (Subpopulation B)

In der Subpopulation B ergab der Interaktionstest für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* in der Inzidenz der Verschlechterung des SDMT um mindestens vier Punkte einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch das Alter.

In keiner der Altersgruppen konnte anhand des relativen Risikos ein statistisch signifikanter Behandlungseffekt beobachtet werden (Tabelle 4-157). Im Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* wird daher nicht von einer Modifikation des Behandlungseffekts ausgegangen.

Tabelle 4-157: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (SDMT)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N = 61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
Inzidenz der Verschlechterung im SDMT nach 12 Monaten					
Interaktionstest[b]				0,034	
Alter 18 – 40 Jahre					
N'	18	13			
Patienten mit Verschlechterung um ≥ 4 Punkte – n (%)	4 (22,2)	4 (30,8)	0,64 [0,13; 3,25] 0,593	0,72 [0,22; 2,37] 0,591	-0,09 [-0,40; 0,23] 0,596
Alter 41 – 55 Jahre					
N'	95	33			
Patienten mit Verschlechterung um ≥ 4 Punkte – n (%)	29 (30,5)	16 (48,5)	0,47 [0,21; 1,05] 0,065	0,63 [0,40; 1,00] 0,051	-0,18 [-0,37; 0,01] 0,070
Alter ≥ 55 Jahre					
N'	14	15			
Patienten mit Verschlechterung um ≥ 4 Punkte – n (%)	8 (57,1)	3 (20,0)	5,33 [1,02; 27,76] 0,047	2,86 [0,94; 8,66] 0,064	0,37 [0,04; 0,70] 0,027

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 7.3.2.

4.3.1.3.15.12 Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* – RCT (Subpopulation B)

In der Subpopulation B ergab der Interaktionstest für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* jeweils Belege für eine mögliche Modifikation des Behandlungseffekts in der Veränderung des Delayed Recall Scores durch das Geschlecht sowie in der Veränderung des Total Recall Scores durch das Alter, das Geschlecht bzw. eine IFN-beta-1b-Vorbehandlung.

<u>Delayed Recall Score</u>

- Mögliche Modifikation durch das Geschlecht: Bei männlichen Patienten wurde zu Monat 12 ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten von BSC festgestellt. Bei weiblichen Patienten war zu keinem der Zeitpunkte eine statistische Signifikanz beobachtet worden (Tabelle 4-158).

Im Endpunkt *Schweregrad der kognitiven Beeinträchtigung (BVMT-R Delayed Recall Score)* wird von keiner Modifikation des Behandlungseffekts durch das Geschlecht ausgegangen.

Tabelle 4-158: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung des BVMT-R Delayed Recall Scores				
Interaktionstest[b]				0,036
Weiblich				
N'	63	34		
Mittelwert zu Studienbeginn (SD)	7,84 (3,23)	6,71 (3,22)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,12 (0,41)	0,10 (0,52)	0,02 [-1,14; 1,19]	0,969
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,97 (0,60)	0,71 (0,78)	0,26 [-1,59; 2,11]	0,781

Männlich				
N'	50	22		
Mittelwert zu Studienbeginn (SD)	7,98 (3,39)	7,95 (3,29)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-1,13 (0,40)	1,16 (0,64)	-2,29 [-3,71; -0,87]	0,002
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,04 (0,58)	1,76 (1,14)	-1,81 [-4,30; 0,69]	0,153

KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 8.1.3.

Total Recall Score

- Mögliche Modifikation durch das Alter: Nur in der Altersgruppe 18 bis 40 Jahre wurde zu Monat 24 ein statistisch signifikanter Effekt zugunsten von BSC festgestellt. Zu Monat 12 war dies in dieser Altersgruppe nicht der Fall. Auch in allen anderen Subgruppen wurde kein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen festgestellt (gilt jeweils für beide Auswertungszeitpunkte) (Tabelle 4-159).

- Mögliche Modifikation durch das Geschlecht: Bei männlichen Patienten wurde zu Monat 12 ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten von BSC festgestellt. Zu Monat 24 war dies nicht der Fall. Bei weiblichen Patienten war zu keinem der Zeitpunkte eine statistische Signifikanz beobachtet worden (Tabelle 4-160).

- Mögliche Modifikation durch eine IFN-beta-1b-Vorbehandlung: Bei Patienten mit IFN-beta-1b-Vorbehandlung wurde zu Monat 24 ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten von Siponimod festgestellt. Zu Monat 12 war dies nicht der Fall. Bei Patienten ohne IFN-beta-1b-Vorbehandlung war der Behandlungseffekt zu keinem Zeitpunkt statistisch signifikant unterschiedlich (Tabelle 4-161).

Da keine konsistenten Effekte beobachtet wurden, wird im Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R Total Recall Score)* von keiner Modifikation des Behandlungseffekts durch Alter, Geschlecht oder IFN-beta-1b-Vorbehandlung ausgegangen.

Tabelle 4-159: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel nach Alter (Subpopulation B)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung des BVMT-R Total Recall Scores				
Interaktionstest[b]				0,016
Alter 18 – 40 Jahre				
N'	16	12		
Mittelwert zu Studienbeginn (SD)	21,44 (9,67)	19,92 (7,46)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-2,92 (1,58)	1,11 (1,95)	-4,03 [-8,96; 0,90]	0,108
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,72 (2,45)	9,02 (3,50)	-9,74 [-18,21; -1,26]	0,025
Alter 41 – 55 Jahre				
N'	86	31		
Mittelwert zu Studienbeginn (SD)	20,37 (8,77)	18 (8,27)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,53 (0,88)	1,12 (1,24)	-1,64 [-4,26; 0,98]	0,217
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,87 (1,14)	0,08 (1,87)	0,79 [-3,32; 4,90]	0,703
Alter > 55 Jahre				
N'	11	13		
Mittelwert zu Studienbeginn (SD)	18,36 (9,39)	17,00 (7,57)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,96 (2,22)	-3,78 (1,72)	4,75 [-0,63; 10,12]	0,083
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	4,06 (3,11)	-0,79 (2,95)	4,85 [-3,60; 13,30]	0,257

KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 8.2.2.

Tabelle 4-160: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**MW-Differenz [95%-KI][a]**	**p-Wert**
Veränderung des BVMT-R Total Recall Scores				
Interaktionstest[b]				0,014
Weiblich				
N'	63	34		
Mittelwert zu Studienbeginn (SD)	20,43 (8,97)	16,85 (7,48)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,68 (0,95)	-1,43 (1,18)	2,11 [-0,50; 4,72]	0,113
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	1,91 (1,35)	-0,29 (1,76)	2,20 [-1,99; 6,38]	0,299
Männlich				
N'	50	22		
Mittelwert zu Studienbeginn (SD)	20,20 (8,92)	20,23 (8,19)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-2,09 (0,92)	2,21 (1,46)	-4,30 [-7,50; -1,11]	0,009
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,00 (1,32)	4,54 (2,65)	-4,54 [-10,31; 1,22]	0,121

KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 8.2.3.

Tabelle 4-161: Ergebnisse für den Endpunkt *Schweregrad der kognitiven Dysfunktion (BVMT-R)* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Subpopulation B)

	Behandlungsgruppen		Vergleich Siponimod vs. BSC	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**MW-Differenz [95%-KI][a]**	**p-Wert**
Veränderung des BVMT-R Total Recall Scores				
Interaktionstest[b]				0,034
IFN-beta-1b-Vorbehandlung: Nein				
N'	72	43		
Mittelwert zu Studienbeginn (SD)	20,76 (8,79)	18,51 (8,46)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,82 (0,85)	0,35 (1,05)	-1,17 [-3,54; 1,20]	0,330
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,49 (1,17)	2,72 (1,61)	-2,23 [-6,01; 1,55]	0,243
IFN-beta-1b-Vorbehandlung: Ja				
N'	41	13		
Mittelwert zu Studienbeginn (SD)	19,56 (9,17)	17,08 (5,62)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,59 (1,10)	-2,46 (2,00)	1,87 [-2,28; 6,02]	0,375
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	1,78 (1,49)	-5,33 (3,02)	7,11 [0,65; 13,58]	0,031

IFN: Interferon; KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 8.2.5.

4.3.1.3.15.13 Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* – RCT (Subpopulation B)

In der Subpopulation B ergab der Interaktionstest für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch eine IFN-beta-1b-Vorbehandlung.

Bei Patienten mit IFN-beta-1b-Vorbehandlung wurde zu Monat 24 ein statistisch signifikanter Unterschied zwischen den Behandlungsgruppen zugunsten von BSC festgestellt. Zu Monat 12 war dies nicht der Fall. Bei Patienten ohne IFN-beta-1b-Vorbehandlung war der Behandlungseffekt zu keinem Zeitpunkt statistisch signifikant unterschiedlich (Tabelle 4-162).

Im Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* wird daher nicht von einer Modifikation des Behandlungseffekts durch eine IFN-beta-1b-Vorbehandlung ausgegangen.

Tabelle 4-162: Ergebnisse für den Endpunkt *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**MW-Differenz [95%-KI][a]**	**p-Wert**
Veränderung der LCVA				
Interaktionstest[b]				0,038
IFN-beta-1b-Vorbehandlung: Nein				
N'	70	42		
Mittelwert zu Studienbeginn (SD)	0,42 (0,27)	0,39 (0,23)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	0,02 (0,03)	-0,01 (0,04)	0,03 [-0,06; 0,11]	0,495
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,00 (0,04)	0,00 (0,05)	0,00 [-0,12; 0,12]	0,978
IFN-beta-1b-Vorbehandlung: Ja				
N'	40	13		
Mittelwert zu Studienbeginn (SD)	0,35 (0,25)	0,40 (0,36)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,02 (0,04)	-0,03 (0,07)	0,01 [-0,14; 0,16]	0,886
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,01 (0,05)	0,23 (0,10)	-0,22 [-0,44; 0,00]	0,048

IFN: Interferon; KI: Konfidenzintervall; LCVA: Low Contrast Visual Acuity; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; SD: Standard Deviation; SE: Standard Error.

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 8.3.5.

4.3.1.3.15.14 Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* – RCT (Subpopulation B)

In der Subpopulation B ergab der Interaktionstest für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* einen Beleg für eine mögliche Modifikation des Behandlungseffekts durch eine unterschiedlich schnell fortschreitende Erkrankung.

In der Auswertung des EQ-5D-Index wurde in der Subgruppe der Patienten mit schnell fortschreitender SPMS zu Monat 24 ein statistisch signifikanter Behandlungseffekt zugunsten von Siponimod erreicht. Zu Monat 12 war der Unterschied nicht statistisch signifikant. Bei Patienten ohne schnell fortschreitende Erkrankung wurde zu keinem Zeitpunkt ein statistisch signifikanter Behandlungsunterschied festgestellt (Tabelle 4-139).

Im Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* wird daher nicht von einer Modifikation des Behandlungseffekts durch den Erkrankungsverlauf ausgegangen.

Tabelle 4-163: Ergebnisse für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D)* aus RCT mit dem zu bewertenden Arzneimittel nach Erkrankungsverlauf (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**MW-Differenz [95%-KI]**	**p-Wert**
Veränderung im EQ-5D-Index				
Interaktionstest[b]				0,037
Schnell fortschreitende Erkrankung: Nein				
N'	88	44		
Mittelwert zu Studienbeginn (SD)	0,60 (0,24)	0,61 (0,24)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-0,03 (0,02)	-0,01 (0,03)		
Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,01 (0,02)	-0,04 (0,03)	0,03 [-0,04; 0,11]	0,381
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-0,05 (0,03)	-0,03 (0,04)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	-0,06 (0,04)	-0,04 (0,06)	-0,02 [-0,16; 0,11]	0,744
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	-0,07 (0,04)	0,01 (0,08)		
Schnell fortschreitende Erkrankung: Ja				
N'	30	14		
Mittelwert zu Studienbeginn (SD)	0,68 (0,22)	0,52 (0,35)		
Monat 6: Adjustierte mittlere Veränderung (SE)[a]	-0,01 (0,04)	-0,08 (0,06)		

Monat 12: Adjustierte mittlere Veränderung (SE)[a]	-0,02 (0,04)	-0,11 (0,06)	0,09 [-0,05; 0,24]	0,203
Monat 18: Adjustierte mittlere Veränderung (SE)[a]	-0,01 (0,05)	-0,16 (0,07)		
Monat 24: Adjustierte mittlere Veränderung (SE)[a]	0,00 (0,06)	-0,28 (0,10)	0,27 [0,04; 0,51]	0,024
Monat 30: Adjustierte mittlere Veränderung (SE)[a]	n. b.	n. b.		

EQ-5D: EuroQol – 5 Dimensionen; KI: Konfidenzintervall; MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; n. b.: nicht bestimmbar; SD: Standard Deviation; SE: Standard Error.

a: Positive Werte zeigen eine Verbesserung an.

b: Aus einem MMRM-Modell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 10.3.6.

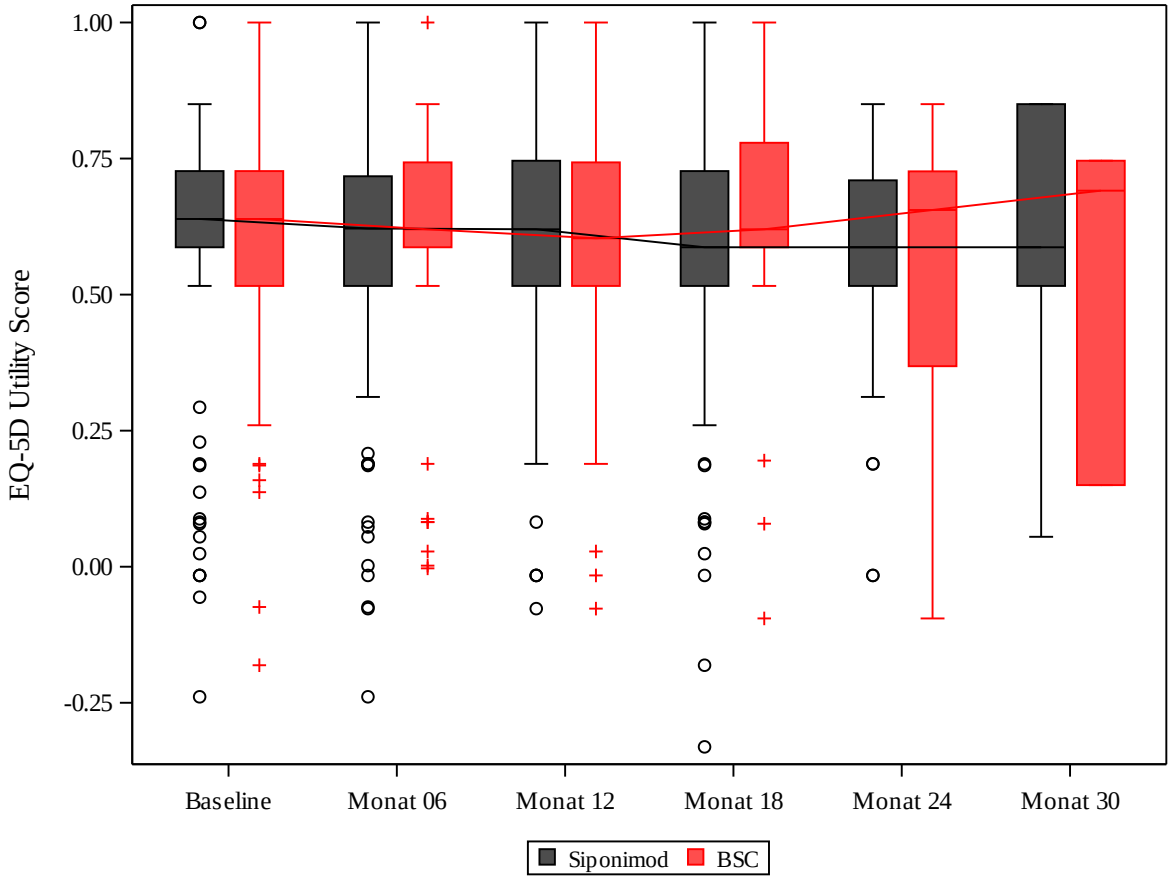


Abbildung 52: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-Index)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Schnell fortschreitende Erkrankung: Nein)

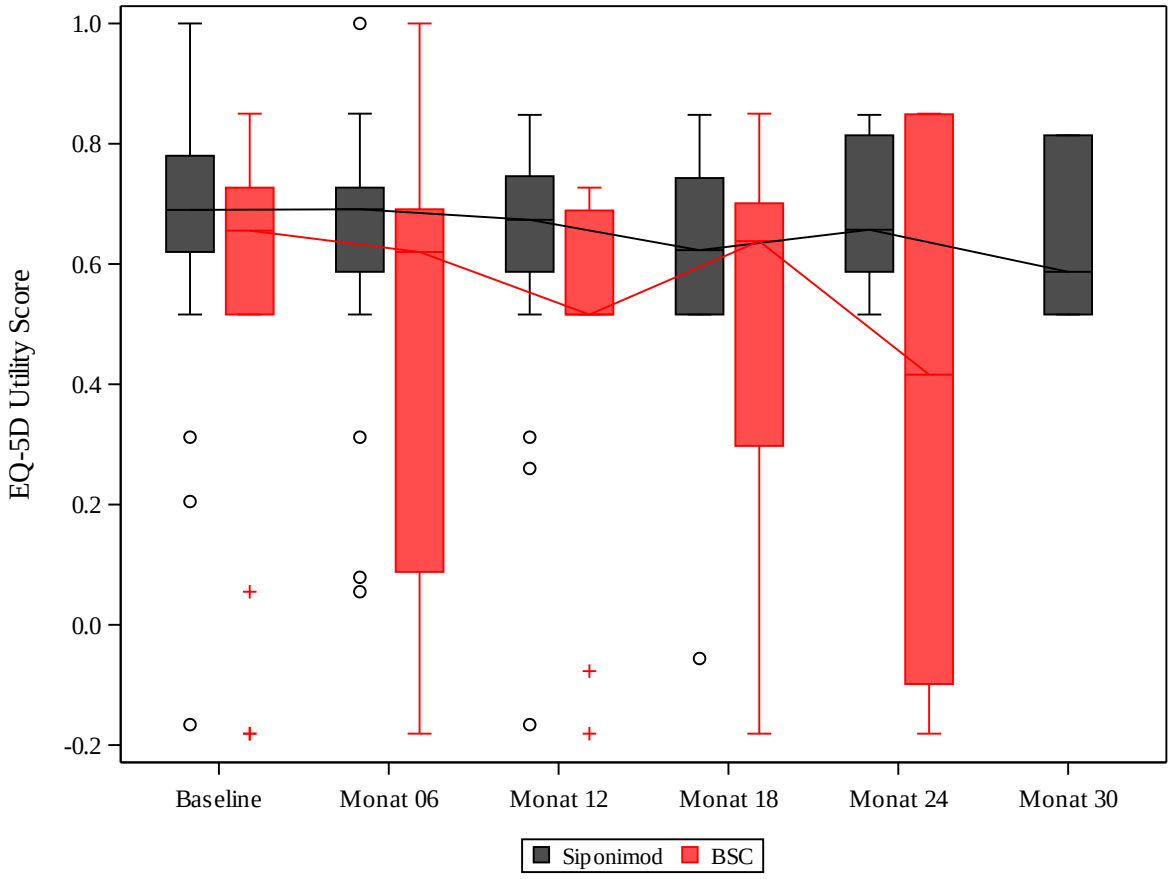


Abbildung 53: Zeitverlauf für den Endpunkt *Gesundheitsbezogene Lebensqualität (EQ-5D-Index)* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Schnell fortschreitende Erkrankung: Ja)

4.3.1.3.15.15 Endpunkt *Unerwünschte Ereignisse* – RCT (Subpopulation B)

In der Subpopulation B ergaben die Interaktionstests für den Endpunkt *Unerwünschte Ereignisse* Belege für eine mögliche Modifikation des Behandlungseffekts durch das Geschlecht in der Gesamtrate der UE, in der Zeit bis zum ersten UE und in der Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse. Außerdem ergaben sich Belege für eine mögliche Effektmodifikation durch das Geschlecht und eine DMT-Vorbehandlung in den UE von besonderem Interesse.

<u>UE (binäre Analyse)</u>

- Mögliche Effektmodifikation durch das Geschlecht: Weibliche Patientinnen zeigten unter Siponimod statistisch signifikant mehr Ereignisse. Bei männlichen Patienten war

kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-164). Von einer Modifikation des Behandlungseffekts wird nicht ausgegangen.

Tabelle 4-164: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B)

EXPAND	**Behandlungsgruppen**		**Vergleich Siponimod vs. BSC**		
	Siponimod (N = 127)	**BSC (N =61)**	**OR [95%-KI][a] p-Wert**	**RR [95%-KI][a] p-Wert**	**RD [95%-KI][a] p-Wert**
UE					
Interaktionstest[b]				0,028	
Männlich					
N'	51	24			
UE – n (%)	43 (84,3)	21 (87,5)	0,77 [0,18; 3,20] 0,716	0,96 [0,80; 1,17] 0,705	-0,03 [-0,20; 0,13] 0,706
Weiblich					
N'	76	37			
UE – n (%)	71 (93,4)	26 (70,3)	6,01 [1,90; 18,95] 0,001	1,33 [1,07; 1,65] 0,010	0,23 [0,07; 0,39] 0,004

KI: Konfidenzintervall; OR: Odds Ratio; n: Anzahl Patienten mit Ereignis; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschte Ereignisse.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 11.3.1.

<u>UE (Zeit bis zum Ereignis)</u>

- Mögliche Effektmodifikation durch das Geschlecht: Weibliche Patientinnen zeigten unter Siponimod ein statistisch signifikant höheres Risiko für UE. Bei männlichen Patienten war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-165). Von einer Modifikation des Behandlungseffekts wird nicht ausgegangen.

Tabelle 4-165: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**HR [95%-KI][c]**	**p-Wert**
UE				
Interaktionstest[a]				0,028
Männlich				
N'	51	24		
Anzahl der Ereignisse	43	21		
Median (in Wochen) [95%-KI][b]	4,1 [2,0; 10,7]	4,4 [1,1; 18,7]		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	14,94 [4,88; 25,00]	5,71 [0,00; 16,38]		
			0,91 [0,54; 1,54]	0,737
Weiblich				
N'	76	37		
Anzahl der Ereignisse	71	26		
Median (in Wochen) [95%-KI][b]	3,6 [2,0; 10,0]	12,9 [5,9; 35,1]		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	4,06 [0,00; 9,10]	29,73 [15,00; 44,46]		
			2,00 [1,27; 3,14]	0,003

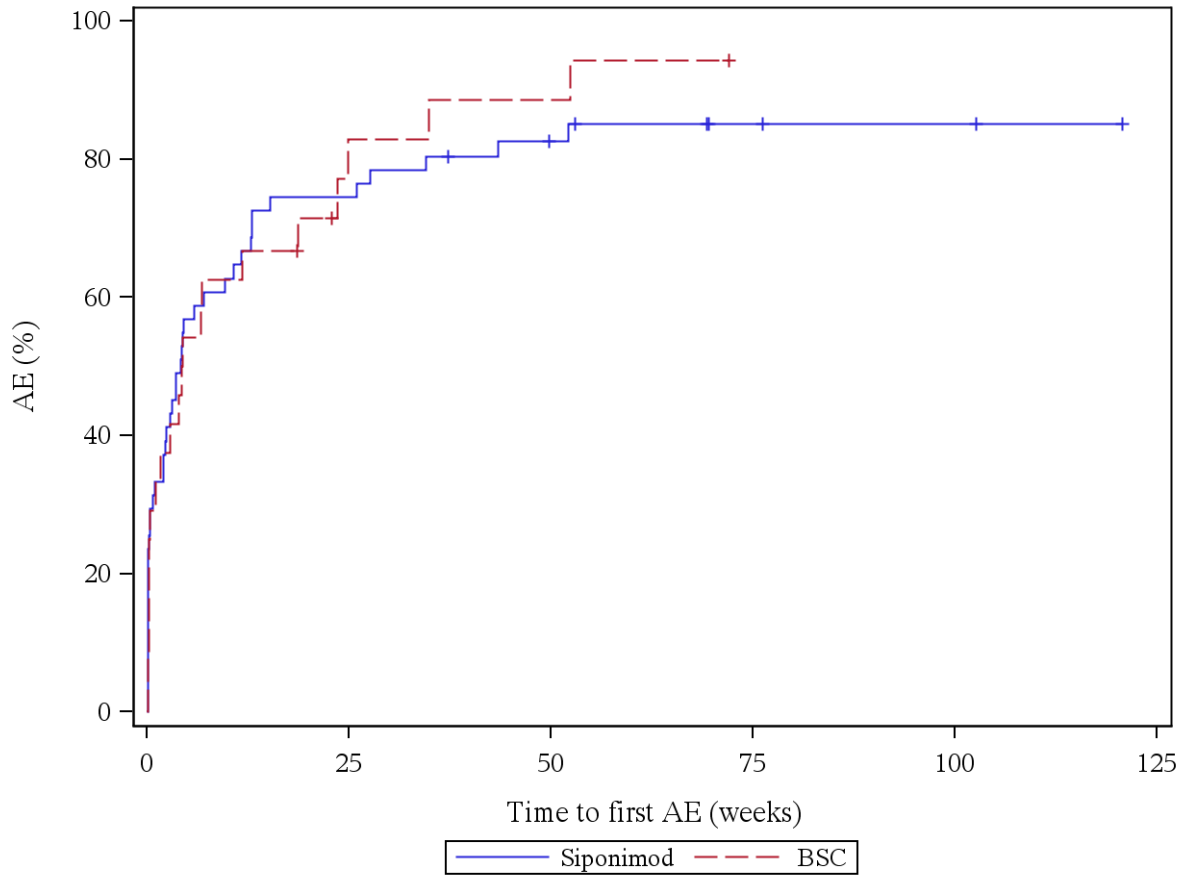
HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; UE: Unerwünschtes Ereignis.

a: Aus einem Cox proportional hazards model; Beleg einer Interaktion $p < 0,05$.

b: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

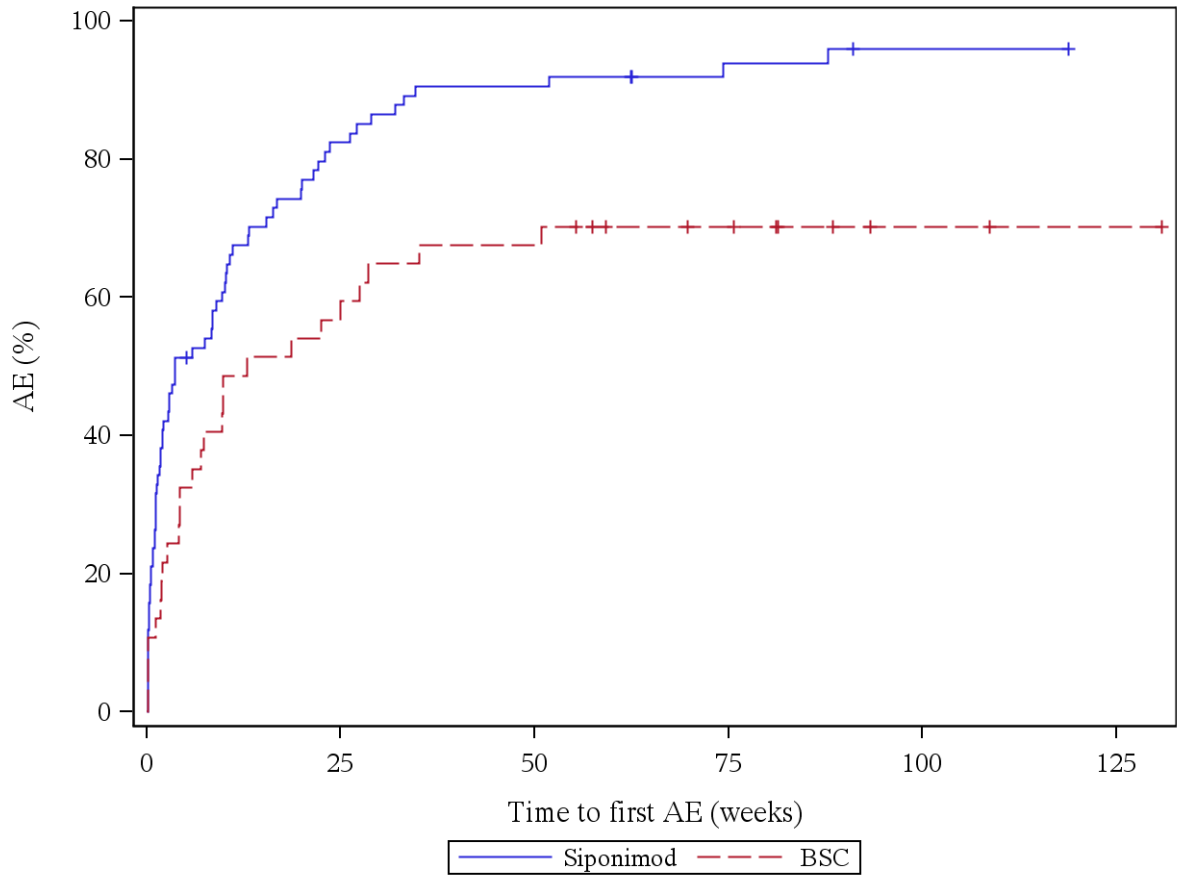
c: Cox proportional hazards model.

Quelle: Anhang 2 zu Modul 4, Tabelle 12.3.1.



Patients at risk (Siponimod/BSC): Week 0: 51/24, Week 4: 26/13, Week 12: 17/8, Week 24: 13/4, Week 36: 10/2, Week 48: 8/2, Week 60: 5/1, Week 72: 3/0, Week 84: 2/0, Week 96: 2/0, Week 108: 1/0, Week 120: 1/0, Week 132: 0/0

Abbildung 54: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Männliche Patienten)



Patients at risk (Siponimod/BSC): Week 0: 76/37, Week 4: 37/28, Week 12: 24/19, Week 24: 13/16, Week 36: 7/12, Week 48: 7/12, Week 60: 6/8, Week 72: 4/7, Week 84: 3/4, Week 96: 1/2, Week 108: 1/2, Week 120: 0/1

Abbildung 55: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Weibliche Patientinnen)

UE ohne krankheitsspezifische UE (Zeit bis zum Ereignis)

- Weibliche Patientinnen zeigten unter Siponimod ein statistisch signifikant höheres Risiko für UE. Bei männlichen Patienten war kein statistisch signifikanter Behandlungseffekt zu beobachten (Tabelle 4-166). Von einer Modifikation des Behandlungseffekts wird nicht ausgegangen.

Tabelle 4-166: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**HR [95%-KI][c]**	**p-Wert**
UE ohne krankheitsspezifische Ereignisse				
Interaktionstest[a]				0,033
Männlich				
N'	51	24		
Anzahl der Ereignisse	41	20		
Median (in Wochen) [95%-KI][b]	12,9 [4,1; 15,3]	6,8 [2,4; 23,6]		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	16,30 [5,44; 27,16]	0,00 [0,00; 0,00]		
			0,81 [0,48; 1,39]	0,451
Weiblich				
N'	76	37		
Anzahl der Ereignisse	67	26		
Median (in Wochen) [95%-KI][b]	10,3 [5,9; 16,7]	22,4 [9,9; 50,9]		
Anteil ereignisfreier Patienten in % – KM-Schätzer [95%-KI][b]	7,15 [0,32; 13,98]	24,32 [6,51; 42,14]		
			1,75 [1,11; 2,77]	0,016

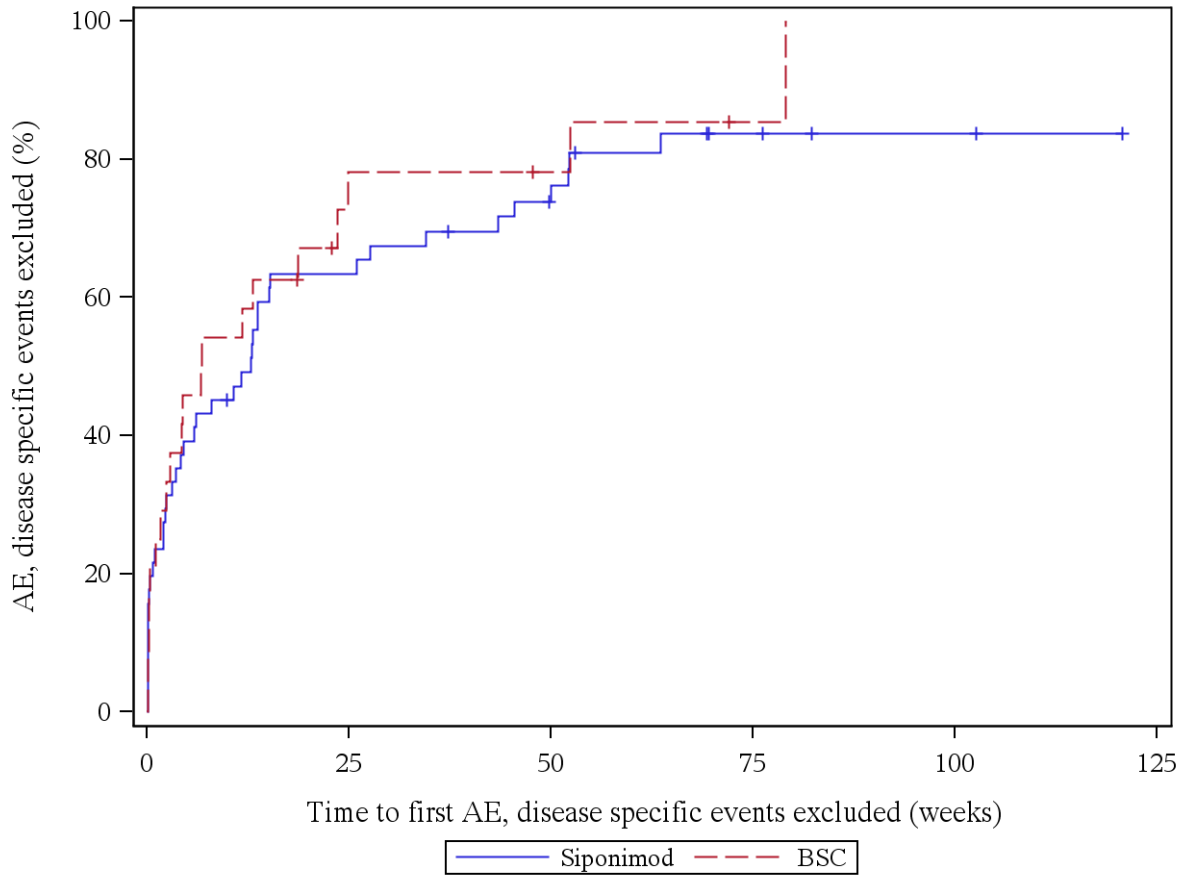
HR: Hazard Ratio; KI: Konfidenzintervall; KM-Schätzer: Kaplan-Meier-Schätzer; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten; UE: Unerwünschtes Ereignis.

a: Aus einem Cox proportional hazards model; Beleg einer Interaktion $p < 0,05$.

b: Greenwood-Formel zur Bestimmung der Konfidenzintervalle.

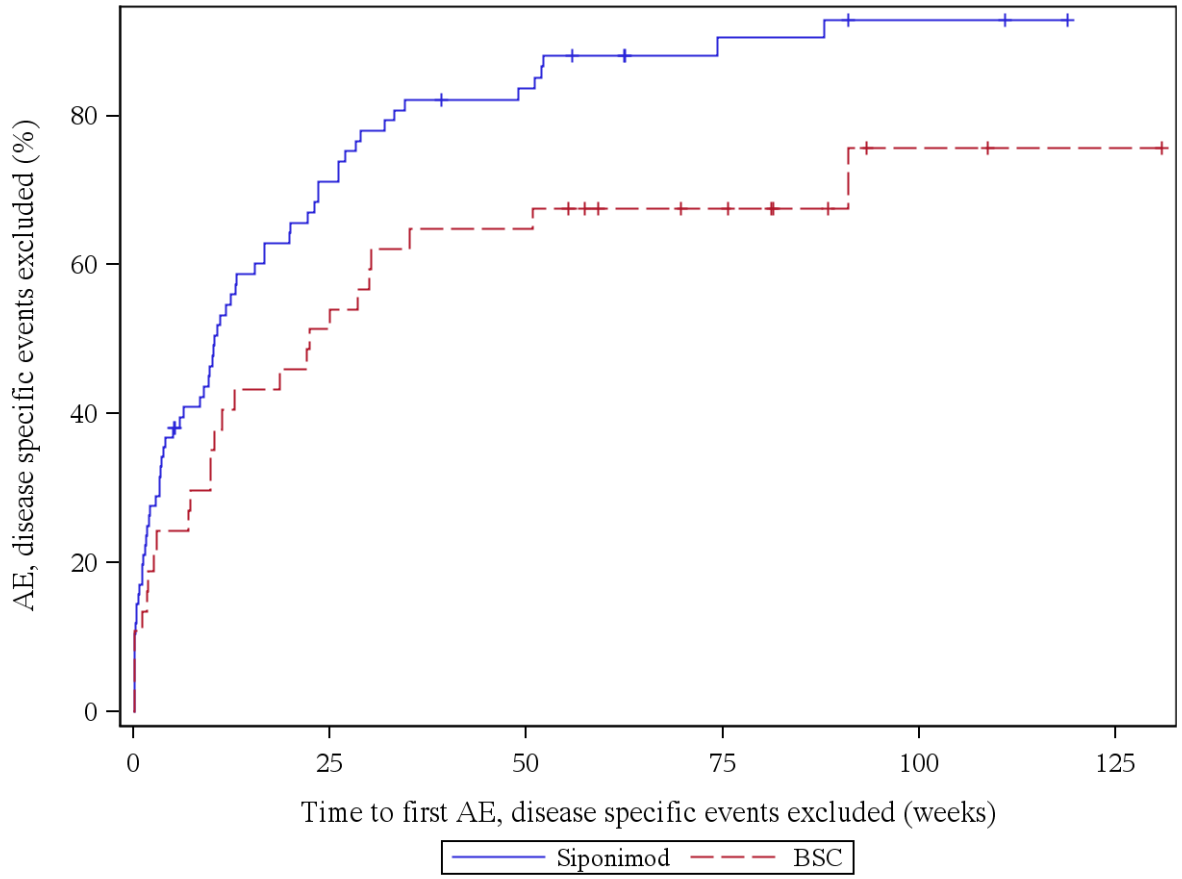
c: Cox proportional hazards model.

Quelle: Anhang 2 zu Modul 4, Tabelle 12.3.1.



Patients at risk (Siponimod/BSC): Week 0: 51/24, Week 4: 33/15, Week 12: 25/10, Week 24: 18/5, Week 36: 15/4, Week 48: 12/3, Week 60: 7/2, Week 72: 4/1, Week 84: 2/0, Week 96: 2/0, Week 108: 1/0, Week 120: 1/0, Week 132: 0/0

Abbildung 56: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Männliche Patienten)



Patients at risk (Siponimod/BSC): Week 0: 76/37, Week 4: 48/28, Week 12: 33/22, Week 24: 21/18, Week 36: 13/13, Week 48: 12/13, Week 60: 7/9, Week 72: 5/8, Week 84: 4/5, Week 96: 2/2, Week 108: 2/2, Week 120: 0/1

Abbildung 57: Kaplan-Meier-Plot für den Endpunkt *Zeit bis zum ersten UE ohne krankheitsspezifische Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel (Subpopulation B – Weibliche Patientinnen)

<u>UE von besonderem Interesse (binäre Analyse)</u>

- Mögliche Effektmodifikation durch das Geschlecht für ‚Hypertonie': In keiner der beiden Subgruppen wurde ein statistisch signifikanter Behandlungseffekt beobachtet (Tabelle 4-167).

- Mögliche Effektmodifikation durch eine DMT-Vorbehandlung für ‚Atemwegserkrankungen': In keiner der beiden Subgruppen wurde ein statistisch signifikanter Behandlungseffekt beobachtet (Tabelle 4-168).

Da kein konsistenter Effekt beobachtet wurde, wird nicht von einer Modifikation des Behandlungseffekts ausgegangen.

Tabelle 4-167: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach Geschlecht (Subpopulation B)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 127)	BSC (N = 61)	OR [95%-KI][a] p-Wert	RR [95%-KI][a] p-Wert	RD [95%-KI][a] p-Wert
UE von besonderem Interesse (Hypertonie)					
Interaktionstest[b]				0,050	
Männlich					
N'	51	24			
Hypertonie – n (%)	5 (9,8)	4 (16,7)	0,54 [0,13; 2,24] 0,394	0,59 [0,17; 2,00] 0,395	-0,07 [-0,24; 0,10] 0,429
Weiblich					
N'	76	37			
Hypertonie – n (%)	12 (15,8)	1 (2,7)	6,75 [0,84; 54,05] 0,041	5,84 [0,79; 43,25] 0,084	0,13 [0,03; 0,23] 0,008

KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 11.3.5.

Tabelle 4-168: Ergebnisse für den Endpunkt *Unerwünschte Ereignisse* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Subpopulation B)

EXPAND	Behandlungsgruppen		Vergleich Siponimod vs. BSC		
	Siponimod (N = 127)	BSC (N = 61)	OR [95%-KI][a] p-Wert	RR [95%-KI][a] p-Wert	RD [95%-KI][a] p-Wert
UE von besonderem Interesse (Atemwegserkrankungen)					
Interaktionstest[b]				0,038	
DMT-Vorbehandlung: Nein					
N'	20	17			
Atemwegsinfektionen – n (%)	2 (10,0)	6 (35,3)	0,20 [0,03; 1,19] 0,063	0,28 [0,07; 1,22] 0,091	-0,25 [-0,52; 0,01] 0,059

DMT-Vorbehandlung: Ja					
N'	107	44			
Atemwegsinfektionen – n (%)	20 (18,7)	5 (11,4)	1,79 [0,63; 5,13] 0,271	1,64 [0,66; 4,11] 0,286	0,07 [-0,05; 0,19] 0,229

DMT: Disease Modifying Therapy; KI: Konfidenzintervall; OR: Odds Ratio; N: Patienten gesamt; n: Anzahl Patienten mit Ereignis; N': Anzahl der in die Auswertung eingegangenen Patienten; RD: Risiko-Differenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.

a: Kann nicht angegeben werden, wenn in keiner Gruppe Ereignisse auftraten.

b: Aus einem logistischen Regressionsmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 11.4.5.

4.3.1.3.15.16 Endpunkt *Läsionslast* – RCT (Subpopulation B)

In der Gesamtpopulation ergaben die Interaktionstests für den Endpunkt *Läsionslast* Belege für eine mögliche Modifikation des Behandlungseffekts durch eine DMT-Vorbehandlung und die Region in der Anzahl neuer bzw. sich vergrößernder T2-Läsionen sowie durch eine DMT-Vorbehandlung, eine IFN-beta-1b-Vorbehandlung und die Region in den Gd-anreichernden T1-Läsionen.

<u>Anzahl neuer bzw. sich vergrößernder T2-Läsionen</u>

- Mögliche Modifikation durch eine DMT-Vorbehandlung: In der Subgruppe der nicht mit DMT vorbehandelten Patienten wurde zu beiden Zeitpunkten ein statistisch signifikanter Behandlungsvorteil von Siponimod festgestellt. In der Subgruppe der mit DMT vorbehandelten Patienten war nur die Auswertung zu Monat 24 statistisch signifikant zugunsten von Siponimod, nicht aber jene zu Monat 12 (Tabelle 4-169).

- Mögliche Modifikation durch die Region: Bei Patienten aus europäischen Zentren wurde zu beiden Zeitpunkten jeweils ein statistisch signifikanter Behandlungsvorteil von Siponimod festgestellt. Bei Patienten aus anderen Regionen war kein statistisch signifikanter Behandlungseffekt festzustellen (Tabelle 4-170).

Da kein konsistenter und eindeutig interpretierbarer Beleg beobachtet werden konnte, wird im Endpunkt *Läsionslast (Anzahl neuer bzw. sich vergrößernder T2-Läsionen)* von keiner Modifikation des Behandlungseffekts durch eines der untersuchten Merkmale ausgegangen.

Tabelle 4-169: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Rate Ratio [95%-KI]**	**p-Wert**
Neue / sich vergrößernde T2-Läsionen				
Interaktionstest[a]				< 0,001
DMT-Vorbehandlung: Nein				
N'	20	16		
Monat 12: Anzahl Läsionen [95%-KI]	1,91 [1,27; 2,87]	16,39 [11,19; 24,03]	0,12 [0,07; 0,20]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,40 [0,12; 1,34]	19,79 [12,53; 31,28]	0,02 [0,01; 0,07]	< 0,001
DMT-Vorbehandlung: Ja				
N'	99	40		
Monat 12: Anzahl Läsionen [95%-KI]	3,57 [2,20; 5,80]	6,96 [4,43; 10,93]	0,51 [0,26; 1,01]	0,052
Monat 24: Anzahl Läsionen [95%-KI]	0,75 [0,30; 1,86]	3,77 [2,16; 6,59]	0,20 [0,07; 0,57]	0,003

DMT: Disease Modifying Therapie; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.

a: Aus dem negativen Binomialmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 9.2.4.

Tabelle 4-170: Ergebnisse für den Endpunkt *Läsionslast (Neue bzw. sich vergrößernde T2-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Rate Ratio [95%-KI]**	**p-Wert**
Neue / sich vergrößernde T2-Läsionen				
Interaktionstest[a]				< 0,001
Europa				
N'	101	48		
Monat 12: Anzahl Läsionen [95%-KI]	2,35 [1,76; 3,14]	9,57 [6,97; 13,14]	0,25 [0,16; 0,38]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,39 [0,25; 0,61]	8,31 [5,15; 13,43]	0,05 [0,02; 0,09]	< 0,001

Andere Regionen				
N'	18	8		
Monat 12: Anzahl Läsionen [95%-KI]	11,06 [4,00; 30,58]	7,23 [2,80; 18,64]	1,53 [0,38; 6,08]	0,546
Monat 24: Anzahl Läsionen [95%-KI]	4,64 [1,28; 16,76]	5,86 [2,36; 14,57]	0,79 [0,16; 3,83]	0,771

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
a: Aus dem negativen Binomialmodell; Beleg einer Interaktion $p < 0,05$.
Quelle: Anhang 2 zu Modul 4, Tabelle 9.2.8.

Gd-anreichernde T1-Läsionen:

- Mögliche Modifikation durch eine DMT-Vorbehandlung: In der Subgruppe der nicht mit DMT vorbehandelten Patienten wurde zu beiden Zeitpunkten jeweils ein statistisch signifikanter Behandlungsvorteil von Siponimod festgestellt. In der Subgruppe der mit DMT vorbehandelten Patienten war nur die Auswertung zu Monat 12 statistisch signifikant zugunsten von Siponimod, nicht aber jene zu Monat 24 (Tabelle 4-171).

- Mögliche Modifikation durch eine IFN-beta-1b-Vorbehandlung: Bei Patienten ohne IFN-beta-1b-Vorbehandlung wurde zu beiden Zeitpunkten jeweils ein statistisch signifikanter Behandlungsvorteil von Siponimod festgestellt. Bei Patienten mit IFN-beta-1b-Vorbehandlung war kein statistisch signifikanter Behandlungseffekt festzustellen (Tabelle 4-172)

- Mögliche Modifikation durch die Region: Bei Patienten aus europäischen Zentren bestand jeweils ein statistisch signifikanter Behandlungsvorteil unter Siponimod zu beiden Auswertungszeitpunkten. Bei Patienten aus nicht-europäischen Zentren wurde zu keinem Zeitpunkt eine statistische Signifikanz erreicht (Tabelle 4-173).

Da kein konsistenter und eindeutig interpretierbarer Beleg beobachtet werden konnte, wird im Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen*) von keiner Modifikation des Behandlungseffekts durch eines der untersuchten Merkmale ausgegangen.

Tabelle 4-171: Ergebnisse für den Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach DMT-Vorbehandlung (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Rate Ratio [95%-KI]**	**p-Wert**
Gd-anreichernde T1-Läsionen				
Interaktionstest[a]				0,006
DMT-Vorbehandlung: Nein				
N'	20	16		
Monat 12: Anzahl Läsionen [95%-KI]	0,03 [0,01; 0,19]	2,37 [1,07; 5,23]	0,01 [0,00; 0,09]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,25 [0,04; 1,68]	2,44 [1,42; 4,18]	0,10 [0,01; 0,75]	0,025
DMT-Vorbehandlung: Ja				
N'	99	40		
Monat 12: Anzahl Läsionen [95%-KI]	0,31 [0,14; 0,68]	1,50 [0,85; 2,64]	0,21 [0,08; 0,56]	0,002
Monat 24: Anzahl Läsionen [95%-KI]	0,33 [0,08; 1,35]	0,56 [0,32; 0,98]	0,58 [0,13; 2,54]	0,474

DMT: Disease Modifying Therapie; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.

a: Aus dem negativen Binomialmodell; Beleg einer Interaktion $p < 0,05$.

Quelle: Anhang 2 zu Modul 4, Tabelle 9.3.4.

Tabelle 4-172: Ergebnisse für den Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach IFN-beta-1b-Vorbehandlung (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Rate Ratio [95%-KI]**	**p-Wert**
Gd-anreichernde T1-Läsionen				
Interaktionstest[a]				0,027
IFN-beta-1b-Vorbehandlung: Nein				
N'	77	45		
Monat 12: Anzahl Läsionen [95%-KI]	0,24 [0,08; 0,70]	1,88 [1,16; 3,08]	0,13 [0,04; 0,41]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,16 [0,06; 0,41]	1,18 [0,72; 1,94]	0,13 [0,04; 0,38]	< 0,001
IFN-beta-1b-Vorbehandlung: Ja				
N'	42	11		
Monat 12: Anzahl Läsionen [95%-KI]	0,33 [0,11; 1,01]	0,94 [0,29; 3,03]	0,35 [0,07; 1,75]	0,202
Monat 24: Anzahl Läsionen [95%-KI]	0,69 [0,13; 3,70]	0,44 [0,14; 1,39]	1,58 [0,20; 12,73]	0,668

IFN: Interferon; KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.

a: Aus dem negativen Binomialmodell; Beleg einer Interaktion p < 0,05.

Quelle: Anhang 2 zu Modul 4, Tabelle 9.3.5.

Tabelle 4-173: Ergebnisse für den Endpunkt *Läsionslast (Gd-anreichernde T1-Läsionen)* aus RCT mit dem zu bewertenden Arzneimittel nach Region (Subpopulation B)

	Behandlungsgruppen		**Vergleich Siponimod vs. BSC**	
EXPAND	**Siponimod (N = 127)**	**BSC (N = 61)**	**Rate Ratio [95%-KI]**	**p-Wert**
Gd-anreichernde T1-Läsionen				
Interaktionstest[a]				< 0,001
Europa				
N'	101	48		
Monat 12: Anzahl Läsionen [95%-KI]	0,12 [0,07; 0,22]	1,58 [0,98; 2,54]	0,08 [0,04; 0,16]	< 0,001
Monat 24: Anzahl Läsionen [95%-KI]	0,11 [0,04; 0,30]	0,97 [0,56; 1,66]	0,11 [0,03; 0,35]	< 0,001

Andere Regionen				
N'	18	8		
Monat 12: Anzahl Läsionen [95%-KI]	1,57 [0,47; 5,20]	1,91 [0,51; 7,18]	0,82 [0,14; 4,90]	0,830
Monat 24: Anzahl Läsionen [95%-KI]	2,50 [0,52; 11,99]	0,83 [0,35; 1,99]	3,01 [0,50; 18,24]	0,230

KI: Konfidenzintervall; N: Patienten gesamt; N': Anzahl der in die Auswertung eingegangenen Patienten.
a: Aus dem negativen Binomialmodell; Beleg einer Interaktion $p < 0,05$.
Quelle: Anhang 2 zu Modul 4, Tabelle 9.3.8.

4.3.1.4 Liste der eingeschlossenen Studien - RCT

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).

EXPAND-Studie:

- Studienbericht (9),
- Zusatzauswertungen Anhang 1 und 2 (51, 52),
- Registereinträge (41-46),
- Publikation (47, 48).

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 sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).*** *Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie*

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der Internetseite des G-BA*
- *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 identifiziert wurden und bewerten Sie darüber hinaus deren Ähnlichkeit. Begründen Sie darauf basierend den Ein- bzw. Ausschluss von Studien für die von Ihnen durchgeführten indirekten Vergleiche. Bewerten Sie das Verzerrungspotenzial der für indirekte Vergleiche herangezogenen Studien. ***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 4-174: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits-bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend.					

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 drei 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 0.**

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

Tabelle 4-175: 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>
	Nicht zutreffend.				

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 4-176: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
	Nicht zutreffend.

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 4-177: 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
Nicht zutreffend.						

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 4-178: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 0)
Nicht zutreffend.	

Nicht zutreffend.

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

- *Homogenität der Ergebnisse: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. 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 inkonsistente Ergebnisse.*

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 4.3.1.3.13.***

Nicht zutreffend.

4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

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).

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/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*

- *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 4-179: 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		
Nicht zutreffend.						

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

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

Tabelle 4-180: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

Studie	<Mortalität>	<Gesundheits-bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend.					

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 4-181: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
	Nicht zutreffend.

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 4-182: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
Nicht zutreffend.				

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 4.3.1.3.13.***

Nicht zutreffend.

4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende 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).

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

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

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

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

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

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

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

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

Nicht zutreffend.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

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

Tabelle 4-183: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

Studie	**<Mortalität>**	**<Gesundheits-bezogene Lebensqualität>**	**<Endpunkt>**	**<Endpunkt>**	**<Endpunkt>**
Nicht zutreffend.					

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

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

Tabelle 4-184: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	**Operationalisierung**
	Nicht zutreffend.

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. 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.

Nicht zutreffend.

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

Nicht zutreffend.

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

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. ***Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.13.***

Nicht zutreffend.

4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen

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).

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.

In die Bewertung des Zusatznutzens im direkten Vergleich wurde eine RCT (Evidenzstufe Ib) eingeschlossen.

Die Studie ist für die zu bewertende Indikation und die relevante Patientengruppe aufgrund der Population, Fallzahl, Studiendauer, Art der Intervention und der Wahl der Endpunkte relevant und erlaubt eine valide Beurteilung des Nutzens und Zusatznutzens.

Das Verzerrungspotenzial war studienspezifisch niedrig, die Studienqualität somit hoch. Das Verzerrungspotenzial war auf Endpunktebene ebenfalls niedrig. Ausnahmen hiervon bilden die Auswertungen zu Monat 24, da hier die Nichtberücksichtigungsanteile hoch waren. Die Auswertungen zu Monat 12 bzw. zu Studienende lassen jedoch eine valide Bewertung des Zusatznutzens zu. Die Auswertungen zu Monat 24 werden ergänzend dargestellt.

In der Studie wurden patientenrelevante Endpunkte zur Mortalität, Morbidität, gesundheitsbezogenen Lebensqualität sowie Sicherheit und Verträglichkeit erhoben. Die für die Bewertung herangezogenen Endpunkte sind valide.

Die Aussagekraft der oben genannten RCT ist somit sehr hoch und damit für eine valide Beurteilung des medizinischen Zusatznutzens uneingeschränkt geeignet.

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.

Für Siponimod ergibt sich in der Zusammenschau der Ergebnisse ein **Hinweis auf einen geringen Zusatznutzen** gegenüber der zweckmäßigen Vergleichstherapie BSC in der Gesamtpopulation (ergänzend dargestellt) sowie in der bewertungsrelevanten Subpopulation B.

Tabelle 4-185: Bewertung des Zusatznutzens bezüglich der untersuchten Endpunkte (Gesamtpopulation und Subpopulation B)

Endpunkt	**Zusatznutzen Siponimod vs. BSC**	
	Gesamtpopulation	**Subpopulation B**
Mortalität		
• Todesfälle gesamt	=	=
Morbidität		

• Bestätigte Krankheitsschübe	+	+
• Behinderungsprogression (EDSS)	+	=
• Behinderungsprogression (MSFC)	=	=
• Schweregrad der kognitiven Dysfunktion (SDMT)	+	+
• Schweregrad der kognitiven Dysfunktion (BVMT-R)	=	=
• Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)	=	=
• Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)	=	=
Gesundheitsbezogene Lebensqualität		
• MSIS-29	+	=
• EQ-5D-Index	+	=
Sicherheit und Verträglichkeit		
• Unerwünschte Ereignisse (UE / SUE / Abbrüche)	– / = / =	– / = / =
• Unerwünschte Ereignisse von besonderem Interesse	−	=
• Suizidalität	=	=

Ergänzend dargestellt:		
Bildgebung		
• Läsionslast	+	+
• Hirnvolumen	+	+

BSC: Best Supportive Care; BVMT-R: Brief Visuospatial Memory Test – Revised; EDSS: Expanded Disability Status Scale; EQ-5D: EuroQol – 5 Dimensionen; LCVA: Low Contrast Visual Acuity; MSFC: Multiple Sclerosis Functional Composite; MSIS: Multiple Sclerosis Impact Scale; MSWS: Multiple Sclerosis Walking Scale; SDMT: Symbol Digital Modalities Test; SUE: Schwerwiegendes Unerwünschtes Ereignis; UE: Unerwünschtes Ereignis.

\+ Vorteil von Siponimod.

= kein Unterschied zwischen den Behandlungen.

– Nachteil für Siponimod.

Hinsichtlich der Mortalität war kein Unterschied zwischen Siponimod und BSC zu beobachten.

Siponimod zeigte in der Gesamtpopulation sowie in der Subpopulation B eine statistisch signifikante Reduktion der jährlichen Rate *bestätigter Krankheitsschübe* und eine statistisch signifikante Reduktion des Schubrisikos.

Im Endpunkt *Behinderungsprogression (EDSS)* zeigte sich, dass die Inzidenz einer Behinderungsprogression in der Gesamtpopulation signifikant reduziert wurde bei gleichzeitig signifikant reduziertem Risiko einer Behinderungsprogression (Zeit bis zur ersten bestätigten Progression). In der Subpopulation B konnte in diesem Endpunkt kein statistisch signifikanter Unterschied festgestellt werden.

Es konnten sowohl in der Gesamtpopulation als auch in der Subpopulation B jeweils eine statistisch signifikante Reduktion *kognitiver Beeinträchtigungen* (gemessen anhand des SDMT) unter Siponimod beobachtet werden.

In der *Gesundheitsbezogenen Lebensqualität* zeigte sich in der Gesamtpopulation ein statistisch signifikanter Vorteil von Siponimod, gemessen anhand des validierten MSIS-29-Fragenbogens sowie des EQ-5D. In der Subpopulation B konnte in diesem Endpunkt kein statistisch signifikanter Unterschied festgestellt werden.

Bei den sicherheitsrelevanten Endpunkten (*Unerwünschte Ereignisse*) zeigte sich in der Gesamtpopulation und in der Subpopulation B eine statistisch signifikant höhere Inzidenz unter Siponimod, die aber vor allem auf Ereignisse mit geringem Schweregrad zurückzuführen ist. Schwere UE waren nicht statistisch signifikant häufiger. SUE, Behandlungsabbrüche aufgrund von UE und patientenrelevante UE von besonderem Interesse traten in der Subpopulation B unter einer Therapie mit Siponimod nicht statistisch signifikant häufiger auf als unter BSC. In der Gesamtpopulation waren UE von besonderem Interesse unter Siponimod statistisch signifikant häufiger, meist aber von milder bis moderater Ausprägung.

Daneben konnte Siponimod sowohl in der Gesamtpopulation als auch in der Subpopulation B die *Läsionslast* (T2- sowie Gd-anreichernde T1-Läsionen) sowie die Hirnatrophie (*Verringerung des Hirnvolumens)* statistisch signifikant reduzieren. Beides sind zentrale Parameter in der Überwachung des Therapieerfolgs in der klinischen Praxis und werden daher ergänzend dargestellt.

Die Aussagekraft der randomisierten, kontrollierten Studie, die der Bewertung zugrunde liegt, ist sehr hoch; damit ist die Studie für eine valide Beurteilung des medizinischen Nutzens und Zusatznutzens uneingeschränkt geeignet. Ebenso lassen sich die Ergebnisse der Studie uneingeschränkt auf den deutschen Versorgungskontext übertragen. Die Subgruppenanalysen ergaben keinen eindeutigen Beleg für eine mögliche Modifikation des therapeutischen Effekts durch Alter, Geschlecht, DMT-Vorbehandlung, IFN-beta-1b-Vorbehandlung, eine unterschiedlich schnell fortschreitende Erkrankung, die Schwere des Erkrankungsverlaufs oder durch die Region.

Ein geringer Zusatznutzen liegt gemäß Verfahrensordnung §5 Absatz 7 Ziffer 3 vor, wenn eine gegenüber der zweckmäßigen Vergleichstherapie bisher nicht erreichte moderate und nicht nur geringfügige Verbesserung des therapierelevanten Nutzens erreicht wird, insbesondere eine Verringerung von nicht schwerwiegenden Symptomen der Erkrankung. Auf Basis der beobachteten Reduktion von Krankheitsschüben und einer Reduktion kognitiver Beeinträchtigungen im Vergleich zur zweckmäßigen Vergleichstherapie ergibt sich daher ein Hinweis auf einen geringen patientenrelevanten Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie in der Subpopulation B (Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe).

Der Zusatznutzen in der Subpopulation B ist aufgrund der Studie der Evidenzstufe Ib, die der Bewertung zugrunde liegt, als gesichert anzusehen.

Für die Subpopulation A (Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, mit aufgesetzten Schüben) liegen keine Daten gegenüber der zweckmäßigen Vergleichstherapie vor. Ein Zusatznutzen ist für die Subpopulation A daher nicht belegt.

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 4-186: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, mit aufgesetzten Schüben (Subpopulation A)	Nicht belegt
Erwachsene Patienten mit SPMS mit aktiver Erkrankung, definiert durch klinischen Befund oder Bildgebung, ohne aufgesetzte Schübe (Subpopulation B)	Hinweis auf einen geringen Zusatznutzen

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

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

Nicht zutreffend.

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

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

Nicht zutreffend.

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

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

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

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

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005[17], Molenberghs 2010[18]). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrundeliegende 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[19]) 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[20]) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

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

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

[17] Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.
[18] 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.
[19] Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.
[20] Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

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.

Aufgrund der Bedeutung Gd-anreichernder T1-Läsionen und der T2-Läsionslast für die Definition der Aktivität der SPMS-Erkrankung und der Bedeutung von Läsionslast und Hirnatrophie als Surrogate für das Fortschreiten der Erkrankung, werden sowohl die Läsionslast als auch die Veränderung des Hirnvolumens ergänzend im Dossier dargestellt. Eine Surrogatvalidierung ist nicht erfolgt.

4.6 Referenzliste

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

1. Novartis. Fachinformation Mayzent (Stand: 02/2020). 2020.
2. Gemeinsamer Bundesausschuss. Niederschrift (finale Fassung) zum Beratungsgespräch gemäß § 8 AM-NutzenV; Beratungsanforderung 2015-B-050. 2015.
3. Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV; Beratungsanforderung 2017-B-205. 2018.
4. Gemeinsamer Bundesausschuss. Niederschrift zum Beratungsgespräch gemäß § 8 AM-NutzenV; Beratungsanforderung 2019-B-049. 2019.
5. Gemeinsamer Bundesausschuss. Ergänzung zu Beratungsgespräch 2019-B-049 nach Positive Opinion. 2019.
6. Roche. Fachinformation Ocrevus (Stand: 06/2019). 2019.
7. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sorensen PS, Thompson AJ, Wolinsky JS, Balcer LJ, Banwell B, Barkhof F, Bebo B, Jr., Calabresi PA, Clanet M, Comi G, Fox RJ, Freedman MS, Goodman AD, Inglese M, Kappos L, Kieseier BC, Lincoln JA, Lubetzki C, Miller AE, Montalban X, O'Connor PW, Petkau J, Pozzilli C, Rudick RA, Sormani MP, Stuve O, Waubant E, Polman CH. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology. 2014;83(3):278-86.
8. European Medicines Agency. Guideline on clinical investigation of medicinal products for the treatment of Multiple Sclerosis. EMA/CHMP/771815/2011. 2007.
9. Novartis. CBAF312A2304 - A multicenter, randomized, double-blind, parallel-group, placebo-controlled variable treatment duration study evaluating the efficacy and safety of Siponimod (BAF312) in patients with secondary progressive multiple sclerosis followed by extended treatment with open-label BAF312 [Studienbericht]. 2018.
10. Wiendl H, Kieseier BC. Multiple Sklerose - Klinik, Diagnostik und Therapie. Brandt T, Hohlfeld R, Noth J, Reichmann H, editors: Kohlhammer; 2010. 14-21; 76-97; 154-7 p.
11. Amato MP, Grimaud J, Achiti I, Bartolozzi ML, Adeleine P, Hartung HP, Kappos L, Thompson A, Trojano M, Vukusic S, Confavreux C. European validation of a standardized clinical description of multiple sclerosis. J Neurol. 2004;251(12):1472-80.

12. Syndulko K, Ke D, Ellison GW, Baumhefner RW, Myers LW, Tourtellotte WW. Comparative evaluations of neuroperformance and clinical outcome assessments in chronic progressive multiple sclerosis: I. Reliability, validity and sensitivity to disease progression. Multiple Sclerosis Study Group. Multiple sclerosis. 1996;2(3):142-56.
13. Tremlett H, Yinshan Z, Devonshire V. Natural history of secondary-progressive multiple sclerosis. Multiple sclerosis. 2008;14(3):314-24.
14. Gross HJ, Watson C. Characteristics, burden of illness, and physical functioning of patients with relapsing-remitting and secondary progressive multiple sclerosis: a cross-sectional US survey. Neuropsychiatr Dis Treat. 2017;13:1349-57.
15. Fischer JSJ, A.J.; Kniker, J.E.; Rudick, R.A.; Cutter, G. Multiple Sclerosis Functional Composite (MSFC) Administration and scoring manual (revised, October 2001): National Multiple Sclerosis Society; 2001.
16. Rudick RA, Cutter G, Reingold S. The multiple sclerosis functional composite: a new clinical outcome measure for multiple sderosis trials. Multiple sclerosis. 2002;8(5):359-65.
17. Brochet B, Ruet A. Cognitive Impairment in Multiple Sclerosis With Regards to Disease Duration and Clinical Phenotypes. Front Neurol. 2019;10:261.
18. Clemens L, Langdon D. How does cognition relate to employment in multiple sclerosis? A systematic review. Mult Scler Relat Disord. 2018;26:183-91.
19. Benedict RH, DeLuca J, Phillips G, LaRocca N, Hudson LD, Rudick R, Multiple Sclerosis Outcome Assessments C. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. Multiple sclerosis. 2017;23(5):721-33.
20. Costa SL, Genova HM, DeLuca J, Chiaravalloti ND. Information processing speed in multiple sclerosis: Past, present, and future. Multiple sclerosis. 2017;23(6):772-89.
21. Kalb R, Beier M, Benedict RH, Charvet L, Costello K, Feinstein A, Gingold J, Goverover Y, Halper J, Harris C, Kostich L, Krupp L, Lathi E, LaRocca N, Thrower B, DeLuca J. Recommendations for cognitive screening and management in multiple sclerosis care. Multiple sclerosis. 2018;24(13):1665-80.
22. Benedict RHB, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. Psychological Assessment. 1996;8(2):145-53.
23. Chahin S, Balcer LJ, Miller DM, Zhang A, Galetta SL. Vision in a phase 3 trial of natalizumab for multiple sclerosis: relation to disability and quality of life. J Neuroophthalmol. 2015;35(1):6-11.
24. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Measuring the impact of MS on walking ability: the 12-Item MS Walking Scale (MSWS-12). Neurology. 2003;60(1):31-6.
25. Motl RW, Snook EM. Confirmation and extension of the validity of the Multiple Sclerosis Walking Scale-12 (MSWS-12). J Neurol Sci. 2008;268(1-2):69-73.
26. Hobart JC, Cano S, Ingram W, Zajicek J. A new path for the MS Walking Scale: MSWS-12 version 2. . Mult Scler J. 2012;18:334.
27. Nilsagard Y, Lundholm C, Gunnarsson LG, Dcnison E. Clinical relevance using timed walk tests and 'timed up and go' testing in persons with multiple sclerosis. Physiother Res Int. 2007;12(2):105-14.
28. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain : a journal of neurology. 2001;124(Pt 5):962-73.
29. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Technology Assessment. 2009;13(12).

30. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Annals of medicine. 2001;33(5):337-43.
31. Walters SJ, Brazier JE. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2005;14(6):1523-32.
32. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, Currier GW, Melvin GA, Greenhill L, Shen S, Mann JJ. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry. 2011;168(12):1266-77.
33. Cook SD, Dhib-Jalbut S, Dowling P, Durelli L, Ford C, Giovannoni G, Halper J, Harris C, Herbert J, Li D, Lincoln JA, Lisak R, Lublin FD, Lucchinetti CF, Moore W, Naismith RT, Oehninger C, Simon J, Sormani MP. Use of Magnetic Resonance Imaging as Well as Clinical Disease Activity in the Clinical Classification of Multiple Sclerosis and Assessment of Its Course: A Report from an International CMSC Consensus Conference, March 5-7, 2010. International journal of MS care. 2012;14(3):105-14.
34. Filippi M, Paty DW, Kappos L, Barkhof F, Compston DA, Thompson AJ, Zhao GJ, Wiles CM, McDonald WI, Miller DH. Correlations between changes in disability and T2-weighted brain MRI activity in multiple sclerosis: a follow-up study. Neurology. 1995;45(2):255-60.
35. Tekok-Kilic A, Benedict RH, Zivadinov R. Update on the relationships between neuropsychological dysfunction and structural MRI in multiple sclerosis. Expert RevNeurother. 2006;6(3):323-31.
36. Popescu V, Agosta F, Hulst HE, Sluimer IC, Knol DL, Sormani MP, Enzinger C, Ropele S, Alonso J, Sastre-Garriga J, Rovira A, Montalban X, Bodini B, Ciccarelli O, Khaleeli Z, Chard DT, Matthews L, Palace J, Giorgio A, De Stefano N, Eisele P, Gass A, Polman CH, Uitdehaag BM, Messina MJ, Comi G, Filippi M, Barkhof F, Vrenken H, Group MS. Brain atrophy and lesion load predict long term disability in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2013;84(10):1082-91.
37. Calabrese M, Agosta F, Rinaldi F, Mattisi I, Grossi P, Favaretto A, Atzori M, Bernardi V, Barachino L, Rinaldi L, Perini P, Gallo P, Filippi M. Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis. Archives of neurology. 2009;66(9):1144-50.
38. Prinster A, Quarantelli M, Lanzillo R, Orefice G, Vacca G, Carotenuto B, Alfano B, Brunetti A, Morra VB, Salvatore M. A voxel-based morphometry study of disease severity correlates in relapsing-- remitting multiple sclerosis. Multiple sclerosis. 2010;16(1):45-54.
39. Lukas C, Minneboo A, de Groot V, Moraal B, Knol DL, Polman CH, Barkhof F, Vrenken H. Early central atrophy rate predicts 5 year clinical outcome in multiple sclerosis. J Neurol Neurosurg Psychiatry. 2010;81(12):1351-6.
40. United Nations. Geographic regions. 2019 [Zuletzt aktualisiert 2019; abgerufen am 23.07.2019]; Abrufbar unter: https://unstats.un.org/unsd/methodology/m49/.
41. Novartis. NCT01665144 - Exploring the Efficacy and Safety of Siponimod in Patients With Secondary Progressive Multiple Sclerosis (EXPAND). Studienregistereintrag: clinicaltrials.gov. 2019 [Zuletzt aktualisiert 09.09.2019; abgerufen am 26.11.2019]; Abrufbar unter: https://clinicaltrials.gov/ct2/show/NCT01665144.
42. Novartis. NCT01665144 - Exploring the Efficacy and Safety of Siponimod in Patients With Secondary Progressive Multiple Sclerosis (EXPAND). Studienregistereintrag: ICTRP. 2019 [Zuletzt aktualisiert 30.09.2019; abgerufen am 26.11.2019]; Abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01665144.

43. Novartis. JPRN-JapicCTI-142447 -Study of efficacy, safety and tolerability data for BAF312 compared to placebo in patients with secondary progressive multiple sclerosis. Studienregistereintrag: ICTRP. 2019 [Zuletzt aktualisiert 16.07.2019; abgerufen am 26.11.2019]; Abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-JapicCTI-142447.
44. Novartis. EUCTR2012-003056-36-HU -Study of efficacy, safety and tolerability data for BAF312 compared to placebo in patients with secondary progressive multiple sclerosis. Studienregistereintrag: ICTRP. 2018 [Zuletzt aktualisiert 19.11.2018; abgerufen am 26.11.2019]; Abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-003056-36-HU.
45. Novartis. EUCTR2012-003056-36-SK -Study of efficacy, safety and tolerability data for BAF312 compared to placebo in patients with secondary progressive multiple sclerosis followed by extended treatment with BAF312. Studienregistereintrag: ICTRP. 2019 [Zuletzt aktualisiert 28.02.2019; abgerufen am 26.11.2019]; Abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-003056-36-SK.
46. Novartis. 2012-003056-36 -Study of efficacy, safety and tolerability data for BAF312 compared to placebo in patients with secondary progressive multiple sclerosis followed by extended treatment with BAF312. Studienregistereintrag: EU-CTR. 2012 [Zuletzt aktualisiert 14.12.2012; abgerufen am 26.11.2019]; Abrufbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-003056-36.
47. Kappos L, Bar-Or A, Cree BAC, Fox RJ, Giovannoni G, Gold R, Vermersch P, Arnold DL, Arnould S, Scherz T, Wolf C, Wallstrom E, Dahlke F, Investigators EC. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. Lancet. 2018;391(10127):1263-73.
48. Kappos L, Bar-Or A, Cree BAC, Fox RJ, Giovannoni G, Gold R, Vermersch P, Arnold DL, Arnould S, Scherz T, Wolf C, Wallstrom E, Dahlke F, Investigators EC. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. Erratum. Lancet. 2018;392:2170.
49. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol. 2011;69:292-302.
50. MS Research and Project Development gGmbH (MSFP). Datenauswertung aus dem MS-Register von SPMS Patienten. Report. 2019.
51. Novartis. Zusatzauswertungen EXPAND - Gesamtpopulation. 2019.
52. Novartis. Zusatzauswertungen EXPAND - Subpopulation B. 2019.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

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

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	07.11.2016	
Zeitsegment	1974 to 2016 November 04	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle[21]] – Strategy minimizing difference between sensitivity and specificity	
#	**Suchbegriffe**	**Ergebnis**
1	Diabetes Mellitus/	552986
2	Non Insulin Dependent Diabetes Mellitus/	195234
3	(diabet* or niddm or t2dm).ab,ti.	714228
4	or/1-3	847068
5	linagliptin*.mp.	1562
6	(random* or double-blind*).tw.	1193849
7	placebo*.mp.	388057
8	or/6-7	1382838
9	and/4,5,8	633

[21] 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

Thema: Randomisierte Studien zu Siponimod bei SPMS

Datenbanken: MEDLINE, EMBASE, COCHRANE Central Register of Controlled Trials, (keine generellen Einschränkungen)

Suchstrategien:

Datenbank	Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present	
Suchoberfläche	Ovid	
Datum der Suche	18.11.2019	
Zeitsegment	1946 to present (November 15, 2019)	
Suchfilter	Filter für randomisierte kontrollierte Studien nach 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.	
#	Suchbegriffe	Ergebnis
1	siponimod.mp.	83
2	(baf312 or baf-312).mp.	30
3	mayzent.mp.	6
4	or/1-3	87
5	randomized controlled trial.pt.	494657
6	randomized.mp.	813877
7	placebo.mp.	209807
8	or/5-7	875788
9	4 and 8	17

Datenbank	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	18.11.2019	
Zeitsegment	1974 to 2019 November 15	
Suchfilter	Filter für randomisierte kontrollierte Studien nach 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.	
#	Suchbegriffe	Ergebnis
1	siponimod.mp.	295
2	(baf312 or baf-312).mp.	135
3	mayzent.mp.	7
4	exp siponimod	281
5	(1230487-85-0 or 1230487-00-9 or 1234627-85-0).rn.	223
6	or/1-5	329
7	random*.tw.	1477478
8	placebo*.mp.	444316
9	double-blind*.tw.	204488
10	or/7-9	1728301
11	6 and 10	119

Datenbank	COCHRANE Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	18.11.2019	
Zeitsegment	1991—October 2019	
#	Suchbegriffe	
1	siponimod.mp.	63
2	(baf312 or baf-312).mp.	50
3	mayzent.mp.	0
4	or/1-3	82

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken 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 jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. clinicaltrials.gov), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank erreichbar ist (z. B. http://www.clinicaltrials.gov), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07.11.2016
Eingabeoberfläche	Standard Search
Suchstrategie	linagliptin OR BI 1356
Treffer	169

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

Thema:

Siponimod: Randomisierte, kontrollierte Studien bei SPMS

Studienregister	clinicaltrials.gov
Internetadresse	http://www.clinicaltrials.gov
Datum der Suche	18.11.2019
Suchstrategie	Search Terms = siponimod OR baf312 OR baf-312
Treffer	12

Studienregister	ICTRP
Internetadresse	http://www.who.int/trialsearch
Datum der Suche	18.11.20198
Suchstrategie	Search Terms= siponimod OR baf312 OR baf-312
Treffer	61 Treffer (21 Studien)

Studienregister	EU Clinical Trials Register
Internetadresse	https://www.clinicaltrialsregister.eu
Datum der Suche	25.11.2019
Suchstrategie	Search Terms= siponimod OR baf312 OR baf-312
Treffer	9

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

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

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

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

Nr.	Referenz	Ausschluss-grund
1	Yumeng Song, Yongfeng Lao, Fuxiang Liang, Jing Li, Bibo Jia, Zixuan Wang, Xu Hui, Zhenxing Lu, Biao Zhou, Wei Luo, Bing Song. Efficacy and safety of siponimod for multiple sclerosis. Protocol for a systematic review and meta-analysis. Medicine. 2019;98:34(e15415).	Review ohne zusätzlich relevante Studie
2	Baldur P. Magnusson, Heinz Schmidli, Nicolas Rouyrre, Daniel O. Scharfstein. Bayesian inference for a principal stratum estimand to assess the treatment effect in a subgroup characterized by postrandomization event occurrence. Statistics in Medicine. 2019;38:4761–4771.	Anderer Studientyp
3	H-C Diener. Siponimod ist wirksamer als Placebo: Ergebnisse einer doppelblinden randomisierten Phase-III-Studie. AMT. 2018; 36:7-8.	Review ohne zusätzlich relevante Studie
4	Weideman AM, Tapia-Maltos MA, Johnson K, Greenwood M and Bielekova B (2017) Meta-analysis of the Age-Dependent Efficacy of Multiple Sclerosis Treatments. Front. Neurol. 8:577.	Review ohne zusätzlich relevante Studie
5	Malihe Safavi, Shekoufeh Nikfar and Mohammad Abdollahi. A Systematic Review of Drugs in Late-Stage Development for the Treatment of Multiple Sclerosis: A Focus on Oral Synthetic Drugs. Inflammation & Allergy - Drug Targets, 2014, 13, 351-366.	Review ohne zusätzlich relevante Studie

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

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

Listen Sie nachfolgend die durch die Studienregistersuche(n)/ Studienergebnisdatenbanksuche(n) identifizierten, aber ausgeschlossenen Registereinträgen 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

ClinicalTrials.gov

Nr.	Referenz	Ausschlussgrund
01	NCT03338998. Novartis Pharmaceuticals. Efficacy, Safety and Tolerability of BAF312 Compared to Placebo in Patients With Intracerebral Hemorrhage (ICH). Zuletzt aktualisiert: 01.10.2019; abgerufen am: 26.11.2018; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT03338998	Andere Population
02	NCT02029274. Novartis Pharmaceuticals. Safety and Efficacy of BAF312 in Dermatomyositis. Zuletzt aktualisiert: 15.01.2019; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT02029274	Andere Population
03	NCT02330965. National Institute of Allergy and Infectious Diseases (NIAID). Mechanistic Studies of Phase III Trial With BAF312 in Secondary Progressive Multiple Sclerosis. Zuletzt aktualisiert: 07.11.2017; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT02029274	Keine Ergebnisse
04	NCT01801917. Novartis Pharmaceuticals. Efficacy and Tolerability of BAF312 in Patients With Polymyositis. Zuletzt aktualisiert: 04.01.2018; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT01801917	Andere Population
05	NCT01565902. Novartis Pharmaceuticals. Pharmacokinetics of BAF312 in Patients With Hepatic Impairment. Zuletzt aktualisiert: 27.03.2018; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT01565902	Andere Population
06	NCT01904214. Novartis Pharmaceuticals. A Single-dose, Open-label, Parallel-group Study to Assess the Pharmacokinetics of BAF312 in Subjects With Renal	Andere Population

	Impairment Compared to Subjects With Normal Renal Function. Zuletzt aktualisiert: 28.02.2018; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT01904214	
07	NCT01185821. Novartis Pharmaceuticals. Long-term Safety, Tolerability and Efficacy of BAF312 Given Orally in Patients With Relapsing-remitting Multiple Sclerosis. Zuletzt aktualisiert: 27.03.2018; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/sho.w/record/NCT01185821	Andere Population
08	NCT01148810. Novartis Pharmaceuticals. Efficacy and Tolerability of BAF312 in Patients With Polymyositis and Dermatomyositis. Zuletzt aktualisiert: 25.04.2019; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT01148810	Andere Population
09	NCT03623243. Novartis Pharmaceuticals. Safety and Tolerability of Conversion From Oral or Injectable Disease Modifying Therapies to Dose-titrated Oral Siponimod in Advancing RMS Patients. (EXCHANGE). Zuletzt aktualisiert: 29.03.2019; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT03623243	Anderer Komparator
10	NCT00879658. Novartis Pharmaceuticals. Safety, Tolerability, Efficacy and Optimal Dose Finding Study of BAF312 in Patients With Relapsing-remitting Multiple Sclerosis. Zuletzt aktualisiert: 02.03.2018; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT00879658	Andere Population
11	NCT00422175. Novartis. Single Dose Study to Assess the Safety, Tolerability, Concentration in Body and Effect of BAF312. Zuletzt aktualisiert: 29.11.2017; abgerufen am: 26.11.2019; abrufbar unter: https://clinicaltrials.gov/ct2/show/record/NCT00422175	Andere Population

ICTRP

Nr	**Referenz**	Ausschluss-grund
01	NCT03623243. Novartis Pharmaceuticals. Exploring the Safety and Tolerability of Conversion From Oral or Injectable Disease Modifying Therapies to Dose-titrated Oral Siponimod in Patients With Advancing Forms of Relapsing Multiple Sclerosis: A 6-month Open Label, Multi-Center Phase IIIb Study. Zuletzt aktualisiert: 08.04.2019; abgerufen am:	Anderer Studientyp

	27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03623243	
02	EUCTR2015-005738-23-DE. Novartis Pharma AG. A phase II, patient and investigator-blinded, randomized, placebo-controlled study to evaluate efficacy, safety and tolerability of BAF312 in patients with stroke due to intracerebral hemorrhage (ICH). Zuletzt aktualisiert: 28.02.2019; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2015-005738-23-DE	Andere Population
03	NCT03338998. Novartis Pharmaceuticals. A Phase II, Patient- and Investigator-blinded, Randomized, Placebo-controlled Study to Evaluate Efficacy, Safety and Tolerability of BAF312 (Siponimod) in Patients With Stroke Due to Intracerebral Hemorrhage (ICH). Zuletzt aktualisiert: 14.10.2019; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT03338998	Andere Population
04	NCT02330965. National Institute of Allergy and Infectious Diseases (NIAID). Mechanistic Studies of Phase III Trial With BAF312 in Secondary Progressive Multiple Sclerosis (AMS04). Zuletzt aktualisiert: 16.12.2017; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02330965	Keine Ergebnisse
05	EUCTR2013-001799-39-HU. Novartis Pharma Services AG. A double blind, randomized, placebo-controlled study to evaluate, safety, tolerability, efficacy and preliminary dose-response of BAF312 in patients with active dermatomyositis. Zuletzt aktualisiert: 11.04.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-001799-39-HU	Andere Population
	EUCTR2013-001799-39-CZ. Novartis Pharma Services AG. A double blind, randomized, placebo-controlled study to evaluate, safety, tolerability, efficacy and preliminary dose-response of BAF312 in patients with active dermatomyositis. Zuletzt aktualisiert: 11.04.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-001799-39-CZ	Duplikat
	EUCTR2013-001799-39-PL. Novartis Pharma Services AG. A double blind, randomized, placebo-controlled study to evaluate, safety, tolerability, efficacy and preliminary dose-response of BAF312 in patients with active dermatomyositis. Zuletzt aktualisiert: 25.04.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-001799-39-PL	Duplikat

	EUCTR2013-001799-39-BE. Novartis Pharma Services AG. A double blind, randomized, placebo-controlled study to evaluate, safety, tolerability, efficacy and preliminary dose-response of BAF312 in patients with active dermatomyositis. Zuletzt aktualisiert: 16.08.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2013-001799-39-BE	Duplikat
	NCT02029274. Novartis Pharmaceuticals. A double blind, randomized, placebo-controlled study to evaluate, safety, tolerability, efficacy and preliminary dose-response of BAF312 in patients with active dermatomyositis. Zuletzt aktualisiert: 19.09.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT02029274	Duplikat
06	NCT01904214. Novartis Pharmaceuticals. A Single-dose, Open-label, Parallel-group Study to Assess the Pharmacokinetics of BAF312 in Subjects With Renal Impairment Compared to Subjects With Normal Renal Function. Zuletzt aktualisiert: 12.03.2018; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01904214	Andere Population
07	ACTRN12613000545763. Commercial sector/Industry Novartis Pharmaceuticals Australia Pty Limited. Open-label study to assess the pharmacokinetics, safety and tolerability of siponimod in healthy subjects with CYP2C9 extensive (EM) and poor metabolizer (PM) phenotype. Zuletzt aktualisiert: 04.01.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=ACTRN12613000545763	Andere Population
08	NCT01801917. Novartis Pharmaceuticals. A Multi-centre Double-blind, Placebo Controlled, Proof of Concept Study to Evaluate the Efficacy and Tolerability of BAF312 in Patients With Polymyositis. Zuletzt aktualisiert: 18.01.2018; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01801917	Andere Population
09	EUCTR2012-002859-42-HU. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis. Zuletzt aktualisiert: 21.11.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-002859-42-HU	Andere Population
	EUCTR2012-002859-42-CZ. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with	Duplikat

	polymyositis. Zuletzt aktualisiert: 05.09.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-002859-42-CZ	
	EUCTR2012-002859-42-PL. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis. Zuletzt aktualisiert: 17.10.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-002859-42-PL	Duplikat
	EUCTR2012-002859-42-BE. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis. Zuletzt aktualisiert: 30.04.2018; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-002859-42-BE	Duplikat
10	NCT01565902. Novartis Pharmaceuticals. A Single-dose, Open-label, Parallel-group Study to Assess the Pharmacokinetics of BAF312 in Subjects With Mild, Moderate and Severe Hepatic Impairment Compared to Healthy Control Subjects. Zuletzt aktualisiert: 19.10.2017; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01565902	Andere Population
	EUCTR2012-000562-37-HU. Novartis Pharma Services AG. A Single-dose, Open-label, Parallel-group Study to Assess the Pharmacokinetics of BAF312 in Subjects With Mild, Moderate and Severe Hepatic Impairment Compared to Healthy Control Subjects. Zuletzt aktualisiert: 12.05.2014; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2012-000562-37-HU	Duplikat
11	EUCTR2011-006043-30-NL. Novartis Pharma Services AG. An open-label, multiple-dose, two-treatment period study to evaluate the effect of oral BAF312 on the pharmacokinetices and pharmacodynamics of a monophasic oral contraceptive in healthy female volunteers. Zuletzt aktualisiert: 25.09.2012; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2011-006043-30-NL	Andere Population
12	EUCTR2009-014392-51-ES. Novartis Farmacéutica S.A. Fase de extensión del estudio CBAF312A2201 con enmascaramiento de la dosis para evaluar la seguridad, tolerabilidad y eficacia a largo plazo de BAF312	Andere Population

	administrado por vía oral una vez al día a pacientes con esclerosis múltiple recurrente. Zuletzt aktualisiert: 07.08.2012; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-014392-51-ES	
	EUCTR2009-014392-51-IT. Novartis Farma. An dose blinded extension study to the CBAF312A2201 study to evaluate long-term safety, tolerability and efficacy of BAF312 given orally once daily in patients with relapsing multiple sclerosis. Zuletzt aktualisiert: 14.03.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-014392-51-IT	Duplikat
	EUCTR2009-014392-51-DE. Novartis Pharma Services AG. An extension study to the CBAF312A2201 study to evaluate long-term safety, tolerability and efficacy of BAF312 given orally once daily in patients with relapsing multiple sclerosis. Zuletzt aktualisiert: 11.04.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-014392-51-DE	Duplikat
	EUCTR2009-014392-51-PL. Novartis Pharma Services AG. An extension study to the CBAF312A2201 study to evaluate long-term safety, tolerability and efficacy of BAF312 given orally once daily in patients with relapsing multiple sclerosis. Zuletzt aktualisiert: 19.12.2016; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-014392-51-PL	Duplikat
	EUCTR2009-014392-51-HU. Novartis Pharma Services AG. An extension study to the CBAF312A2201 study to evaluate long-term safety, tolerability and efficacy of BAF312 given orally once daily in patients with relapsing multiple sclerosis. Zuletzt aktualisiert: 09.01.2017; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-014392-51-HU	Duplikat
	EUCTR2009-014392-51-FI. Novartis Pharma Services AG. An extension study to the CBAF312A2201 study to evaluate long-term safety, tolerability and efficacy of BAF312 given orally once daily in patients with relapsing multiple sclerosis.Zuletzt aktualisiert: 09.01.2017; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2009-014392-51-FI	Duplikat

	NCT01185821. Novartis Pharmaceuticals. A Dose Blinded Extension Study to the CBAF312A2201 Study to Evaluate Long-term Safety, Tolerability and Efficacy of BAF312 Given Orally Once Daily in Patients With Relapsing-remitting Multiple Sclerosis. Zuletzt aktualisiert: 16.12.2017; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01185821	Duplikat
13	EUCTR2008-006311-21-SE. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis and dermatomyositis. Zuletzt aktualisiert: 17.06.2013; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-006311-21-SE	Andere Population
	EUCTR2008-006311-21-GB. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis and dermatomyositis. Zuletzt aktualisiert: 10.12.2019; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-006311-21-GB	Duplikat
	EUCTR2008-006311-21-HU. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis and dermatomyositis. Zuletzt aktualisiert: 09.10.2012; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-006311-21-HU	Duplikat
	EUCTR2008-006311-21-CZ. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis and dermatomyositis. Zuletzt aktualisiert: 09.10.2012; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-006311-21-CZ	Duplikat
	NCT01148810. Novartis Pharmaceuticals. A Multi-centre, Double-blind, Placebo Controlled, Proof of Concept Study to Evaluate the Efficacy and Tolerability of BAF312 in Patients With Polymyositis and Dermatomyositis. Zuletzt aktualisiert: 19.02.2015; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT01148810	Duplikat

14	NCT00879658. Novartis Pharmaceuticals. A Phase II, Double-blind, Randomized, Multi-center, Adaptive Dose-ranging, Placebo-controlled, Parallel-group Study Evaluating Safety, Tolerability and Efficacy on MRI Lesion Parameters and Determining the Dose Response Curve of BAF312 Given Orally Once Daily in Patients With Relapsing-remitting Multiple Sclerosis. Zuletzt aktualisiert: 12.03.2018; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00879658	Andere Population
15	EUCTR2008-008719-25-DE. Novartis Pharma Services AG. A phase II, double-blind, randomized, multi-center, adaptive dose-ranging, placebo-controlled, parallel-group study evaluating safety, tolerability and efficacy on MRI lesion parameters and determining the dose response curve of BAF312 given orally once daily in patients with relapsing-remitting multiple sclerosis. Zuletzt aktualisiert: 01.05.2012; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-008719-25-DE	Andere Population
	EUCTR2008-008719-25-FI. Novartis Pharma Services AG. A phase II, double-blind, randomized, multi-center, adaptive dose-ranging, placebo-controlled, parallel-group study evaluating safety, tolerability and efficacy on MRI lesion parameters and determining the dose response curve of BAF312 given orally once daily in patients with relapsing-remitting multiple sclerosis. Zuletzt aktualisiert: 01.05.2012; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-008719-25-FI	Duplikat
	EUCTR2008-008719-25-HU. Novartis Pharma Services AG. A phase II, double-blind, randomized, multi-center, adaptive dose-ranging, placebo-controlled, parallel-group study evaluating safety, tolerability and efficacy on MRI lesion parameters and determining the dose response curve of BAF312 given orally once daily in patients with relapsing-remitting multiple sclerosis. Zuletzt aktualisiert: 01.05.2012; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-008719-25-HU	Duplikat
	EUCTR2008-008719-25-IT. Novartis Farma. A phase II, double-blind, randomized, multi-center, adaptive dose-ranging, placebo-controlled, parallel-group study evaluating safety, tolerability and efficacy on MRI lesion parameters and determining the dose response curve of BAF312 given orally once daily in patients with relapsing-remitting multiple sclerosis. Zuletzt aktualisiert: 01.05.2012; abgerufen am: 27.11.2019; abrufbar unter:	Duplikat

	http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-008719-25-IT	
	EUCTR2008-008719-25-ES. Novartis Farmaceutica S.A. Estudio fase II, a doble ciego, aleatorizado y multicéntrico, con esquema adaptativo de dosis, controlado con placebo y de grupos paralelos, para evaluar la eficacia, tolerabilidad y seguridad en base a parámetros de lesión observados mediante resonancia magnética, así como determinar la curva dosis - respuesta de BAF312 administrado oralmente en dosis única diaria en pacientes con Esclerosis Múltiple remitente recurrente. Zuletzt aktualisiert: 01.05.2012; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=EUCTR2008-008719-25-ES	Duplikat
16	NCT00422175. Novartis. A First-in-Human Study for BAF312: A Two Parts, Single Center, Randomized, Double-Blind, Placebo-Controlled Ascending Single Dose Study to Explore the Safety, Tolerability, Pharmacokinetics and Pharmacodynamics of Oral BAF312 in Healthy Volunteers. Zuletzt aktualisiert: 19.02.2015; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=NCT00422175	Andere Population
17	JPRN-JapicCTI-142590. Novartis Pharma K.K. A Double Blind, Randomized, Placebo-controlled Study to Evaluate, Safety, Tolerability, Efficacy and Preliminary Dose-response of BAF312 in Patients With Active Dermatomyositis. Zuletzt aktualisiert: 02.04.2019; abgerufen am: 27.11.2019; abrufbar unter: http://apps.who.int/trialsearch/Trial2.aspx?TrialID=JPRN-JapicCTI-142590	Andere Population

EU-CTR

Nr.	**Referenz**	**Ausschlussgrund**
01	2012-002859-42. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis. Zuletzt aktualisiert: 04.08.2016; abgerufen am: 27.11.2019; abrufbar unter: //www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-002859-42	Andere Population
02	2013-001799-39. Novartis Pharma Services AG. A double blind, randomized, placebo-controlled study to evaluate, safety, tolerability, efficacy and preliminary dose-response of BAF312 in patients with active dermatomyositis. Zuletzt aktualisiert: 16.10.2014; abgerufen am: 27.11.2019; abrufbar unter:	Andere Population

	//www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2013-001799-39	
03	2009-014392-51. Novartis Pharma Services AG. An extension study to the CBAF312A2201 study to evaluate long-term safety, tolerability and efficacy of BAF312 given orally once daily in patients with relapsing multiple sclerosis. Zuletzt aktualisiert: 10.10.2016; abgerufen am: 27.11.2019; abrufbar unter: //www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2009-014392-51	Andere Population
04	2015-005738-23. Novartis Pharma AG. A phase II, patient and investigator-blinded, randomized, placebo-controlled study to evaluate efficacy, safety and tolerability of BAF312 in patients with stroke due to intracerebral hemorrhage (ICH). Zuletzt aktualisiert: 15.06.2018; abgerufen am: 27.11.2019; abrufbar unter: //www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2015-005738-23	Andere Population
05	2011-006043-30. Novartis Pharma Services AG. An open-label, multiple-dose, two-treatment period study to evaluate the effect of oral BAF312 on the pharmacokinetices and pharmacodynamics of a monophasic oral contraceptive in healthy female volunteers. Zuletzt aktualisiert: 27.02.2012; abgerufen am: 27.11.2019; abrufbar unter: //www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2011-006043-30	Andere Population
06	2012-000562-37. Novartis Pharma Services AG. A Single-dose, Open-label, Parallel-group Study to Assess the Pharmacokinetics of BAF312 in Subjects With Mild, Moderate and Severe Hepatic Impairment Compared to Healthy Control Subjects. Zuletzt aktualisiert: 04.03.2014; abgerufen am: 27.11.2019; abrufbar unter: //www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2012-000562-37	Andere Population
07	2008-006311-21. Novartis Pharma Services AG. A multi-centre, double-blind, placebo controlled, proof of concept study to evaluate the efficacy and tolerability of BAF312 in patients with polymyositis and dermatomyositis. Zuletzt aktualisiert: 13.06.2012; abgerufen am: 27.11.2019; abrufbar unter: //www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2008-006311-21	Andere Population
08	2008-008719-25. Novartis Pharma Services AG. A phase II, double-blind, randomized, multi-center, adaptive dose-ranging, placebo-controlled, parallel-group study evaluating safety, tolerability and efficacy on MRI lesion parameters and determining	Andere Population

	the dose response curve of BAF312 given orally once daily in patients with relapsing-remitting multiple sclerosis. Zuletzt aktualisiert: 04.05.2011; abgerufen am: 27.11.2019; abrufbar unter: //www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number: 2008-008719-25	

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-187 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 4-187 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-187 (Anhang): Studiendesign und -methodik für Studie EXPAND

Item[a]	Charakteristikum	Studieninformation
Studienziel		
2 b	Genaue Ziele, Fragestellung und Hypothesen	Zweck der Studie CBAF312A2304 (im Folgenden: Studie A2304) war die Gewinnung von doppelt-verblindeten und kontrollierten Daten zur Wirksamkeit, Sicherheit und Verträglichkeit bei Patienten mit SPMS; weiterhin die Gewinnung von Langzeitdaten zur Sicherheit, Verträglichkeit und Wirksamkeit von Siponimod.
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	Die Kernphase bildete eine randomisierte, doppelt verblindete, placebokontrollierte, multizentrische Studie im Parallelgruppen-Design mit SPMS-Patienten. Geeignete Patienten wurden im Verhältnis 2:1 randomisiert und erhielten entweder Siponimod oder Placebo.
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	Das Originalprotokoll datiert vom 12. Juli 2012. Das Protokoll wurde viermal per Amendment geändert. **Amendment 1** Zum Zeitpunkt von Amendment 1 (26. Mai 2014) lief die Rekrutierung noch, über 1.200 Patienten waren gescreent worden. Die wesentlichen Veränderungen sind im Folgenden zusammengefasst: • Im Originalprotokoll war bestimmt, dass aufgrund eines im Vorfeld geplanten DMC-Reviews von kardiologischen Monitoring-Ergebnissen zur Therapieinitiierung (basierend auf Daten der ersten Hälfte rekrutierter Patienten) die Monitoring-Vorgaben in der zweiten Rekrutierungshälfte bei Patienten mit normalem kardiovaskulären Status gelockert werden können. Basierend auf den verfügbaren kardiologischen Monitoring-Ergebnissen aus der Therapieinitiierung bestätigte das DMC, dass die sechsstündige Nachbeobachtung nach der ersten Dosis an den Tagen 1 und 7 bei Patienten in normalem kardiologischem Zustand entfallen kann. Die kontinuierliche Datenerfassung mit mobiler kardialer Telemetrie und Langzeit-EKG wurde aber für alle Patienten befürwortet, um die Erhebung jeglicher relevanter Ereignisse während der Therapieinitiierung zu erleichtern. • Harmonisierung der Behandlungsphase zwischen Protokoll und eCRF. • Erlaubnis von Betablockern. • Modifizierung der Ein- und Ausschlusskriterien hinsichtlich Empfängnisverhütung auf Basis der Rückmeldungen von Gesundheitsbehörden. Weiterhin wurden Ausschlusskriterien für zwei weitere Medikamente formuliert, die erst nach Abschluss des Originalprotokolls zur Behandlung

		der MS zugelassen worden waren (Dimethylfumarat und Teriflunomid). • Streichung einer explorativen kognitiven Substudie. • Aktualisierung der Details bezüglich potenzieller CYP2C9-Inhibitoren. • Jeder Patient, der irgendeine Dosis in der Titrationsphase versäumt hatte, wurde verpflichtet, die Titration von vorne zu beginnen. **Amendment 2** Zum Zeitpunkt von Amendment 2 (5. März 2015) war die Rekrutierung nahezu abgeschlossen. > 2.000 Patienten waren gescreent und > 1.500 Patienten randomisiert. Die wesentlichen Änderungen sind im Folgenden zusammengefasst: • Aktualisierung der Regeln und des Timings für die geplante Überprüfung der Futility-Analyse. • Ergänzung klärender und spezifischer Hinweise für die Begleittherapie hinsichtlich des möglichen Absetzens einer Behandlung mit Betablockern bei Patienten mit einem Ruhepuls ≤ 50 bpm vor dem Start der Behandlung mit Siponimod/Placebo. • Aktualisierung der Details bezüglich der Notfallentblindung. • Erwähnung des Abschlusses des verblindeten Daten-Reviews für eine Reevaluierung der Stichprobengröße. Hier wurde insbesondere bestätigt, dass die Zahl der einzuschließenden Patienten nicht mehr steigen würde. **Amendment 3** Zum Zeitpunkt von Amendment 3 (6. Oktober 2015) war die Rekrutierung abgeschlossen. Die wesentlichen Veränderungen sind im Folgenden zusammengefasst: • Eine Extensionsphase wurde ins Studiendesign aufgenommen. • Wenn zutreffend, wurden die Bezeichnungen der Kernphase (Core Part) und Extensionsphase (Extension Part) in den Protokollelementen aufgenommen. • Das Kriterium für das Ende der Kernphase der Studie wurde aktualisiert. Konkret sollte die Kernphase der Studie ungefähr drei Jahre nach der Randomisierung des ersten Patienten abgeschlossen werden. Zu diesem Zeitpunkt, so die Abschätzung, wären mindestens 374 Patienten mit einer nach drei Monaten bestätigten Behinderungsprogression erfasst worden, und bei mehr als 95 % der Patienten läge die Randomisierung mindestens ein Jahr (oder länger) zurück. **Amendment 4** Die wesentlichen Änderungen aus Amendment 4 (23. März 2016) sind im Folgenden zusammengefasst: • Spezifizierung und Begründung der Gesamtstudiendauer im Protokoll, um die

Item[a]	Charakteristikum	Studieninformation
		unterschiedlichen Vorschriften in den verschiedenen Ländern aufzunehmen. • Die Visite zu Monat 1 der Extensionsphase soll für diejenigen Patienten optional sein, die in der Kernphase bis unmittelbar vor Beginn der Extensionsphase bereits für mindestens sechs Monate unverblindet Siponimod erhalten hatten. Damit sollen wiederholte Untersuchungen bei diesen Patienten vermieden werden. • Klargestellt wurde, dass die Visite zu Tag 4 der Extensionsphase nur bei denjenigen Patienten stattfinden sollte, für die ein Monitoring der Erstdosierung notwendig war und bei denen zu diesem Zweck ein Langzeit-EKG geschrieben wurde. Die Erfassung der Vitalparameter war optional, da der Hauptzweck dieser Visite die Anbringung des neuen Langzeit-EKG-Gerätes war. • Klargestellt wurde, dass die Visite zu Tag 7 für diejenigen Patienten nicht notwendig war, die aus der unverblindeten Behandlung mit Siponimod ohne eine Behandlungspause von vier Tagen in die Extensionsphase eintraten. Grund ist, dass bei diesen Patienten kein Monitoring der Erstdosierung notwendig war. • Klargestellt wurde, dass die Erfassung der Biomarkerprobe zu Visite 301 nicht notwendig war, sofern diese Visite innerhalb der vier Wochen nach der EOS-Visite stattfand. • Entfernt wurde der Passus, dass eine Fortsetzung in der Extensionsphase für „mindestens ein Jahr nach Zulassung“ erfolgen soll. Damit sollte Konsistenz mit dem Untersuchungsplan über 84 Monate gewahrt werden. Ergänzt wurde die wissenschaftliche Rationale für eine Extensionsphase über 84 Monate. Ebenfalls entfernt wurde der Verweis auf eine andere Open-Label-Studie zu Siponimod, damit der Endpunkt dieser Studie nicht gefährdet wird. • Aus dem Untersuchungsplan wurde die Visite zu Monat 84 entfernt und als EOS-Visite bezeichnet. Diese Änderung untermauert die Gesamtstudiendauer von 84 Monaten. Die EOS-Visite / Visite zu Monat 84 sollte bei jedem Patienten stattfinden, bei dem alle anderen Visiten in der Extensionsphase (bis Monat 78) vorgenommen wurden, oder bei denjenigen Patienten, die die Studie egal aus welchem Grund vorzeitig abgebrochen haben. • Die SUE-Terminologie wurde angepasst, und bei den ethischen Erwägungen wurde eine Ergänzung vorgenommen. Damit sollte den neuesten Richtlinien eines neuen Protokoll-Templates entsprochen werden. • Der Passus zur frühzeitigen Beendigung der Studie wurde aktualisiert, um damit Gründe basierend auf

Item[a]	Charakteristikum	Studieninformation
		der Risiko-/Nutzenabwägung sowie regulatorische oder medizinische Gründe aufzunehmen. Damit sollte der aktuellen Vorlage entsprochen werden. • Entfernt wurde der Passus zum elektronischen EDSS, um damit dem Studiensponsor einen Wechsel auf andere EDSS-Erfassungen zu ermöglichen. • Im Statistikteil wurden Änderungen vorgenommen, um den Baseline-Wert als Kovariate für die T25W-Analyse einzuschließen (um Konsistenz mit anderen Analysen herzustellen). Außerdem sollten damit Änderungen bezüglich der Analysen der UE und der Analysen der Leberfunktionstests vorgenommen werden. Da diese Ergänzungen (Amendments) nur geringfügig war und vor der Verblindung der Studie veröffentlicht wurden, wird angenommen, dass sie die Interpretation der Studienergebnisse nicht beeinflussen. **Andere Änderungen in der Durchführung der Studie** Das NESC-System (Neurostatus eScoring) kam zum Einsatz, um in dieser Studie die Daten des EDSS zu erfassen. Damit sollten Schwankungen und Rechenfehler reduziert und die Datenqualität verbessert werden. Die Daten des EDSS wurden im Behandlungszentrum mittels eines elektronischen Erfassungstools (NESC iPad) erfasst. Durch einen integrierten Algorithmus war für den Erfasser der EDSS-Daten im Zentrum ein sofortiges Feedback (bis zu vier Durchgänge) möglich. Nachdem der Erfasser der EDSS-Daten im Zentrum die Erhebung abgeschlossen hatte und nachdem die Daten zu einer zentralen NESC-Database übertragen worden waren, wurden jegliche noch bestehenden Inkonsistenzen von einem EDSS-Experten überprüft. Sofern notwendig, konnten von diesem weitere Rückfragen zum EDSS-Erfasser im Zentrum übermittelt werden. Ganz zu Anfang der Studie war die NESC-iPad-Technologie noch nicht verfügbar. Die EDSS-Daten wurden in Papierdokumenten erfasst, nach Veröffentlichung des NESC-eScoring wurden sie dort eingegeben. Kurz vor Schließung der Datenbank wurden die Aktivitäten des Neurostatus eScoring der Kernstudie vorzeitig beendet. Damit konnte die Abstimmung zwischen der Novartis-Datenbank und externen EDSS-Daten nicht abgeschlossen werden. Zur Erfassung des Einflusses von Unstimmigkeiten wurde eine Sensitivitätsanalyse vorgenommen.
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	**Einschlusskriterien** Patienten mussten für den Einschluss in die Studie folgende Kriterien erfüllen: 1. Bevor irgendwelche Erhebungen vorgenommen wurden, musste eine schriftliche Einverständniserklärung vorliegen.

Item[a]	Charakteristikum	Studieninformation
		2. Patienten oder Patientinnen im Alter zwischen 18 und einschließlich 60 Jahren (zum Zeitpunkt des Screenings). 3. Schubförmig-remittierende Multiple Sklerose (RRMS) in der Anamnese gemäß den überarbeiteten McDonald-Kriterien von 2010. 4. SPMS, definiert durch eine Behinderungsprogression (über mindestens sechs Monate) ohne auftretende Schübe oder unabhängig von Schüben. • schriftlich durch den Prüfarzt bestätigt, dass die Krankheit in das Stadium der Progredienz eingetreten ist (gemäß Studiendefinition), und zwar mindestens sechs Monate vor der Rekrutierung. 5. Behinderungsgrad beim Screening mit einem EDSS-Score von 3,0 bis 6,0 (einschließlich). 6. Dokumentierte EDSS-Progression in den zwei Jahren vor der Studie in Höhe von ≥ 1 Punkt bei Patienten mit EDSS < 6,0 zum Zeitpunkt des Screenings und ≥ 0,5 Punkte bei Patienten mit EDSS ≥ 6,0 zum Zeitpunkt des Screenings. Sofern die dokumentierten EDSS-Scores nicht vorlagen, musste zum zentralen Review Folgendes eingereicht werden: schriftliche Zusammenfassung der klinischen Evidenz der Behinderungsprogression in den letzten zwei Jahren sowie eine retrospektive Bewertung des EDSS-Scores aus Daten im Zeitraum von bis zu zwei Jahren vor dem Screening. 7. Kein Hinweis auf einen Schub oder eine Kortikosteroid-Behandlung im Zeitraum von drei Monaten vor der Randomisierung. **Entscheidungsprozess zur Eignung** Wie oben beschrieben, besagte Einschlusskriterium 6, dass sich bei Patienten in den vorangegangenen zwei Jahren eine Behinderungsprogression gezeigt haben musste. Dabei musste bei Patienten mit einem EDSS von < 6,0 zu Studienbeginn eine EDSS-Progression von ≥ 1 Punkt vorliegen ODER bei Patienten mit einem EDSS von ≥ 6,0 zu Studienbeginn ein Anstieg um ≥ 0,5 Punkte. Falls die der Studie vorausgehendenden Erfassungen des EDSS aus den zwei Jahren vor der Studie nicht verfügbar waren, um diese Bewertung zu ermöglichen, wurde durch das Entscheidungskomitee ein zentraler Review der verfügbaren klinischen Evidenz der Behinderungsprogression vorgenommen (Appendix 16.1.1-Protokollteil 8.5). In diesen Fällen wurde durch den Prüfarzt ein Formular zur Evidenz der Behinderungsprogression vollständig ausgefüllt und eingereicht. In diesem Bogen wurde die vorherige Evidenz aus anderen Quellen zusammengefasst, so etwa Befunde aus neurologischen Untersuchungen, Krankheitsgeschichte et cetera. Damit war es dem Entscheidungskomitee möglich zu bewerten, ob der Patient geeignet war. **Ausschlusskriterien** Patienten, die irgendeines der Ausschlusskriterien erfüllen, waren nicht für die Studie geeignet. Vom Prüfarzt wurden

Item[a]	Charakteristikum	Studieninformation
		keine weiteren Ausschlüsse vorgenommen, womit sichergestellt werden sollte, dass die Studienpopulation für alle geeigneten Patienten repräsentativ war. Grundsätzlich waren keine erneuten Patientenrekrutierungsversuche erlaubt (d. h. nachdem der Patient ursprünglich als ungeeignet eingestuft worden war). In bestimmten Fällen, bei denen sich die Verfassung des Patienten geändert hatte, war der Versuch einer erneuten Rekrutierung erlaubt (z. B. waren Patienten von der Studie ausgeschlossen, die negativ auf Varizella-Zoster-Virus-ImmunglobulinG-Antikörper (IgG) waren; allerdings konnten diese Patienten zwei Monate nach Impfung eingeschlossen werden, sofern sie alle Bewertungskriterien zum Screening wiederholt absolviert haben und bestanden haben.). Falls ein Patient einen oder mehrere der Laborwerte (oder andere) Bewertungskriterien nicht erfüllte, konnte(n) die Auswertung(en) innerhalb des Screeningprozesses nach Ermessen des Prüfarztes wiederholt werden. Falls die Kriterien dann erfüllt wurden, konnte der Patient eingeschlossen werden, vorausgesetzt, dass die Erhebungen innerhalb der Screening-Phase abgeschlossen wurden. 1. Patienten mit aktiver chronischer Erkrankung des Immunsystems außer MS (oder stabiler chronischer Erkrankung des Immunsystems außer MS, jedoch mit Immuntherapie behandelt), oder Patienten mit bekanntem Immunschwächesyndrom. 2. Schwangere oder stillende (laktierende) Frauen. 3. Gebärfähige Frauen, definiert als alle Frauen, die physiologisch in der Lage sind, schwanger zu werden, außer sie wenden während der Studienbehandlung und bis 30 Tage danach hocheffektive Verhütungsmethoden an. 4. Malignität in irgendeinem Organsystem (außer Basalzellkarzinom der Haut) in den letzten fünf Jahren, behandelt oder unbehandelt, unabhängig davon, ob es Hinweise eines lokalen Wiederauftretens oder auf Metastasen gab. 5. Diabetes mellitus, außer bei guter Kontrolle und ohne Komplikationen in den Organen, wie etwa verminderte Nierenfunktion, signifikante Netzhautpathologie oder Neuropathie. 6. Diagnose eines Makulaödems während der Phase vor der Randomisierung. Patienten mit Makulaödem in der Anamnese durften an der Studie unter der Voraussetzung teilnehmen, dass sie bei der augenärztlichen Untersuchung beim Screening kein Makulaödem aufwiesen. 7. Patienten mit aktiven systemischen bakteriellen oder viralen Infektionen oder Pilzinfektionen. Oder Patienten, bei denen eine AIDS-Erkrankung bekannt ist oder die HIV-positiv sind. 8. In der Screening-Phase positiv auf serologische Marker einer Hepatitis A, B, C und E getestet, die auf eine akute oder chronische Infektion hindeuten.

Item[a]	Charakteristikum	Studieninformation
		9. Beim Screening negativer Test auf VZV-IgG-Antikörper (Varizella-Zoster-Virus). 10. Impfung mit Lebendimpfstoffen oder abgeschwächten Lebendimpfstoffen (einschließlich Varizella-Zoster-Virus oder Masern) im Zeitraum von zwei Monaten vor der Randomisierung. 11. Behandlung mit einem der im Folgenden aufgeführten Medikamente (Anmerkung: falls vorher mit IFN-beta oder Glatirameracetat behandelt worden war, war keine Auswaschphase notwendig): • Siponimod, unabhängig vom Zeitpunkt. • Fingolimod in einem Zeitraum von zwei Monaten vor der Randomisierung oder Fingolimod-Therapie über einen Zeitraum von mehr als sechs Monaten. • Immunglobulin intravenös in einem Zeitraum von zwei Monaten vor der Randomisierung. • Dimethylfumarat in einem Zeitraum von zwei Monaten vor der Randomisierung. • Natalizumab in einem Zeitraum von sechs Monaten vor der Randomisierung. • Immunsuppressiva/Chemotherapie (z. B. Azathioprin, Methotrexat) in einem Zeitraum von sechs Monaten vor der Randomisierung. • Cyclophosphamid in einem Zeitraum von einem Jahr vor der Randomisierung. • Rituximab, Ofatumumab, Ocrelizumab oder Cladribin in einem Zeitraum von zwei Jahren vor der Randomisierung. • Alemtuzumab, unabhängig vom Zeitpunkt. • Jegliche Art von Mitoxantron-Behandlung in einem Zeitraum von zwei Jahren vor der Randomisierung oder Hinweis auf Kardiotoxizität in Folge einer Mitoxantron-Gabe oder eine kumulative Mitoxantron-Gabe von mehr als 60 mg/m^2. • Teriflunomid in einem Zeitraum von zwei Jahren vor der Randomisierung (außer wenn die Teriflunomid-Konzentration im Plasma null war oder ohne bedeutsame biologische Signifikanz) ODER innerhalb eines Zeitraums von zwei Wochen vor der Randomisierung nach einer erfolgreichen beschleunigten Elimination (wie in der Fachinformation beschrieben). • Lymphknoten-Bestrahlung, Knochenmarkstransplantation oder andere immunsupprimierende Behandlungen, deren Wirkung potenziell über sechs Monate anhält (unabhängig vom Zeitpunkt). 12. Vom Prüfarzt bestätigte medizinisch instabile Verfassung (unabhängig von der Art). 13. Eine(r) der folgenden Zustände oder Behandlungen, die Auswirkungen auf die Herzfunktion gehabt haben können:

Item[a]	Charakteristikum	Studieninformation
		• aktuelle oder anamnestisch bekannte signifikante Herzerkrankung, einschließlich Herzinsuffizienz (NYHA-Klassen II – IV), Myokarditis, Kardiomyopathie, Angina pectoris oder Myokardinfarkt (innerhalb von sechs Monaten), Schlaganfall (innerhalb von sechs Monaten), transitorische ischämische Attacke (TIA; innerhalb von sechs Monaten), dekompensierte Herzinsuffizienz, die einer Behandlung in einer Klinik bedarf (innerhalb von sechs Monaten) oder unkontrollierte arterielle Hypertonie. • Erregungsleitungsstörungen oder Rhythmusstörungen, einschließlich kompletter Linksschenkelblock, Sinusarrest oder sinuatrialer Block, symptomatische Bradykardie, Sick-Sinus-Syndrom, AV-Block zweiten Grades vom Typ Mobitz II oder höhergradiger AV-Block (entweder in der Anamnese oder beim Screening ermittelt), außer der Patienten hat einen funktionierenden Herzschrittmacher. • Herzrhythmusstörungen, die einer Behandlung bedürfen, oder eine Herzsynkope in der Anamnese. • Behandlung mit Antiarrythmika der Klassen Ia oder III. • Gesundheitszustand, der einer medikamentösen Behandlung bedarf, die ihrerseits einen AV-Block oder AV-Überleitungsstörungen verursacht haben könnte. Ausgenommen sind Beta-Blocker. • Patienten, die bei Randomisierung (zu Behandlungsbeginn) herzfrequenzsenkende Kalziumkanalblocker einnehmen oder andere Wirkstoffe, die die Herzfrequenz gesenkt haben können. • PR-Intervall >230 msek. Long-QT-Syndrom oder QTcF-Intervallverlängerung (> 450 msek bei Männern oder > 470 msek bei Frauen) im EKG zum Screeningzeitpunkt. • schwerwiegende Dysfunktion des vegetativen Nervensystems. • Herzerkrankung, die einer Kathederablation bedarf. • andere Herzerkrankungen oder Behandlungen, die signifikante Auswirkungen auf die Patientensicherheit haben können (bestimmt durch den Prüfarzt). 14. Eine der folgenden Lungenerkrankungen: • schwere Erkrankung der Atemwege, akut oder in der Anamnese, einschließlich chronischer obstruktiver Lungenerkrankung (COPD) oder Lungenfibrose. • Tuberkulose; ausgenommen sind eine erfolgreich behandelte Tuberkulose oder eine prophylaktische Behandlung nach positiver PPD-Hautreaktion in der Anamnese.

Item[a]	Charakteristikum	Studieninformation
		• schweres Asthma oder Asthma, das einer regelmäßigen Behandlung mit oralen Steroiden bedarf. 15. Eine der folgenden Lebererkrankungen, die vor der Randomisierung vorgelegen hat/haben: • Alkoholmissbrauch in der Anamnese, chronische Erkrankung der Leber oder der Galle. • gesamtes oder konjugiertes Bilirubin über dem 1,5-Fachen des oberen Normwertes (ULN), soweit dies nicht in Verbindung mit dem Gilbert-Meulengracht-Syndrom steht. • Alkalische Phosphatase (ALP) über dem 1,5-Fachen des oberen Normwerts (ULN). • Aspartat-Aminotransferase (AST), Alanin-Aminotransferase (ALT) oder Gamma-Glutamyltransferase (GGT) > 3-Fachen des oberen Normwertes (ULN). 16. Eine der folgenden Laborwertabweichungen vom Normbereich, die vor der Randomisierung vorgelegen haben: • Serumkreatinin > 1,7 mg/dl (150 µmol/l). • Leukozytenzahl (WBC) < 3.500/mm^3 (< 3,5 × 10^9/l). • Lymphozytenzahl < 800/mm^3 (< 0,8 × 10^9/l). • Serum-Kalium > ULN (oberer Normwert). • andere klinisch signifikante Laborbewertung (d. h. Hypomagnesiämie oder Hypokaliämie). 17. Eine der folgenden neurologischen/psychiatrischen Erkrankungen, die vor der Randomisierung vorgelegen haben: • Patienten mit der Angabe „ja“ bei Punkt 4 oder 5 des Abschnitts zu Selbstmordgedanken des eC-SSRS (electronic Columbia Suicide Severity Rating Scale). Sofern diese Gedanken in den letzten sechs Monaten aufgetreten sind oder es irgendeine „Ja“-Angabe im Abschnitt zu suizidalem Verhalten gab, außer „nichtsuizidales selbstverletzendes Verhalten“, sofern dieses Verhalten in den letzten zwei Jahren aufgetreten ist (falls dies zutraf, wurde der Patient zur weiteren Untersuchung und/oder Behandlung sofort an einen Psychologen/Psychiater überwiesen). • Missbrauch in der Anamnese (Alkohol oder Drogen) oder anderer Faktor (d. h. schwerwiegende psychiatrische Beschwerden), der die Fähigkeit des Patienten zur Kooperation oder die Compliance im Studienablauf beeinträchtigen könnte. • Progrediente neurologische Erkrankung (außer MS), die die Studienteilnahme gefährden könnte oder die die Einnahme von Medikamenten erfordern würde, die laut Studienprotokoll nicht erlaubt wären. 18. Patienten, die sich nicht einer MRT-Untersuchung unterziehen können. 19. Anwendung anderer Prüfpräparate bei Einschluss in die Studie oder im Zeitraum der vorangegangenen 30 Tage oder

Item[a]	Charakteristikum	Studieninformation
		innerhalb von fünf Eliminationshalbwertszeiten oder bis der erwartete pharmakodynamische Effekt auf den Ausgangswert zurückgegangen ist (je nachdem, was länger ist). 20. Bekannte Überempfindlichkeit gegenüber einem der Studienmedikamente oder gegenüber Medikamenten aus ähnlichen chemischen Klassen. 21. Homozygot auf CYP2C9*3 (Test beim Screening) oder Verweigerung eines Tests auf den Haplotyp CYP2C9*3. 22. Anwendung von Begleitmedikationen, die starke CYP2C9-Induktoren sind (einschließlich des Zeitraums von vier Wochen vor der Initialdosis). 23. Jegliche andere Erkrankung oder Gesundheitszustände, die entweder die Studienteilnahme gemäß Studienprotokoll oder die Fähigkeit der Patienten zur Kooperation oder die Compliance in den Studienabläufen beeinträchtigt haben könnten. **Metabolismus durch CYP2C9-Interferenz mit den Siponimod-Spiegeln** CYP2C9*3-homozygote Patienten (Test beim Screening) wurden aus der Studie ausgeschlossen. Nicht erlaubt war die Verwendung von Begleitmedikationen, die starke CYP2C9-Induktoren sind (einschließlich des Zeitraums von vier Wochen vor der Initialdosis). Starke CYP2C9-Inhibitoren konnten signifikante Auswirkungen auf die Siponimod-Exposition haben. Angezeigt ist somit eine erhöhte Wachsamkeit hinsichtlich möglicher unerwünschter Ereignisse. Starke CYP2C9-Induktoren durften nicht mit Siponimod verabreicht werden. Grund ist, dass keine Informationen hinsichtlich einer möglichen Reduktion der Wirkung von Siponimod im Falle einer zu geringen Exposition aufgrund einer CYP2C9-Induktion vorliegen. Für alle Patienten wurde der CYP2C9-Genotyp bei Studieneintritt wie folgt bestimmt: • WT/WT (Wildtyp CYP2C9*1/*1, extensiver Metabolisierer) • *2/WT (CYP2C9*1/*2, extensiver Metabolisierer) • *2/*2 (CYP2C9*2/*2, intermediärer Metabolisierer) • *3/WT (CYP2C91/*3, intermediärer Metabolisierer) • *2/*3 (CYP2C9*2/*3, schwacher Metabolisierer) (Gemäß Ausschlusskriterium 21 oben wurden schwache CYP2C9*3/*3-Metabolisierer aus der Zielpatientenpopulation ausgeschlossen).

Item[a]	Charakteristikum	Studieninformation
4b	Studienorganisation und Ort der Studiendurchführung	Die multizentrische Studie wurde an 292 Kliniken und spezialisierten MS-Zentren in 31 Ländern durchgeführt.
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	Gruppe 1: Siponimod 2 mg Gruppe 2: Placebo In der Kernphase begannen die Patienten die Behandlung mit Auftitration von Siponimod oder entsprechendem Placebo: Tag 1: 0,25 mg; Tag 2: 0,25 mg; Tag 3: 0,5 mg; Tag 4: 0,75 mg; Tag 5: 1,25 mg; Tag 6: 2,0 mg. Bei Patienten mit bestätigter Lymphozytenzahl von $< 0{,}2 \times 10^9$/l wurde die Dosis auf 1 mg/Tag reduziert.
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	**Primäres Ziel** Primär sollte die Wirksamkeit von Siponimod bei Patienten mit SPMS im Vergleich mit Placebo gezeigt werden, und zwar hinsichtlich der Verzögerung der Zeit bis zur nach drei Monaten bestätigten Behinderungsprogression, gemessen anhand des EDSS. **Wesentliche sekundäre Ziele** • Wirksamkeit von Siponimod im Vergleich mit Placebo hinsichtlich der Verzögerung der Zeit bis zu einer bestätigten Verschlechterung nach drei Monaten, und zwar mindestens um 20 % im T25W seit Studienbeginn. • Wirksamkeit von Siponimod im Vergleich mit Placebo in der Verringerung des Anstiegs des T2-Läsionsvolumens seit Studienbeginn. **Zusätzliche sekundäre Ziele** • Untersuchung der Wirksamkeit von Siponimod im Vergleich mit Placebo hinsichtlich der Verzögerung der Zeit bis zur nach sechs Monaten bestätigten Behinderungsprogression, gemessen anhand des EDSS. • Untersuchung der Wirksamkeit von Siponimod im Vergleich mit Placebo hinsichtlich einer Verringerung der Häufigkeit von bestätigten Schüben, gemessen anhand der jährlichen Schubrate (ARR, Annualized Relapse Rate). Weiterhin die Bewertung der Zeit bis zum ersten Schub und der Anteil schubfreier Patienten. • Untersuchung der Wirkung von Siponimod im Vergleich mit Placebo hinsichtlich des von Patienten berichteten MSWS-12. • Untersuchung der Wirksamkeit von Siponimod im Vergleich mit Placebo in Hinblick auf die inflammatorische Aktivität und die Krankheitslast, ermittelt durch herkömmliches MRT: T1-

		Gadolinium-(Gd)-anreichernde Läsionen, neue oder sich vergrößernde T2-Läsionen, Hirnvolumen. • Untersuchung der Wirksamkeit von Siponimod im Vergleich mit Placebo hinsichtlich der nach drei Monaten bestätigten Behinderungsprogression, gemessen anhand des EDSS in folgenden Subgruppen im Vergleich: o SPMS-Patienten mit bzw. ohne aufgesetzte(n) Schübe(n). o Patienten mit schnell fortschreitender Erkrankung, definiert durch eine Veränderung von 1,5 Punkten im EDSS in den zwei Jahren vor Studienstart, bzw. bei solchen Patienten, die dieses Kriterium nicht erfüllten. o Patienten mit moderatem und schwerem Erkrankungsverlauf, definiert durch einen MSSS-Wert (Multiple Sclerosis Severity Score) von 4 oder höher zu Studienbeginn, bzw. bei Patienten, die nicht dieses Kriterium erfüllten. • Ermittlung der Sicherheit und Verträglichkeit von Siponimod im Vergleich mit Placebo. **Explorative Ziele** • Untersuchung der Wirkung von Siponimod im Vergleich mit Placebo in folgenden von Patienten berichteten Endpunkten: o Gesundheitsbezogene Lebensqualität (QoL), gemessen anhand des MSIS-29 o Gesundheitsbezogene Lebensqualität (QoL), gemessen anhand des EQ-5D • Untersuchung der Wirksamkeit von Siponimod im Vergleich mit Placebo in definierten Kognitionstests: o PASAT o SDMT o BVMT-R • Untersuchung der Wirksamkeit von Siponimod im Vergleich mit Placebo hinsichtlich der Entwicklung akuter Läsionen hin zu chronischen Black Holes, ermittelt im MRT. • Untersuchung der Wirksamkeit von Siponimod im Vergleich mit Placebo hinsichtlich des MSFC-z-Scores. • Untersuchung der Wirksamkeit von Siponimod im Vergleich mit Placebo hinsichtlich der Verzögerung der Zeit: o bis zur nach drei Monaten bestätigten Verschlechterung um mindestens 20 % im T25W oder

Item[a]	Charakteristikum	Studieninformation
		○ bis zur nach drei Monaten bestätigten Behinderungsprogression, gemessen anhand des EDSS-Scores. ○ bis zur nach drei Monaten bestätigten Verschlechterung um mindestens 20 % im 9-HPT an mindestens einer Hand (dominant oder nicht-dominant). • Untersuchung der Charakteristika zu Studienbeginn, die mit einem positiven Behandlungsansprechen assoziiert sind. Damit sollen die Subgruppen der klinisch relevanten Responder definiert werden. • Untersuchung des Zusammenhangs zwischen den Endpunkten zur Behinderungsprogression und der Wirkstoffkonzentration/Lymphozytenzahl. • Untersuchung des Zusammenhangs zwischen ausgewählten Sicherheitsparametern und der Wirkstoffkonzentration/Lymphozytenzahl. • Untersuchung der pharmakokinetischen Eigenschaften (PK) von Siponimod. • Untersuchung der Wirkungen von Siponimod im Vergleich mit Placebo auf die nach drei Monaten bestätigte Behinderungsprogression, gemessen anhand des EDSS, in folgenden Subgruppen im Vergleich: ○ Patienten, die vorher mit Interferon-beta 1b behandelt wurden bzw. nicht damit behandelt wurden. ○ Unvorbehandelte Patienten bzw. Patienten, die vorab mit krankheitsmodifizierenden Medikamenten behandelt worden sind.
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Keine; Änderungen zur EDSS-Erfassung wurden in Abschnitt 3b beschrieben.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	Die Studie wurde mit einer Power von 90 % angelegt, um eine 30-prozentige Reduktion des Risikos einer nach drei Monaten bestätigten Behinderungsprogression zu bestimmen (Hazard Ratio: 0,70). Zur Anwendung kam dabei ein Log-Rank-Test mit einem zweiseitigen Alpha-Level von 5 %. Die Randomisierung erfolgte im Verhältnis von 2:1 Siponimod zu Placebo. Unter Annahme, dass der Anteil der Patienten mit einer Behinderungsprogression nach zwei Jahren in der Placebogruppe bei 0,30 liegt, einer Zwei-Jahres-Dropout-Rate von 20 % und einer Einschlussrate von 100 Patienten pro Monat, waren 1.530 Patienten und eine Gesamtstudiendauer von 42 Monaten notwendig, um mindestens 374 Patienten mit Behinderungsprogression zu erfassen. Dies sollte die notwendige Power gewährleisten. In dieser Berechnung wurden für das Ereignis und die Drop-Out-Raten exponentielle Verteilungsannahmen getroffen. Geändert wurde das Protokoll, um das Kriterium zum Abbruch der Kernphase zu aktualisieren. Ursprünglich wäre

Item[a]	Charakteristikum	Studieninformation
		gestoppt worden, sobald 374 Patienten eine Behinderungsprogression gehabt hätten. Änderung des Stopp-Kriteriums auf ungefähr drei Jahre nach Randomisierung des ersten Patienten und einer Ereigniszahl von mindestens 374. Es wurde angenommen, dass nach ungefähr drei Jahren mehr als 374 Ereignisse beobachtet worden waren (siehe Zusammenfassung des Amendments 3). Erwartungsgemäß sollte dies den leichten Power-Verlust aufgrund der Alpha-Adjustierung für die Interimsanalyse kompensieren und eine Power zum Ende der Kernphase von mindestens 90 % erreicht sein.
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	**Verblindeter Review der Fallzahl** Vor der unverblindeten Futility-Interimsanalyse wurde ein verblindeter Review der Fallzahl vorgenommen. Vor Abschluss der Rekrutierung wurde ein Review der gepoolten Daten zur Behinderungsprogression (basierend auf dem EDSS) vorgenommen. Damit sollten die Annahmen zur Berechnung der Fallzahl auf Basis der Behinderungsprogression neu bewertet werden. Nachdem die verblindeten Ergebnisse vorlagen, entschied man sich, wie im Protokoll geplant fortzufahren. **Interimsanalyse** Die in diesem Abschnitt erwähnte Interimsanalyse wurde für das DMC (Data Monitoring Committee) durchgeführt. Aus diesem Grund wurden die Analysen in zwei separaten Analyseplänen beschrieben, Studienberichte zu diesen Analysen wurden nicht erstellt. Vorgenommen wurden die Analysen von einer externen CRO, die nicht anderweitig in die Studie involviert war. Die regelmäßigen halbverblindeten Interimsanalysen zu den Sicherheitsdaten wurden dem DMC (ungefähr zweimal im Jahr) vorgelegt. Ebenfalls vorgelegt wurde eine Interimsanalyse zur Futility-Bewertung. Ursprünglich war im Protokoll zum Zeitpunkt der Interimsanalyse eine Bewertung erlaubt, nach der die Studie bei gezeigter Wirksamkeit frühzeitig beendet werden konnte. Zur Anwendung kam hier eine O'Brien-Fleming-Grenze mit einer Lan-DeMets Alpha Spending function. Vor der Durchführung der Interimsanalyse zur Futility wurde jedoch entschieden, dass die Studie auch bei Erreichen der Abbruchgrenze fortgesetzt werden sollte. Damit sollte die Sammlung ausreichender Langzeitdaten zur Sicherheit und Wirksamkeit ermöglicht werden. Trotzdem wurde in der finalen Analyse der kritische Wert unter Einbezug des durch die Interimsanalyse verlorenen Alpha-Levels sowie der tatsächlichen Korrelation der Teststatistiken berechnet. Damit sollte ein kumulativer Typ-I-Fehler erreicht werden, der geringer als das Signifikanzniveau war (das heißt: kleiner als 5 % für einen zweiseitigen Test). Die Interimsanalyse zur Futility sollte vorgenommen werden, wenn mindestens 50 % der benötigten Anzahl an Patienten mit einer nach drei Monaten bestätigten

Item[a]	Charakteristikum	Studieninformation
		Behinderungsprogression in der Datenbank vorlagen. Während der Kernphase war es Novartis nicht möglich, die unverblindeten Ergebnisse einzusehen. Das DMC empfahl, dass die Studie fortgesetzt werden sollte. Das heißt, sie sollte nicht aufgrund von Futility abgebrochen werden.
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Geeignete Patienten wurden mittels IRT (Interactive Response Technology) randomisiert.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Patienten wurden auf die Gruppen im Zufallsprinzip nach dem Verhältnis 2:1 verteilt, und zwar mittels Blockrandomisierung mit einer Blockgröße von 6. Die Randomisierung wurde nach Regionen stratifiziert.
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	An Tag 1 (Visite zur Randomisierung) wurden geeignete Patienten mittels IRT (Interactive Response Technology) auf Siponimod oder Placebo randomisiert. Der Prüfarzt oder sein Stellvertreter nahm mit dem IRT Kontakt auf, nachdem bestätigt war, dass der Patient die Einschlusskriterien erfüllte und unter keines der Ausschlusskriterien fiel. Das IRT teilte dem Patienten eine Randomisierungsnummer zu. Diese ordnete den Patienten einem Behandlungsarm zu und wurde mit einer eindeutigen Medikationsnummer für die erste Packung der Studienbehandlung verknüpft, die an den Patienten ausgegeben wurde. Die Randomisierungsnummer wurde dem telefonischen Übermittler nicht mitgeteilt. Die Randomisierungsdaten wurden bis zum Zeitpunkt der Entblindung streng vertraulich behandelt. Mit Ausnahmen (Mitglieder des DMC und Pharmakokinetiker) waren die Daten niemandem zugänglich, der an der Studie beteiligt war.
10	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?	Um sicherzustellen, dass die Behandlungszuteilung unverzerrt war und für die Patienten und das Prüfpersonal verblindet war, wurden die Randomisierungsnummern nach folgendem Verfahren generiert. Der Anbieter des IRT fertigte eine Randomisierungsliste an. Dabei kam ein validiertes System zur Anwendung, das die Zufallszuteilung der Patientennummern zu den Randomisierungsnummern automatisiert vornahm. Diese Randomisierungsnummern wurden mit den verschiedenen Behandlungsarmen und wiederum mit den Medikationsnummern verknüpft. Eine separate Medikationsliste wurde vom Novartis Drug Supply Management bzw. unter dessen Verantwortung erstellt. Dabei kam ein validiertes System zum Einsatz, das die zufällige Zuteilung der Medikationsnummern zu den Packungen automatisiert vornahm, die das zu untersuchende Medikament enthielten. Siehe auch Abschnitt 9.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c)	Patienten, Prüfpersonal, Personen, die die Bewertungen vornahmen, sowie die Datenanalysten blieben hinsichtlich der Identität der Therapie verblindet, und zwar bis zur Schließung

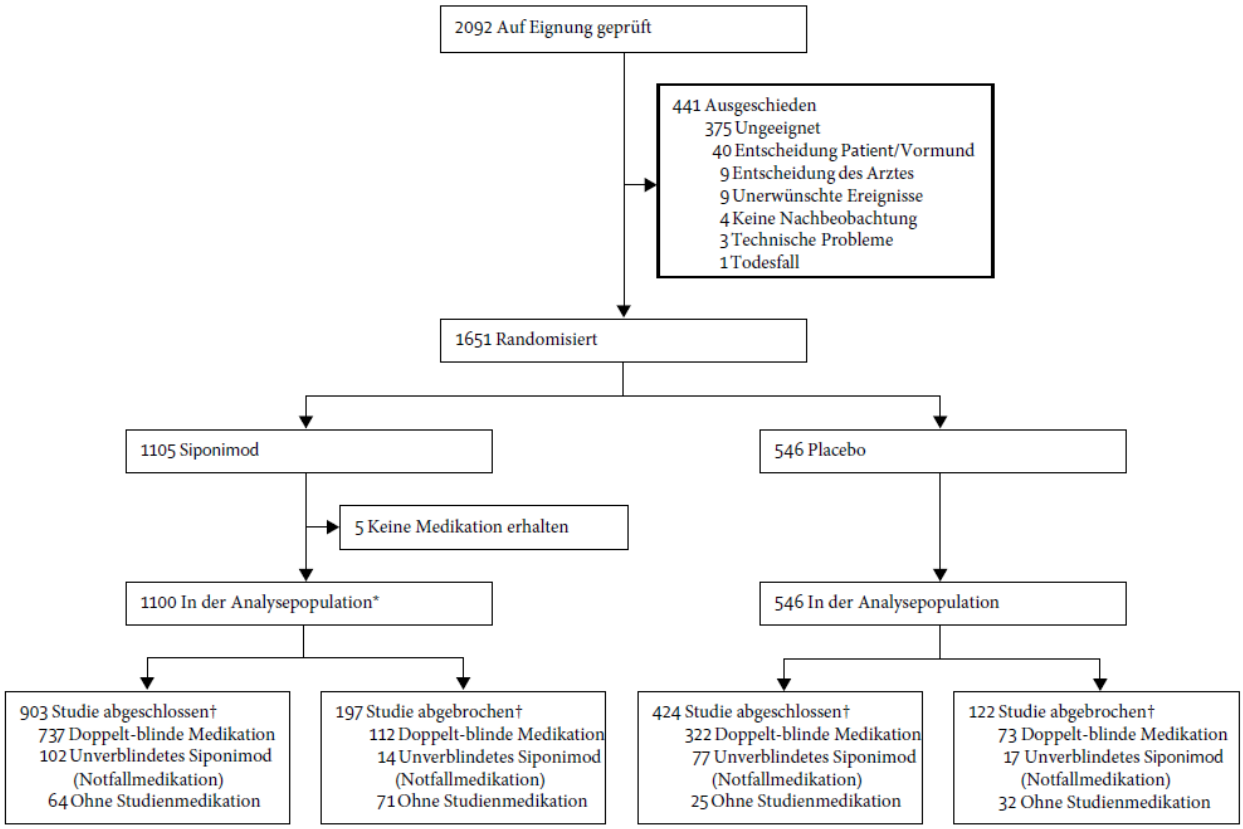
Item[a]	Charakteristikum	Studieninformation
	diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	der Datenbank in der Kernphase. Dabei kamen folgende Methoden zur Anwendung: 1. Bis zur vorgesehenen Entblindung blieben die Randomisierungsdaten streng geheim. Sie waren für niemanden, der an der Studie beteiligt war, einsehbar. Allerdings mit folgenden Ausnahmen: • Mitglieder des DMC, unabhängige Statistiker und unabhängige Programmierer. • Pharmakokinetiker konnten auf die Randomisierungscodes derjenigen Patienten zugreifen, für die PK-Stichproben genommen wurden. Die PK-Ergebnisse wurden bis zur Schließung der Datenbank geheim gehalten. 2. Die Identität der Behandlungen war verdeckt Eine Entblindung war gestattet, falls beim Patienten Notfallsituationen auftraten sowie zum Abschluss der Kernphase der Studie. Die Erhebung des EDSS wurde von einem unabhängigen EDSS-Bewerter vorgenommen. Um die Verblindung aufrecht zu erhalten, wurden zwei verschiedene Datenbanken für die Hauptdaten sowie die Daten zum Behandlungsbeginn aufgesetzt. Die Prüfzentren wurden über das iPortal über SUSARs (Suspected Unexpected Serious Reactions) in Kenntnis gesetzt. Damit hatte das Personal des Prüfzentrums Zugriff auf die veröffentlichten SUSARs zu ihren jeweiligen Novartis-Studien. Im Verlauf der Studie wurde bemerkt, dass das iPortal-System versehentlich unverblindete SUSARs an zwei Zentren übermittelt hatte. Dadurch war es dort potenziell möglich, die Behandlungen ihrer Patienten zu sehen (PIDs A2304-4033-002, A2304-7081-020 und A2304-7081-005). Diese potenziellen Entblindungen wurden derart bewertet, als dass sie keine Auswirkungen auf die Gesamtintegrität der Studie haben. Grund war, dass alle drei Patienten die doppelt-verblindete Therapie abgeschlossen hatten, bevor es zu dieser versehentlichen Entblindung kam.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Die Identität der Behandlungen wurde verblindet, indem Studienmedikamente verwendet wurden, bei denen folgende Faktoren identisch waren: Verpackung, Etikett, Verabreichungsplan, Aussehen, Geschmack und Geruch.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	Mit einer Randomisierung von Siponimod zu Placebo im Verhältnis 2:1 und einer Beobachtung von mindestens 374 Fällen einer nach drei Monaten bestätigten Behinderungsprogression hatte die Studie eine Power von 90 %, um damit eine 30-prozentige Verringerung des Risikos einer nach drei Monaten bestätigten Behinderungsprogression zu ermitteln. Dabei kam ein Log-Rank-Test mit einem zweiseitigen Signifikanzniveau von 5 % zum Einsatz. Die Unterschiede zwischen Siponimod und Placebo hinsichtlich der Zeit bis zur nach drei Monaten bestätigten Behinderungsprogression und hinsichtlich der Zeit bis zur

		nach drei Monaten bestätigten Verschlechterung um mindestens 20 % im T25W wurden mittels Cox-Modell (Proportional hazard model) und Log-Rank-Test analysiert. Bei dem Cox-Modell (Proportional hazard model) wurden diejenigen Patienten von den Analysen ausgeschlossen, bei denen Angaben zu Kovariaten fehlten. Die Annahmen des Cox-Modells (Proportional hazard model) wurden getestet. Für den primären Endpunkt waren Behandlung, Land, EDSS zu Studienbeginn und SPMS-Gruppe (mit oder ohne aufgesetzte Schübe gemäß Definition zu Studienbeginn) Kovariaten im Cox-Modell. Für den sekundären Endpunkt, der Zeit bis zur nach drei Monaten bestätigten Verschlechterung (mindestens 20 % im T25W), waren Behandlung, Land, EDSS zu Studienbeginn, T25W zu Studienbeginn und die SPMS-Gruppe (mit oder ohne aufgesetzte Schübe gemäß Definition zu Studienbeginn) Kovariaten im Cox-Model. Die Risikoreduktion wurde wie folgt abgeleitet: (1 – Hazard Ratio) × 100. Die Analyse des T2-Läsionsvolumens und der prozentualen Veränderung des Hirnvolumens erfolgte mittels eines Mixed Model for Repeated Measures, die Visite war dabei der kategorielle Faktor. Zur Erfassung des T2-Läsionsvolumens wurde das Modell für Folgendes adjustiert: Behandlung, Land, Alter, T2-Läsionsvolumen bei Studienbeginn, Anzahl Gd-anreichernder T1-Läsionen zu Studienbeginn sowie SPMS-Gruppe (mit oder ohne aufgesetzte Schübe gemäß Definition zu Studienbeginn). Der adjustierte Mittelwert ist definiert als Veränderung des T2-Läsionsvolumens seit Studienbeginn. Zur Erfassung der prozentualen Veränderung im Hirnvolumen wurde das Modell auf folgende Faktoren adjustiert: Behandlung, Land, Alter, normalisiertes Hirnvolumen zu Studienbeginn, Anzahl Gd-anreichernder T1-Läsionen zu Studienbeginn, T2-Läsionsvolumen zu Studienbeginn sowie SPMS-Gruppe (mit oder ohne Schübe gemäß Definition zu Studienbeginn). Der adjustierte Mittelwert bezieht sich auf die prozentuale Veränderung des Hirnvolumens seit Studienbeginn. Die Anzahl der Läsionen und die ARR wurden durch negative Binomial-Regression geschätzt. Der primäre Endpunkt wurde bei einem adjustierten Alpha-Level von 0,0434 basierend auf einer O'Brien-Fleming-Korrektur getestet. Die wesentlichen sekundären Endpunkte wurden hierarchisch bei einem zweiseitigen Signifikanzniveau von 0,05 getestet. Die zusätzlichen sekundären Endpunkte wurden analysiert mit einem nominalen Signifikanzniveau von 0,05 ohne Korrektur der Multiplizität oder hierarchischen Test. Primäre und sekundäre Endpunkte wurden im FAS analysiert, das aus allen randomisierten und behandelten Patienten bestand. Gemäß dem Intention-to-treat-Prinzip wurden alle Daten verwendet, unabhängig von einer frühzeitigen Beendigung der Einnahme des verblindeten Studienmedikaments. Die Sicherheit wurde im Safety Set ermittelt. Dieses umfasste alle Patienten, die mindestens eine Dosis des Studienmedikaments genommen hatten. Analysiert wurden

Item[a]	Charakteristikum	Studieninformation
		die Patienten nach der tatsächlich erhaltenen Behandlung. Unerwünschte Ereignisse beim doppelt-verblindeten Studienmedikament bestanden definitionsgemäß 30 Tage nach Absetzen des Medikaments fort. Schwerwiegende Unerwünschte Ereignisse wurden für die Kernphase der Studie einschließlich der Ereignisse unter unverblindetem Siponimod in der Kernphase berichtet.
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	Folgende Patienten-Subgruppen wurden definiert: • Demographische Daten zu Studienbeginn und Behandlungshistorie (Geschlecht, vorherige Behandlung mit IFN-beta 1b, vorherige Behandlung mit einer krankheitsmodifizierenden MS-Therapie, [Vorherige Behandlung mit IFN-beta wurde als Post-hoc-Analyse ergänzt]). • Krankheitscharakteristika zu Studienbeginn: o SPMS-Patienten mit oder ohne aufgesetzte Schübe im Zeitraum von zwei Jahren vor der Screening-Visite. Dies wird als „Baseline-Definition“ für mit bzw. ohne Schübe im Studienbericht erfasst. o Patienten mit schnell fortschreitender Erkrankung (definiert anhand anamnestischer EDSS-Scores; das heißt, eine Veränderung von ≥ 1,5 in den zwei Jahren vor der Studie oder zu Studienbeginn). Alle Patienten, deren Erkrankung im Adjudizierungsprozess als progredient eingestuft worden war, wurden nicht in die Subgruppe der Patienten mit schnell fortschreitender Erkrankung aufgenommen. o Erkrankungsverlauf: Patienten mit MSSS ≥ 4 wurden in der Subgruppe mittelschwer/schwer eingeteilt. o Anzahl Gd-anreichernder T1-Läsionen zu Studienbeginn (0; ≥ 1). • Patienten mit oder ohne mindestens einem bestätigten Schub zu irgendeinem Zeitpunkt an Tag 1 oder nach Tag 1.
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben,	Siehe Flussdiagramm im Anschluss an die Tabelle.

Item[a]	Charakteristikum	Studieninformation
	c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	Siehe Flussdiagramm im Anschluss an die Tabelle.
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Vom 5. Februar 2013 bis 2. Juni 2015 wurden 1.651 Patienten randomisiert.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Kernphase der Studie wurde planmäßig beendet. Die Nachbeobachtung der Patienten in der Extensionsphase läuft noch.
a: nach CONSORT 2010.		

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



*Ein Patient (randomisiert auf Siponimod) wurde wegen fehlender Einwilligungserklärung zu Studieneinschluss nicht in die Auswertungen einbezogen. †Doppelt-verblindete, placebo-kontrollierte Kernphase.

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 4-188 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie EXPAND

Studie: EXPAND (CBAF312A2304)

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Studienbericht	CSR
Publikation	Kappos 2018

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:
Bei der EXPAND-Studie handelt es sich um eine randomisierte, kontrollierte Studie mit adäquat erzeugter Randomisierungssequenz. Die Gruppenzuteilung erfolgte verdeckt und zentralisiert durch ein Interactive Response System (IRS).

1.
<u>für randomisierte Studien:</u> Adäquate Erzeugung der Randomisierungssequenz

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

<u>für nicht randomisierte Studien:</u> Zeitliche Parallelität der Gruppen

☐ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

2.
<u>für randomisierte Studien:</u> Verdeckung der Gruppenzuteilung („allocation concealment“)

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; falls unklar oder nein, <u>obligate</u> Begründung für die Einstufung:

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:

Alle Patienten waren verblindet und bekamen entweder einmal täglich Siponimod oder das entsprechende Placebo, jeweils oral eingenommen. Verum und Placebo waren in Aussehen, Geschmack, Geruch und Verpackung identisch.

behandelnde bzw. weiterbehandelnde Personen:

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Alle in die Studie involvierten Personen (Prüfarzt, Personal am Zentrum, untersuchender Arzt, mit der Auswertung betraute Personen) waren zwischen Randomisierung und Datenbankschluss verblindet (mit Ausnahme des unabhängigen Statistikers und des unabhängigen Programmierers sowie der Mitglieder des Data Monitoring Committee, DMC).

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

☐ **ja** ☐ **unklar** ☒ **nein**

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

Im CSR dargestellte Subgruppenanalysen wurden nicht a-priori im Prüfplan definiert.

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

☒ **ja** ☐ **nein**

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

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

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Da die Randomisierungssequenz adäquat erzeugt wurde, die Gruppenzuteilung verdeckt erfolgte, Patienten und Behandler verblindet waren und keine sonstigen das Verzerrungspotenzial beeinflussenden verzerrenden Punkte vorliegen, wird das Verzerrungspotenzial auf Studienebene insgesamt als niedrig bewertet. Die post-hoc durchgeführten Subgruppenanalysen im CSR führen hierbei zu keiner relevant erhöhten Verzerrung.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Endpunkt: Mortalität

1. **Verblindung der Endpunkterheber**

   ☒ **ja** ☐ **unklar** ☐ **nein**

   Angaben zum Kriterium; obligate Begründung für die Einstufung:

   Der Endpunkterheber (Prüfarzt) war 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** ☐ **nein**

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

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

   ☒ **ja** ☐ **nein**

   Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Mortalität* als niedrig bewertet.

Endpunkt: Bestätigte Krankheitsschübe

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Patienten waren angehalten, unverzüglich ihren behandelnden Arzt zu kontaktieren, wenn neurologische Symptome auftraten bzw. sich verschlechterten. Zusätzlich wurde jeder Patient zu jeder planmäßigen Visite nach aufgetretenen Symptomen befragt. Waren diese aufgetretenen Symptome schubbedingt, erfolgte eine neurologische Untersuchung durch den Prüfarzt sowie eine EDSS-Bewertung durch einen unabhängigen EDSS-Bewerter, der nicht in die Behandlung des Patienten involviert war. Sowohl die Patienten als auch der behandelnde und der unabhängig beurteilende Bewerter waren bezüglich der Behandlung 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** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine

ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Bestätigte Krankheitsschübe* als niedrig bewertet.

Endpunkt: Behinderungsprogression (EDSS)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung

Die Behinderungsprogression wurde anhand des EDSS-Wertes eingeschätzt Die Einschätzung des EDSS-Werts wurde vom unabhängigen EDSS-Bewerter, der nicht in die Behandlung des Patienten involviert war, vorgenommen. Der EDSS-Bewerter war ebenfalls 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** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Behinderungsprogression (EDSS)* als niedrig bewertet.

Endpunkt: Behinderungsprogression (MSFC)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die MSFC-Bewertung wurde entweder vom unabhängigen EDSS-Bewerter oder von qualifiziertem Personal, das nicht in die Behandlung des Patienten involviert war, durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☐ **ja** ☐ **unklar** ☒ **nein**

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

Der Nichtberücksichtigungsanteil liegt zu Monat 12 insgesamt bei unter 30 % mit Ausnahme der Auswertung des MSFC-z in der BSC-Gruppe der Subpopulation B (32,8 %). Der Unterschied in den Nichtberücksichtigungsanteilen zwischen den Behandlungsgruppen liegt unter 10 % mit Ausnahme der Auswertung des MSFC-z und des T25W in der Subpopulation B (13,9 % bzw. 13,8 %). Zu Monat 24 liegen die Nichtberücksichtigungsanteile bei über 30 %. Die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 liegen in beiden Populationen bei über 30 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Die Endpunkterheber waren verblindet und es gab weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte. Da das ITT-Prinzip nicht in allen Auswertungen adäquat

umgesetzt wurde, gilt folgende Bewertung des Verzerrungspotenzials: Für die Auswertung zu Monat 12 ist für den MSFC-z und den T25W in der Subpopulation B mit einem hohen Verzerrungspotenzial zu rechnen. Für alle Auswertungen in der Gesamtpopulation und für den PASAT-3 sowie den 9-HPT in der Subpopulation B ist das Verzerrungspotenzial niedrig. Für die Auswertung zu Monat 24 lagen die Nichtberücksichtigungsanteile bei über 30 %, hier ist mit einem hohen Verzerrungspotenzial zu rechnen.

Endpunkt: Schweregrad der kognitiven Dysfunktion (SDMT)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die SDMT-Bewertung wurde entweder vom unabhängigen EDSS-Bewerter oder von qualifiziertem Personal, das nicht in die Behandlung des Patienten involviert war, durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

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

Der Anteil der Patienten ohne Berücksichtigung in der Auswertung liegt für die Gesamtpopulation zu Monat 12 unter 30 % und ist zwischen den Behandlungsgruppen nicht relevant verschieden (Unterschied in den Nichtberücksichtigungsanteilen < 10 %). Dies gilt im gleichen Maß für die im Dossier dargestellte Subpopulation B. Jedoch liegen die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 in beiden Populationen bei über 30 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Schweregrad der kognitiven Dysfunktion*

(SDMT) für die Auswertung zu Monat 12 als niedrig bewertet. Für die Auswertung zu Monat 24 lagen die Nichtberücksichtigungsanteile bei über 70 %, hier ist mit einem hohen Verzerrungspotenzial zu rechnen.

Endpunkt: Schweregrad der kognitiven Dysfunktion (BVMT-R)

1. **Verblindung der Endpunkterheber**

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die BVMT-R-Bewertung wurde entweder vom unabhängigen EDSS-Bewerter oder von qualifiziertem Personal, das nicht in die Behandlung des Patienten involviert war, vorgenommen.

2. **Adäquate Umsetzung des ITT-Prinzips**

☒ **ja** ☐ **unklar** ☐ **nein**

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

Der Anteil der Patienten ohne Berücksichtigung in der Auswertung liegt für die Gesamtpopulation zu Monat 12 unter 30 % und ist zwischen den Behandlungsgruppen nicht relevant verschieden (Unterschied in den Nichtberücksichtigungsanteilen < 10 %). Dies gilt im gleichen Maß für die im Dossier dargestellte Subpopulation B. Jedoch liegen die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 in beiden Populationen bei über 30 %.

3. **Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine**

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Schweregrad der kognitiven Dysfunktion (BVMT-R)* für die Auswertung zu Monat 12 als niedrig bewertet. Für die Auswertung zu Monat 24 lagen

die Nichtberücksichtigungsanteile bei über 30 %, hier ist mit einem hohen Verzerrungspotenzial zu rechnen.

Endpunkt: Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die LCVA-Bewertung wurde entweder vom unabhängigen EDSS-Bewerter oder von qualifiziertem Personal, das nicht in die Behandlung des Patienten involviert war, durchgeführt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

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

Der Anteil der Patienten ohne Berücksichtigung in der Auswertung liegt für die Gesamtpopulation zu Monat 12 unter 30 % und ist zwischen den Behandlungsgruppen nicht relevant verschieden (Unterschied in den Nichtberücksichtigungsanteilen < 10 %). Dies gilt im gleichen Maß für die im Dossier dargestellte Subpopulation B. Jedoch liegen die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 in beiden Populationen bei über 30 %

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Schweregrad der Beeinträchtigung des Sehvermögens (LCVA)* für die Auswertung zu Monat 12 als niedrig bewertet. Für die Auswertung zu Monat

24 lagen die Nichtberücksichtigungsanteile bei über 30 %, hier ist mit einem hohen Verzerrungspotenzial zu rechnen.

Endpunkt: Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die MSWS-12 erfasst die Gehfähigkeit über einen Fragebogen mit zwölf Fragen. Sie werden vom verblindeten Patienten selbst beantwortet.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

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

Der Anteil der Patienten ohne Berücksichtigung in der Auswertung liegt für die Gesamtpopulation zu Monat 12 unter 30 % und ist zwischen den Behandlungsgruppen nicht relevant verschieden (Unterschied in den Nichtberücksichtigungsanteilen < 10 %). Dies gilt im gleichen Maß für die im Dossier dargestellte Subpopulation B. Jedoch liegen die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 in beiden Populationen bei über 30 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Schweregrad der Beeinträchtigung der Gehfähigkeit (MSWS-12)* für die Auswertung zu Monat 12 als niedrig bewertet. Für die Auswertung zu

Monat 24 lagen die Nichtberücksichtigungsanteile bei über 30 %, hier ist mit einem hohen Verzerrungspotenzial zu rechnen.

v
Endpunkt: Gesundheitsbezogene Lebensqualität (MSIS-29)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der MSIS-29 besteht aus zwei Komponenten, der Physical Impact Scale und der Psychological Impact Scale. Beide werden vom verblindeten Patienten selbst eingeschätzt.

2. Adäquate Umsetzung des ITT-Prinzips

☐ **ja** ☐ **unklar** ☒ **nein**

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

Der Anteil der Patienten ohne Berücksichtigung in der Auswertung liegt für die Gesamtpopulation zu Monat 12 unter 30 % und ist zwischen den Behandlungsgruppen nicht relevant verschieden (Unterschied in den Nichtberücksichtigungsanteilen < 10 %). Der Anteil der Patienten ohne Berücksichtigung in der Auswertung der Subpopulation B zu Monat 12 liegt unter 30 %, wobei der Unterschied in den Nichtberücksichtigungsanteilen bei über 10 % liegt. Die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 liegen in beiden Populationen bei über 30 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als hoch bewertet. Die Endpunkterheber waren verblindet und es gab weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte. Da das ITT-Prinzip in der Subpopulation B nicht adäquat

umgesetzt wurde, wird das Verzerrungspotenzial des Endpunkts *Gesundheitsbezogene Lebensqualität (MSIS-29)* als hoch bewertet.

Endpunkt: Gesundheitsbezogene Lebensqualität (EQ-5D)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Der EQ-5D besteht aus zwei Komponenten, der Health State Classification und der Visuellen Analogskala. Beide werden vom verblindeten Patienten selbst eingeschätzt.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

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

Der Anteil der Patienten ohne Berücksichtigung in der Auswertung liegt für die Gesamtpopulation zu Monat 12 unter 30 % und ist zwischen den Behandlungsgruppen nicht relevant verschieden (Unterschied in den Nichtberücksichtigungsanteilen < 10 %). Dies gilt im gleichen Maß für die im Dossier dargestellte Subpopulation B. Jedoch liegen die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 in beiden Populationen bei über 30 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Gesundheitsbezogene Lebensqualität (EQ-*

5D) für die Auswertung zu Monat 12 als niedrig bewertet. Für die Auswertung zu Monat 24 lagen die Nichtberücksichtigungsanteile bei über 30 %, hier ist mit einem hohen Verzerrungspotenzial zu rechnen.

Endpunkt: Unerwünschte Ereignisse

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Das Auftreten von Unerwünschten Ereignissen (UE) wurde bei jeder Visite durch nicht-direktive Befragung der Patienten erfasst. Zudem konnten UE auch jederzeit durch den Patienten selbst spontan berichtet oder anhand von Untersuchungen wie Labormessungen nach definierten Kriterien festgestellt werden. Die Patienten und Behandler waren 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** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

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

Die Beobachtungsdauer war in den Studienarmen nicht unterschiedlich.

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

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt *Unerwünschte Ereignisse* wird als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkterheber (Prüfarzt) verblindet war, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde und eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war. Die Kriterien für Unerwünschte Ereignisse waren im Prüfplan a priori definiert. Die Beobachtungsdauer war in den Studienarmen nicht unterschiedlich. Insgesamt kann von einem objektivierbaren Endpunkt

ausgegangen werden und es ist nur mit einem niedrigen daraus resultierenden Verzerrungspotenzial zu rechnen.

Endpunkt: Suizidalität (C-SSRS)

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die C-SSRS wurde über ein interaktives Voice-Response-System erhoben. Die Fragen wurden direkt durch die Patienten beantwortet. Die Patienten waren 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** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

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

Die Beobachtungsdauer war in den Studienarmen nicht unterschiedlich.

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

☒ **niedrig** ☐ **hoch**

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt *Suizidalität* wird als niedrig bewertet, da es auf Studienebene niedrig war, der Endpunkt durch ein Voice-Response-System erhoben wurde, das ITT-Prinzip bei der Analyse adäquat umgesetzt wurde und eine ergebnisgesteuerte Berichterstattung unwahrscheinlich war. Insgesamt kann von einem objektivierbaren Endpunkt ausgegangen werden und es ist nur mit einem niedrigen daraus resultierenden Verzerrungspotenzial zu rechnen.

Endpunkt: Läsionslast

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die MRT-Scans wurden zentral von einem verblindeten MRT-Auswerter auf Qualität, Vollständigkeit und Einhaltung der MRT-Vorgaben geprüft und anschließend ausgewertet. Eine unvollständige oder inkorrekte MRT-Aufnahme wurde wiederholt. Die Auswertung der Scans erfolgte nach einer standardisierten Prozedur.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

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

Der Anteil der Patienten ohne Berücksichtigung in der Auswertung liegt für die Gesamtpopulation zu Monat 12 unter 30 % und ist zwischen den Behandlungsgruppen nicht relevant verschieden (Unterschied in den Nichtberücksichtigungsanteilen < 10 %). Dies gilt im gleichen Maß für die im Dossier dargestellte Subpopulation B. Jedoch liegen die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 in beiden Populationen bei über 30 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte

vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Läsionslast* für die Auswertung zu Monat 12 als niedrig bewertet. Für die Auswertung zu Monat 24 lagen die Nichtberücksichtigungsanteile bei 40 %, hier ist mit einem hohen Verzerrungspotenzial zu rechnen.

Endpunkt: Hirnvolumen

1. Verblindung der Endpunkterheber

☒ **ja** ☐ **unklar** ☐ **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die MRT-Scans wurden zentral von einem verblindeten MRT-Auswerter auf Qualität. Vollständigkeit und Einhaltung der MRT-Vorgaben geprüft und anschließend ausgewertet. Eine unvollständige oder inkorrekte MRT-Aufnahme wurde wiederholt. Die Auswertung der Scans erfolgte nach einer standardisierten Prozedur.

2. Adäquate Umsetzung des ITT-Prinzips

☒ **ja** ☐ **unklar** ☐ **nein**

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

Der Anteil der Patienten ohne Berücksichtigung in der Auswertung liegt für die Gesamtpopulation zu Monat 12 unter 30 % und ist zwischen den Behandlungsgruppen nicht relevant verschieden (Unterschied in den Nichtberücksichtigungsanteilen < 10 %). Dies gilt im gleichen Maß für die im Dossier dargestellte Subpopulation B. Jedoch liegen die Nichtberücksichtigungsanteile für die Auswertungen zu Monat 24 in beiden Populationen bei über 30 %.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

☒ **ja** ☐ **unklar** ☐ **nein**

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

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

☒ **ja** ☐ **nein**

Angaben 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:

Das Verzerrungspotenzial der Studie wird insgesamt als niedrig bewertet. Da die Endpunkterheber verblindet waren, das ITT-Prinzip adäquat umgesetzt wurde und weder Hinweise auf eine ergebnisgesteuerte Berichterstattung noch sonstige das Verzerrungspotenzial beeinflussende Punkte

vorliegen, wird auch das Verzerrungspotenzial des Endpunkts *Hirnvolumen* für die Auswertung zu Monat 12 als niedrig bewertet. Für die Auswertung zu Monat 24 lagen die Nichtberücksichtigungsanteile bei über 50 %, hier ist mit einem hohen Verzerrungspotenzial zu rechnen.

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

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

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

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

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

☐ **nein**: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.
→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

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

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

☐ **nein**: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

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

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

☐ **ja**: Die Gruppen wurden zeitlich parallel verfolgt.

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

☐ **nein**: Die Gruppen wurden nicht zeitlich parallel verfolgt.

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

2.

für randomisierte Studien:

Verdeckung der Gruppenzuteilung („allocation concealment“)

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

☐ **unklar**: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau.

☐ **nein**: Die Gruppenzuteilung erfolgte nicht verdeckt.

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

für nicht randomisierte Studien:

Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

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

☐ **unklar**: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau.

☐ **nein**: Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.

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

3. Verblindung von Patienten und behandelnden Personen

Patient:

☐ **ja:** Die Patienten waren verblindet.

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

☐ **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

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

☐ **unklar:** Es finden sich keine diesbezüglichen Angaben.

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

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

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

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

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

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:
- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*
  Zulässige Gründe sind:
  - *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
  - *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
  - *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche" Endpunkte nicht berichtet sind.*

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

☐ **ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

☐ **unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

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

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

---

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

z. B.
- *zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien*
- *intransparenter Patientenfluss*
- *Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:*
  - *Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).*
  - *Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein.*
  - *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 Patientenanzahlen 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: